TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Dienstleistungs- und Technologiemarketing

Social Networks and Online Communities – Managing User Acquisition, Activation and Retention

Patrick Rohrmeier

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften
der Technischen Universität München zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr. Isabell M. Welpe

Prüfer der Dissertation:

1. Univ.- Prof. Dr. Florian von Wangenheim
2. Prof. Dr. René Algesheimer,
   Universität Zürich / Schweiz

Die Dissertation wurde am 27.12.2011 bei der Technischen Universität München eingereicht
und durch die Fakultät für Wirtschaftswissenschaften am 02.05.2012 angenommen.

Summary

Online social communities are among the most visited websites on the Internet. Although there has been enormous growth in the past years, these online venues operate in an increasingly saturated market with fierce competition. Community operators fight for active users, who interact and share content with other users. In order to ensure long-term success, community operators need to keep a sufficient number of active users in their community. Thereby, community managers are concerned with three challenges: the acquisition of new users, the activation of their members (making them contribute more), and the retention of the existing users.

Despite past research on online social communities and the broad interest in social influence across different disciplines, surprisingly little is known about how the users' social structural context affects participation in online communities. Because users share relationships with each other and interact in the community, it is of specific interest to understand how these relationships and the position of the users in the overall online network influence their behavior. Therefore, the present dissertation provides an empirical investigation of the effects of the social context on the users' perceptions, active participation behavior and loyalty to the online community. Because managing user acquisition, activation and retention is an important task for community managers, these different facets are examined in three empirical studies. *Study 1* investigates how users coming from different interpersonal acquisition channels differ in their attitudes towards the community and their behavior. *Study 2* observes the effect of the users' social structural context, attitudes and motivations on their active participation in the community. *Study 3* is concerned with the effect of the user's position and engagement in the online network on the decision to leave the online community.

Study 1 demonstrates that users coming from different interpersonal acquisition channels differ in their attitudes and behavior. The study reveals that receivers of word-of-mouth show more favorable attitudes and behavior towards the online community than users acquired by firm-initiated personal selling activities. Word-of-mouth referred users can identify more with the community, provide more recommendations to other people to use the community and show higher participation behavior on the platform than those coming from personal selling. Moreover, a differentiation in online and offline word-of-mouth recommendations reveals that

offline-referred users are more beneficial for the online community than online-referred users. Offline-referred users are more satisfied with the community, provide more offline recommendations and participate more compared to users who received online referrals. The social relationship between the sender and the receiver of the marketing message is of particular relevance to understand these effects. Nevertheless, the study also shows that these channel differences are partly explained by the users' different motives to use the community. Word-of-mouth referred users in general, and offline-referred users in particular, are more driven by social needs. This affects their attitudes and behavior towards the platform. This knowledge on interpersonal acquisition channels helps community operators to better understand what kinds of users are attracted and how to use different acquisition channels to build a balanced customer portfolio.

Study 2 provides insights on the drivers of active user participation, which is described by the active contribution to the community and interaction with other users. Based on social capital theory, it is demonstrated that both structural and attitudinal factors influence user behavior. The position of the user in the network, for example how central he is to other users, affects his active participation. Users with many contacts and users whose contacts do not share many relationships with each other participate more. In addition, meeting many of the online contacts in the offline world leads to less participation. Further, attitudinal factors yield additional insights, influencing participation in combination with objective network data. Here, satisfaction and identification with the community are found to be important drivers of active user behavior. Importantly, the study also shows that users with different motivations are affected differently in their participation. It is therefore critical to understand what types of users exist in the community and how they are motivated to use the platform. Overall, it is observed that the individuals' network structures play an important role in explaining participation behavior across different user groups.

Study 3 investigates user retention. Because the online community market becomes more saturated, it is more important for community managers to retain their users to keep a critical mass using the community. This study demonstrates that the user's network position, the configuration of the user's network of contacts and the engagement of the user in the online community influence the likelihood of defection. The results show that users with a central position and a close network of friends stay longer in the community. In addition, a lower share of contacts who already left the platform and a lower number of contacts from the same region and of the same gender lead to a higher likelihood that the focal user stays in the online social community. The users' active participation and being a verified member also affect loyalty. These are important insights for community managers because they help to assess the user's risk of leaving the community and take measures to keep them using the platform. Interestingly, some of the effects decrease over time, which has important implications for community management and the timing of marketing actions.

Overall, the present dissertation shows that the social context and the relationships to other users of the online network contribute to a better understanding of user behavior. The results suggest that community operators make use of structural and attitudinal data in order to build a strategic online community management. Based on this information, specific marketing measures should be developed to acquire, activate and retain users. First, users should be stimulated to provide word-of-mouth recommendations to other people. Nevertheless, community managers should use different interpersonal acquisition channels to attract a balanced customer portfolio consisting of different types of users in order to reduce the risk of being dependent on just one user group. Second, community managers should use structural and attitudinal data to personalize information for individual users on the platform and help them to integrate better in the online network. This data also helps to identify the most important users and evaluate the risk of leaving the community.

Theoretically, this dissertation underlines the value of social relationships for user behavior. Social capital and related social concepts can help to better understand both structural and attitudinal aspects in online social communities. The different facets and dimensions of social relationships and social capital, which are investigated in the three empirical studies, are linked to develop an integrated understanding of how social capital and user behavior works in online social communities.

Acknowledgements

The work at hand was accepted as doctoral thesis by the Technische Universität München in spring 2012. It was a long journey until I held the final version of this dissertation in my hands. However, it was a challenging, enriching and fun time thanks to all my mentors, peers, colleagues and research partners. I had the pleasure to work together with skilled and helpful people, always good for advice and deep topical discussions.

First, I would like to thank my PhD advisor, Prof. Florian von Wangenheim, for accepting me as a doctoral candidate at the chair for service and technology marketing. He continuously provided constructive support by challenging my ideas and opening up new ways to think about scientific issues. His timely feedback was valuable throughout all phases of the dissertation. Without his advice, the dissertation would not be what it is today. I also want to thank Prof. René Algesheimer, my second supervisor, for fruitful discussions and comments, his great expertise and support in bringing the empirical and theoretical work one step further. Our first meeting helped to shape the focus of my dissertation and his continuous feedback always helped to improve my thesis. Many thanks go to Prof. Isabell Welpe, who agreed to be the chairperson of my dissertation committee.

In addition, the whole team of the service and technology marketing chair at the TUM was just great. They were always there to have discussions on research topics and to have fun in the breaks in between. I am very grateful that I could work together with the best research partner I can imagine – thank you Ralf, Alexej, Stefan and Andreas for your unfailing efforts, the productive discussions and passion to support my dissertation.

I also want to thank my friends and colleagues for their feedback on my first versions of the dissertation: Dana, Lucie, Michael, Stefan and Tracie. Without you, I would not have finished in time. Of course, my whole family was important to me during this time, as they motivated me to pursue my doctoral thesis and to do always my best. Also, thanks to my employer, Booz & Company, for giving me the chance to take a break from daily consulting life and to work on my topic of passion.

Most importantly, my special thanks go to my girlfriend Nadine, who supported me with infinite patience, advice and love. It was the time we had together in Munich, which made me enjoy the dissertation even more.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADF	Asymptotically Distribution-Free
AIC	Akaike Information Criterion
AMOS	Analysis of Moment Structures (Statistical Software Package)
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AVE	Average Variance Extracted
B2B	Business-to-Business
B2C	Business-to-Consumer
BCa CI	Bias-corrected and accelerated Confidence Interval
BIC	Bayesian Information Criterion
cf.	confer (“compare”)
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Confidence Interval
Coef.	Coefficient
CR	Composite Reliability
e.g.	for example (exempli gratia)
et al.	et alii
eWOM	electronic Word-of-Mouth
HR	Hazard Ration
i.e.	that is (id est)
IAB	Interactive Advertising Bureau
MANCOVA	Multivariate Analysis of Covariance
MANOVA	Multivariate Analysis of Variance
NB	Negative Binomial
NBRM	Negative Binomial Regression Model
NFI	Normed Fit Index
OLS	Ordinary Least Squares
PH	Proportional Hazard
PRM	Poisson Regression Model
PS	Personal Selling

RMSEA	Root Mean Square of Error Approximation
S.E. / Std.Err.	Standard Error
SEM	Structural Equation Modeling
SRMR	Standardized Root Mean Squared Residual
Std.Dev.	Standard Deviation
US	United States (of America)
VIF	Variance inflation factor
WOM	Word-of-Mouth
ZINB	Zero-inflated negative binomial
ZIP	Zero-inflated Poison

1 Introduction

1.1 Relevance of Online Social Communities for Marketing

In the past years, marketing transformed significantly. Although not a new medium, the Internet gained more attention from marketers as its usage further increased and users became more engaged in the online environment. Internet advertising spending increased by 15 % from 2009 to 2010 to reach US$ 26.0 billion in the US, surpassing newspaper advertising spending to become the second largest advertising media behind television (IAB 2011). This highlights the growing importance of the Internet as a marketing and communications channel for companies and organizations. Most interestingly, the former one-to-many communication of the marketers changed into a many-to-many communication where users are becoming more and more active in their behavior on the World Wide Web. Thereby social media plays an important role, which takes on many different forms including weblogs, microblogging, Internet forums, wikis, recommendation platforms, social bookmarking, or online social communities. For such services, user-generated content is a key ingredient, which increasingly attracts the attention of marketers and consumers. A Forrester Research study (Bernoff 2010) shows that 83 % of the US online adults are already classified as active social media users, who create content, use social networking sites, post rankings and comments, tag and vote for websites, or consume social media content.

Among different social media channels, online social communities gained vast attention in recent years. They can be described as social aggregations of people who form personal relationships online through communication and interaction, often using public or semi-public user profiles (Boyd and Ellison 2008; Rheingold 1993).[1] Online social communities significantly increased in popularity, where Facebook is today's largest social networking site on the Internet with more than 800 million members, who each produce more then 90 pieces of content on average per month (Facebook 2011). Besides Facebook, there are hundreds of other online communities in the market. According to Hampton et al. (2011), nearly half of all American adults use at least one social networking site, with a growing trend. On such sites, users can share their knowledge, experiences, information, and digital goods with their peers through an easy-to-use online interface.

[1] The terms online communities, virtual communities, online social communities, and social networking sites are used synonymously, although there might be some differences in definition by different authors (cf. chapter 2).

This development provides great potential for marketers and online community operators to gain value from these social conglomerations of people. In fact, online communities can be used to directly and indirectly generate revenue in many different ways. For example, community operators can monetize their members' attention and behavior through advertising, membership fees, or selling transactions, while companies can use such platforms to gain customer insights, increase organizational efficiency through knowledge transfer, or increase customer loyalty towards products and brands (e.g., Algesheimer, Dholakia, and Herrmann 2005; Clemons 2009). In order to capture the economic potential and to be successful in an increasingly competitive marketplace, an online community must remain attractive for its members and partners. Therefore, community operators need to keep a critical mass of active users and ensure that new content is generated continuously and enough communication takes place within the community.

This dissertation has the objective of providing insights for online community operators on how to become more successful, by answering the following questions: Which acquisition channels work well in acquiring new users? What makes their users contribute more content on the platform? What keeps their users remaining in the community for a longer period of time? All three aspects are of significant value for the community operator, third parties who are interested in advertising or transactions, and also for the community users themselves, as this keeps their community alive. In the remainder of this introductory chapter, the objectives and research questions are addressed, the relevance of this study is discussed, and an outline of the dissertation is presented.

1.2 Objectives and Research Questions

In the past 20 years, researchers from different disciplines have investigated the effects, dynamics and value of online social communities for users, operators and companies. A multitude of scientific and practical work emerged in fields like marketing (e.g., Dholakia, Bagozzi, and Pearo 2004; Woisetschlaeger, Hartleb, and Blut 2008), management (e.g., Armstrong and Hagel 1996), social-psychology (e.g., Lin 2006), sociology (e.g., Wellman et al. 1996), or information systems (e.g., Wasko and Faraj 2005). In marketing, the social context in which customers act has become more important in recent years, which is promoted by the network approach of marketing (Achrol and Kotler 1999; Algesheimer and Wangenheim 2006). Online social communities provide a rich social environment because users are embedded in a network of social contacts, which is often visible through friend lists and the explicit presentation of contacts on user profiles. In order to understand the key drivers of an online community's success, community operators must take the social context and its influence for their individual users into consideration.

In fact, social network analysis gained increasing popularity in marketing to explain social behavior and its effects (e.g., Brown and Reingen 1987; Katona, Zubcsek, and Sarvary 2011). As users connect and interact with each other in online communities, the users' location within the network and their relationship to the community are expected to influence their behavior (Tsai and Ghoshal 1998). Despite the prominence of the network perspective in marketing and the growing popularity of online communities as a research object to investigate customer behavior, only few studies integrated the effects of the user's network and its social influence to investigate user behavior in online communities. In particular, recent research incorporated a user's network structure to explain adoption behavior (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011), the influence of a user's log-in behavior on other user's activity (Trusov, Bodapati, and Bucklin 2010), and the overall success of open source software communities (Toral et al. 2009). Although these research studies provide valuable insights to our understanding of the effects of social influence within online communities, many knowledge gaps still exist.

This dissertation has the objective of gaining additional insights on how social aspects and connections between users in online social communities affect important factors of success for community operators. To do this, three basic membership stages are identified, on which community operators need to take action. First, community operators need to attract new users. This means, using different channels to contact and convince people to join the community. Customer acquisition should thereby function as an important process that should bring the "right" customers to the firm (Blattberg and Deighton 1996; Hansotia and Wang 1997). Second, after the users sign up for the online community, it is important to make them contribute to the platform and interact with other users. Providing enough new content and interaction to keep the current user base interested in the community should be a central goal. Third, because of an increasing saturation of the online community market, it is crucial to keep the existing users. Retention of community members fosters the existence of a critical mass and sufficient traffic on the platform.

The objective is to investigate the effect of the social context on user perceptions and behavior in online social communities. Therefore, this dissertation provides empirical evidence on influencing factors of attitudinal and behavioral outcomes, as well as recommendations on how to establish an active online community environment.[2] Because community management deals with users in different phases of their life cycle, three empirical studies are conducted, each addressing one stage of the generic membership development process[3] mentioned above: user acquisition, user activation, and user retention.

1.2.1 The Acquisition of Users in Online Social Communities

For online communities, it is important to maintain a sufficient number of users to keep the community alive and ensure interaction between its users. In order to reach this goal, new users must be acquired not only during the growth phase of an online community, but also in a more mature state, because existing users stop using the community and new users can provide additional content. Therefore, an online community's goal should be to continuously acquire new members. To achieve this, several marketing communication channels exist (e.g., Borden 1964; Chen and Xie 2008; Duncan and Moriarty 1998), through which an online community can potentially attract new users.

Specifically, interpersonal communication is an effective means for acquiring new customers and especially important for interactive online services. Past research has already shown that word-of-mouth (WOM) is more effective than traditional marketing channels (e.g., Herr, Kardes, and Kim 1991; Trusov, Bucklin, and Pauwels 2009). This dissertation calls for a more differentiated view. Considering the social context of interpersonal communications, this study compares receivers of WOM referrals with users acquired via the personal selling channel. As far as is known from past research, only the different effects of WOM recommendations and personal selling concerning the decision making process have been investigated (e.g., Katz and Lazarsfeld 1955), and the effectiveness of WOM and other marketing channels in general have been compared (e.g., Schmitt, Skiera, and van den Bulte 2011; Villanueva, Yoo, and Hanssens 2008). Therefore, this study provides a new perspective on how customer- and firm-initiated communications differ. Further, a distinction of WOM in offline and online referrals is proposed. By this, post-adoption behavior and attitudes for online- and offline-referred users are compared. This comparison has received no academic attention so far. In addition, customer motivations for participating in the online service are investigated, in order to take channel-specific user characteristics into account. The motives

[2] The empirical studies are presented in chapters 5, 6 and 7.
[3] The membership development process and the users' roles in their life cycle are described in more detail in chapter 2.4.3.

to use the online service are compared by interpersonal acquisition channel and are tested for the impact of motivations as mediating factors between acquisition channel and the users' perceptions and behavior. The results provide valuable insights on how communication channels differ and what kinds of users, in terms of their attitudes and behavior towards the online community, are attracted.

Altogether, three research questions will be answered: (1) Do WOM-referred users differ in their post-adoption attitudes and behavior compared to users acquired by personal selling? (2) Do offline-referred users differ in their post-adoption attitudes and behavior compared to online-referred users? (3) Do users from different acquisition channels differ in their motivation to use the online community, and does this motivation mediate the effects on user attitudes and behavior?

1.2.2 The Activation of Users in Online Social Communities

After the users have been acquired by the online community, it is important to facilitate their participation within the community. Active user participation is the central element for an online community's success, independent of the community's business model and orientation. From a marketing perspective, community participation can significantly impact the customers' loyalty and commitment to products and brands through brand and consumption communities (e.g., Algesheimer, Dholakia, and Herrmann 2005; McAlexander, Schouten, and Koenig 2002; Woisetschlaeger, Hartleb, and Blut 2008) or create value for other members through user conversations and product reviews (e.g., Godes and Mayzlin 2004; Hennig-Thurau et al. 2004; Nambisan and Baron 2007). For many online communities, member participation is not only a means to strengthen a brand or create content in a cost-efficient way, it is also an important outcome itself, as it affects revenues through advertising (Clemons 2009; Trusov, Bodapati, and Bucklin 2010). Therefore, managing participation behavior is an important marketing objective.

Despite the fact that social connections are the foundation of online communities, an in-depth analysis of the effects of the users' positions in the network on their participation behavior is still missing. This study investigates the effect of online social structures and user attitudes on participation behavior. Although recent research utilized actor networks to identify the influence of users on other members in the community and to examine diffusion in social networks (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Trusov, Bodapati, and Bucklin 2010), there is a need to better understand the impact on an individuals' active community participation. Centrality measures, ego-network density and the network's online-offline configuration are included in the analysis, which have not been used to get insights on active user participation so far. Further, the value of additional attitudinal factors for explaining active participation behavior in combination with objective network data is investigated. Thereby, social capital theory is used as an adequate frame to study the influence of these structural and attitudinal factors. The effects on participation are compared for two user groups based on their motivations to participate. This differentiation is expected to provide further insights on how user groups need to be treated according to their needs.

Based on the relevance of online community participation, three main research questions are addressed: (1) What are the structural drivers for active user participation? (2) How do attitudinal drivers affect active user participation in the presence of structural drivers? (3) How does user motivation affect the relationship of the structural and attitudinal factors on active user participation?

1.2.3 The Retention of Users in Online Social Communities

Registered members use the online community as long as their perceived benefits of the membership are higher than their costs. Because there are many other online communities in the market, built around different topics and geographies, they compete for the users' attention. Though there has been enormous growth in the past years, this development is slowing down, where the trend shows some social media fatigue among certain user groups of such sites (Gartner 2011). Within a more saturated market it is important for community managers to retain their users, thus keeping a critical mass using the community. Therefore, community operators need to know what influences user defections.

Recent research emphasized the importance of social influence on adoption and retention behavior (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Nitzan and Libai 2011). However, past research lacks empirical evidence of why online community users leave the platform. This study is the first to investigate the impact of dynamic social structures on user defection in the context of online community services. Based on social theories, the effects of online social structures and community engagement are investigated. Specifically, the users' position in the network, the configuration of the users' current network, and the participation and engagement in the community are tested for their effect on user defection. An important finding is that effects can change over time, thus these changes are given explicit attention.

Getting a better understanding about user defection, three main research questions are addressed: (1) What are the structural drivers for members to defect from the online community? (2) How does community engagement affect the members' defection from the online community? (3) How do the effects on user defection change over time?

Overall, in order to keep an online community active and attractive for its users, community providers need to concentrate on all three stages – acquisition, activation and retention. All three studies are centered around the influence of the individual users' network positions and their social context within the online community. Thereby, all three contribute to a better understanding of the importance of social relationships and how the community operator can use this knowledge to manage the community in a way to make it more successful. Failing at a single point can have significant impact on the overall performance of the community, and therefore lead to decreasing value. For example, if the users in the community are less active, this results in less content and interaction on the platform. Consequently, more members will be dissatisfied and become completely inactive, not using the platform anymore. In addition, less content contributed can also decrease the attractiveness of the online community for new users, which hinders prospective users to join the community. This illustrates that community operators continuously need to track and manage the activity of their community. The different stages are related to one another, so that success depends on the entire membership development process.

1.3 Structure and Approach of the Dissertation

Because of the importance of user acquisition, user activation and user retention for long term community success, the research study provides different perspectives on the membership development process through the analysis of these three topics. Multiple theoretical concepts are used to facilitate the development of the research hypotheses. Most importantly, social concepts help to understand the interrelations of the observed drivers and outcomes in the online community. The empirical studies are conducted in order to test the theoretically derived hypotheses. In this way, insights are generated on the three generic stages of the process of community building and maintenance for marketing managers of online social communities. The dissertation is structured as follows:

Because of the large academic attention in recent years, there exists a plethora of relevant research on online social communities. This research covers a wide spectrum of topics. One important aspect is the participation of users in such communities, its influencing factors and consequences. In **section 2** a basic understanding will be developed about what online social communities are, how users participate in the communities, and which relevant antecedents and consequences of participation have been identified in prior research. Therefore, online social communities are defined, different types of communities and different types of members are described, and the membership development process and different roles in the users' life cycle are introduced. At the end of this section an overview of the most relevant factors for community participation is provided.

Section 3 includes the theoretical background for the research questions. Relevant concepts of social theories are introduced for a better understanding of the drivers of community success. First, an overview of interpersonal communication in marketing research is given. This is particularly relevant for the user acquisition study. Second, important social network concepts are described. Here, the increased popularity of social network analysis underlines the value of such analytical tools for understanding individual behavior in interactive social contexts, like online communities. Further, an introduction to social capital theory is given. In social sciences, social capital theory has gained significant attention, as it helps to understand what drives the value of relationships between actors within certain social environments, like companies, inter-firm networks, virtual groups, or online communities. As online communities share the characteristic of being highly connective and being based on social relationships, it has an inherent social value for the community. Additional social concepts, including social identity, social exchange, social presence and collective behavior, are introduced, which add to the theoretical foundation of the empirical studies. Overall, different social theoretical concepts are discussed, which provide a basis for a better understanding of user behavior in online social communities.

After the theoretical basis has been laid out, the three empirical studies are described. The empirical work is based on data from a local online social community. The research design, the research object and the data used in the empirical studies are explained in **section 4**.

Sections 5, 6 and 7 describe different perspectives on user attitudes and participation, which contribute to each of the three stages – acquisition, activation, and retention. First, the effect of different acquisition channels on the post-adoption attitudes and behavior of the users is tested in **section 5**. Second, an empirical model to identify the most important predictors of active user participation in the online social community is investigated in **section 6**.

And third, factors influencing user defection are observed in **section 7**. All studies base the analysis on the social context in which the individual users are acting. In each section, the context of the respective empirical study is described and the results are presented. Each section includes the relevance of the empirical analysis for marketing and online community research, links the research focus to the theory presented in section 3, theoretically develops propositions on the influencing factors of user attitudes and participation in the online community, and tests these propositions by means of adequate empirical models. The results are presented and the implications and limitations of the study are discussed.

Section 8 includes a summary of the results. The results from the three empirical studies are integrated into an overall discussion of the findings. Prospects on future research are given at the end of this section.

2. Online Social Communities
- Description and Definition
- Types and Classification
- Community Value for Marketing
- User Benefits and User Types
- Online Community Participation

3. Social Context of Communication and Relationships - Theoretical Perspective
- Interpersonal Communication Channels
- Social Networks
- Social Capital Theory
- Further Social Concepts

4. Three Empirical Studies on Network Structure, Attitudes and Behavior in Online Social Communities
- Objectives and Research Approach
- Research Design and Data Collection

5. User Acquisition: A Comparison of Inter-personal Acquisition Channels for Online Social Communities
- Relevance
- Theoretical Background and Hypotheses
- Study Description
- Statistical Model and Results
- Discussion, Implications, and Limitations

6. User Activation: The Effects of Individual Network Structure, Attitudes and Motivation on User Participation
- Relevance
- Theoretical Background and Hypotheses
- Study Description
- Statistical Model and Results
- Discussion, Implications, and Limitations

7. User Retention: The Effects of Social Structure and Engagement on Defection from Online Social Communities
- Relevance
- Theoretical Background and Hypotheses
- Study Description
- Statistical Model and Results
- Discussion, Implications, and Limitations

8. Conclusion – The Influence of Acquisition, Structure, and Attitudes on User Behavior in Online Social Communities
- Summary of Empirical Studies
- General Discussion of the Results
- Future Research Directions

Figure 1: Structure of the Dissertation

2 Online Social Communities

Online communities and social networking sites gained broad attention from practitioners and scholars in recent years. The usage of these online venues has experienced rapid growth in the first decade of the new century. For example, Facebook is today's largest community site in the World Wide Web with more than 800 million members (Facebook 2011). Sites like Facebook or LinkedIn are among the most frequently visited websites on the Internet (Alexa Internet 2011). Thus, they play a significant role in the daily (Internet) life of many people. Today, nearly half of all American adults are members of at least one online social community (Hampton et al. 2011). Because of their high reach and growing importance, it is not surprising that scholars from different disciplines become interested in online communities as a research topic. In fact, the literature on online communities increased dramatically, comprising over 300 publications in the past 15 years (Laine 2009). Because of the multi-disciplinary interest and the multitude of different online communities in the marketplace, it is important to have a common understanding of what online communities are. This chapter provides a definition of online communities, discusses different types of communities and roles of members along their life cycle to better understand this phenomenon. Further, a broad literature review of the constituents, antecedents and consequences of online community participation is given.

2.1 Description and Definitions of Online Social Communities

Online communities are not a new phenomenon. Long before social networking sites like Facebook or MySpace emerged, early forms of online communities existed. Back in the 1960s and 1970s first computer networks were established to facilitate the connection between geographically dispersed people (Wellman et al. 1996). Over the years, different computer communication systems have been developed to interconnect users. For example, Usenet newsgroups, bulletin board systems, E-mail lists or chat systems have been used as some kinds of virtual communities to exchange information and connect people electronically

(Hauben and Hauben 1997; Rheingold 1993).[4] Although different computer-supported social networks previously existed, the research stream of online (or virtual) communities started to develop in the early 1990s. Howard Rheingold (1993, p. 5) was one of the first authors to form the term “virtual communities”, defining them as “social aggregations that emerge from the Net when enough people carry on public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace”. The most recent forms of online communities are social networking sites. They can be described as “web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.” (Boyd and Ellison 2008, p. 211). The difference between the definitions of online communities by Rheingold (1993) and social networking sites by Boyd and Ellison (2008) shows that definitions depend on the focus of research and the orientation of the online community, sometimes resulting in little overlap between definitions. With the growth of online communities, many scholars from different disciplines began to investigate various online social gatherings under the label of online communities. Because of the different research perspectives and the various types of online groups observed, a versatile set of terms and definitions emerged. With respect to the focus of research, different dimensions and specific aspects of online communities are more highlighted than others. In past literature, no consistent definition of online communities can be found. Therefore, several definitions are compared in this section to identify the most important dimensions and achieve a common understanding of online communities.

In fact, there exists a multitude of definitions of online communities in research. Table 1 gives an overview of relevant definitions of online communities and social networking sites. This list of definitions is neither exclusive nor exhaustive. Moreover, it should provide different perspectives on how online communities can be understood. A review of these selected definitions reveals the most important dimensions of an online social community, which are shared by several authors. These include user relationships, social interaction, common interests, and virtual environments.

[4] This dissertation does not claim to exhaustively discuss the history of online communities. For an overview of the historical development of computer networks, online communities and social networking sites see for example Boyd and Ellison (2008), Hauben and Hauben (1997), Rheingold (1993), and Wellman et al. (1996).

Rheingold (1993, p. 5)	"Virtual communities are social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace."
Hagel and Armstrong (1997, p. 143)	"Virtual communities are groups of people with common interests and needs who come together on line. Most are drawn by the opportunity to share a sense of community with like-minded strangers, regardless of where they live. But virtual communities are more than just a social phenomenon. What starts off as a group drawn together by common interests ends up as a group with a critical mass of purchasing power, partly thanks to the fact that communities allow members to exchange information on such things as a product's price and quality."
Kozinets (1999, p. 254)	"Virtual communities of consumption' are a specific subgroup of virtual communities that explicitly center upon consumption-related interests. They can be defined as 'affiliative groups whose online interactions are based upon shared enthusiasm for, and knowledge of, a specific consumption activity or related group of activities."
Preece (2000, p. 10)	"An online community consists of: - People, who interact socially as they strive to satisfy their own needs or perform special roles, such as leading or moderating. - A shared purpose, such as an interest, need, information exchange, or service that provides a reason for the community. - Policies, in the form of tacit assumptions, rituals, protocols, rules, and laws that guide people's interactions. - Computer systems, to support and mediate social interaction and facilitate a sense of togetherness."
Balasubra-manian and Mahjan (2001, p. 108)	"[…] we define a virtual community (in a relatively neutral way) as any entity that exhibits all of the following characteristics: 1. It is constituted by an aggregation of people. 2. Its constituents are rational utility-maximizers. 3. Its constituents interact with one other without physical collocation, but not every constituent necessarily interacts with every other constituent. 4. Its constituents are engaged in a (broadly defined) social-exchange process that includes mutual production and consumption (e.g., mutual dissemination and perusal of thoughts and opinions). While each of its constituents is engaged in some level of consumption, not all of them are necessarily engaged in production. Such social exchange (as opposed to monetary or material exchange) is a necessary, but not always the only, component of interaction between the constituents of the entity. 5. The social interaction between constituents revolves around a well-understood focus that comprises a shared objective (e.g., environmental protection), a shared property/identity (e.g., a national culture or a lifestyle choice), or a shared interest (e.g., a hobby)."

Table 1: Overview of Different Definitions of Online Social Communities

Bagozzi and Dholakia (2002, p. 3)	“We view virtual communities to be mediated social spaces in the digital environment that allow groups to form and be sustained primarily through ongoing communication processes.”
Ridings, Gefen and Arinze (2002, p. 273)	“Virtual communities can be defined as groups of people with common interests and practices that communicate regularly and for some duration in an organized way over the Internet through a common location or mechanism.”
Wiertz and de Ruyter (2007, p. 349)	“We define commercial online communities as firm-hosted online aggregations of customers who collectively co-produce and consume content about a commercial activity that is central to their interest by exchanging intangible resources.”
Boyd and Ellison (2008, p. 211)	“We define social network sites as web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.”

Table 1: Overview of Different Definitions of Online Social Communities (continued)

User Relationships. The central aspect of an online community is the social relationship between its members. Rheingold (1993) already emphasized that virtual communities are “social aggregations” on the Net and form personal relationships. With regards to a more general definition of communities as “[…] networks of interpersonal ties that provide sociability, support, information, a sense of belonging, and social identity” (Wellman 2001, p. 228), it becomes clear that it is the social connection that constitutes a community, be it either physically or in the virtual space. Those relationships build the social context for interaction between the users. In fact, most definitions considered in Table 1 explicitly describe the social connection between people in online communities as a main element.

Social Interaction. Closely related to the relationship between members is the aspect that these members interact and communicate with each other in the community, thereby exchanging ideas, information and knowledge. It is the ongoing communication that forms and maintains these relationships (Bagozzi and Dholakia 2002). The members of an online community are engaged in social-exchange processes, but not all members are necessarily engaged in content production (Balasubramanian and Mahjan 2001). Thus, online communities need social interaction between their members, which does not always include constant activity of all members, but can also occur through information consumption.

Common Interests. A specific orientation and topic of the community facilitates users to share certain interests and feelings. On this way, online communities bring together people and make them form relationships based on their common interests (e.g., Hagel and Armstrong 1997; Ridings, Gefen, and Arinze 2002). The community's purpose and the associated common interests of its users give the reason for the community to exist (Preece 2000). Often there is a theme which constitutes the community, for example sports, travel, jobs, or motherhood. Thus, interaction between members is often related to information exchange on particular topics like exchanging recipes and cooking tips in a community about culinary matters (e.g., de Valck et al. 2007) or problem-solving for specific services and products (e.g., Dholakia et al. 2009).

Virtual Environment. As the name "online" community already indicates, the formation of groups takes place in the online domain. Although social ties from online and offline contexts often intertwine (Wellman and Hampton 1999), the relationship and interaction of its members occurs online. In online communities, both strangers and people one already knows can meet and interact. Even when people meet offline, the online community is referenced to their relationships and communications in the online world. Thereby, the online community website is the interface for users meeting in the virtual space. The web-based service and technology facilitates the online relationships and social interactions, and accommodates the virtual community of people (Boyd and Ellison 2008; Preece 2000).

Besides these four main elements of online communities, other aspects are included in some definitions. For example, Preece (2000) underlines the relevance of policies, rituals and rules. They are important to ensure appropriate behavior of its members in the community. Often such rules and rituals are developed as tacit norms over time. Other authors emphasize the commercial orientation of online communities. For example, Hagel and Armstrong (1997) directly relate online communities to economic outcomes. But as Kozinets (1999) notes, commercial communities, like communities of consumption, are interpreted as a subgroup of online communities. Therefore, the commercial purpose is not a constitutive element of an online community. It rather describes a certain type of online community.

The definitions also show how online communities developed over time. In the first online communities, like bulleting boards and Usenet groups, people simply exchanged knowledge and information on certain topics. In these early days most online communities used only text-based communication, though written word is still an important aspect of most newly emerged communities (Bagozzi and Dholakia 2002). In recent years, there has been a continuous development of online community and social network technology, which led to new functionalities allowing for an increased activity of users on the platform. The emergence of social networking sites provides extended functionality, like providing recommendations and rate products and services, play social games, or use status updates to express current feelings and opinions. As the definition of Boyd and Ellison (2008) points out, the functionality to maintain a personal user profile and the visibility of connections between people is one of the latest developments of online communities, which became standard in many cases. These developments do not undermine the original understanding of online communities, but rather emphasize its ability to connect people and pursue social interactions. At their core, both online communities and social networking sites still fulfill the basic purpose of enabling users to communicate and connect with each other, building personal networks, as well as sharing user-generated content (Enders et al. 2008). Thereby, social networking sites can be seen as a unique type of online communities, with similarities regarding their socializing aspects, but some dissimilarity in their motives and business models (e.g., Boyd and Ellison 2008; Trusov, Bucklin, and Pauwels 2009). Based on the different aspects and definitions, the general terms "online community" or "online social community" are mostly used in the remainder of this dissertation. Nevertheless, other terms like virtual community and social networking site are also used as synonyms when talking about online social communities in the following.

2.2 Types and Classification of Online Social Communities

The existence of various online community definitions is not surprising, because online communities can greatly differ in their purpose and objective. In the online community market, a diverse set of online community types can be found. Because online communities have been investigated from different research disciplines, various classification schemes on how to describe the different types of communities have been proposed. Therein, online community researchers set specific focuses in their choice of criteria to their classification approach.

Armstrong and Hagel (1996) provide one of the first typologies for online communities. They base their classification of online communities on the purpose for which they are organized and the needs which are addressed by each type. They define four types, which are designed towards customer-oriented communities: communities of interest, communities of relationship building, communities of transaction, and communities of fantasy. Communities of interest are formed by individuals who interact with each other based on some shared interest, expertise, or passion such as sports, entertainment, or traveling. The formation of personal connections is strongly developed in communities of relationships, which are formed by individuals with a need to come together and share life experiences. Communities of transaction facilitate economic exchanges through information sharing related to those transactions. Communities of fantasy provide people with the opportunity to develop environments and personalities in imaginary worlds of fantasy. All four types are directed towards the fulfillment of specific user needs. However, for online communities it is important to enable users to satisfy multiple needs (Hagel and Armstrong 2006). For example, a community about motherhood would first cover the interest in information about preparation and possible issues of becoming a mother. Additionally, the connection to other users can help solve problems experienced in different life phases, providing social support as a community of relationship. Further, the community can provide valuable information about child related products. Because online communities can cover more than one of these four community types, a clear distinction is hard to make. However, most online communities primarily focus on one of the four types.

Moreover, Hagel and Armstrong (1997) also propose a more business driven typology. Thereby, online communities can be differentiated into consumer-focused and business-to-business (B2B) communities. Differentiating further, consumer communities can be focused as geographic, demographic and thematic communities. B2B communities are further distinguished into vertical industry, functional, geographic and business category communities.

Another approach is proposed by Porter (2004). She first distinguishes between member-initiated and organization-sponsored communities. This view focuses on the establishment of the community. Further, with respect to the relationship orientation, member-initiated communities can be social or professional, while organization-sponsored communities are allocated to commercial, non-profit and government communities.

The major limitation of these typologies is that the categories do not clearly distinguish to each other, but instead online communities might frequently fall in more than one category. Some community types overlap and a hierarchical typology is not always applicable. For example in Hagel and Armstrong’s (1997) business driven typology, thematic orientation on motherhood includes a demographic focus on young women as well. In another example, organization sponsored commercial communities, like Facebook, can also facilitate social interaction. Because large social networking sites are organization sponsored a further differentiation of what they offer in their commercial communities is missing in the typology of Porter (2004). In this respect, a hierarchical clustering seems to be informative, but not always helpful.

Other researchers did not develop certain classification schemes, but defined specific types of communities according to their focus and the study context. For example, online communities are defined as communities of practice (Matthwick, Wiertz, and de Ruyter 2008), communities of consumption (Kozinets 1999), electronic knowledge repositories (Kankanhalli, Tan, and Wei 2005), brand communities (Algesheimer, Dholakia, and Herrmann 2005; Woisetschlaeger, Hartleb, and Blut 2008), problem-solving communities (Dholakia et al. 2009), online travel communities (Wang and Fesenmaier 2004a) or social networking sites (Trusov, Bucklin, and Pauwels 2009). All these various forms of online communities cover one or more user needs and are directed to a specific purpose or topic. Specific labels are useful to narrow down the focus of research, but they do not contribute significantly to form new typologies of online communities. For example, brand communities can be categorized as firm-initiated, commercial communities fulfilling the function of communities of interest and transaction. Social networking sites are also firm-initiated, commercial and facilitate relationships. At the same time they might also focus on specific target groups according to their age and geography. These examples illustrate that all together, the dimensions of different typologies can help to better understand the different contexts in which online communities work. For this reason, a more comprehensive framework of different dimensions is provided here, based on existing typologies as well as new categories defined from practical examples. Figure 2 provides a summary of the different dimensions.

Initiator. The initiator of a community can be an organization, which includes companies, non-profit organizations, and government organizations (Porter 2004). On the other hand, members or private persons could initiate communities based on their interests and ideas. Though the border between member-initiated and organization-sponsored communities can be blurry, as the community might originate as member-initiated, but then grows to a business in its own and therefore becoming company- or organization-sponsored, as is the case of Facebook[5].

[5] http://www.facebook.com

<table>
<tr><th>Dimension</th><th colspan="20">Category</th></tr>
<tr><td>Initiator</td><td colspan="10">Member-initiated</td><td colspan="10">Organization-sponsored</td></tr>
<tr><td>Living Environment</td><td colspan="10">Private Life</td><td colspan="10">Professional Life</td></tr>
<tr><td>Commercial Orientation</td><td colspan="10">Non-Commercial</td><td colspan="10">Commercial</td></tr>
<tr><td>Community Function</td><td colspan="10">Core Business</td><td colspan="10">Support Function</td></tr>
<tr><td>User Segment</td><td colspan="10">Consumer-oriented</td><td colspan="10">Business-to-business</td></tr>
<tr><td>Content Focus</td><td colspan="5">Geographic</td><td colspan="5">Demographic</td><td colspan="5">Thematic</td><td colspan="5">General</td></tr>
<tr><td>Needs</td><td colspan="4">Social (Relationship)</td><td colspan="4">Informational (Interest)</td><td colspan="4">Entertainment (Fantasy)</td><td colspan="4">Economic (Transaction)</td><td colspan="4">Other</td></tr>
<tr><td>Content</td><td colspan="10">User-generated content</td><td colspan="10">Hybrid content</td></tr>
<tr><td>Access</td><td colspan="10">Open</td><td colspan="10">Closed
(only with invitation)</td></tr>
</table>

Figure 2: Categorization Scheme for Online Communities

Living Environment. An intuitive distinction of communities can be made into professional or job-related communities and communities centered around the private or social life of its members (Porter 2004). Professional communities could be directed towards a specific profession, for example lawyers discussing job-related matters such as newest developments in tax law. More generally, professional communities also include business networks like LinkedIn[6] and Xing[7], which are focused on connecting business contacts. Social communities are established to connect people outside their profession. For instance, motherhood communities are clearly categorized as communities about one's private life. Although professional communities can become social and vice versa, the focus of most communities on this dimension is obvious.

Commercial Orientation. Online communities can be differentiated according to their commercial orientation, and can be either commercially driven or not. Most large social networking sites are clearly commercially driven as their objective is to achieve profit. On the other hand, non-commercial communities include examples like member-initiated brand communities, which are established based on their passion for a specific brand, like car or motorcycle brands. However, if such brand communities are firm-initiated by the owner of

[6] http://www.linkedin.com
[7] http://www.xing.com

the brand, they are commercially driven, because the firm at least wants to increase the loyalty and identification with the brand.

Community Function. The function of the community relates to its overall objective. When the online community is the core business, the community itself needs to reach certain economic objectives. For example, Facebook's social networking site is its core business and the company generates revenues through its community. Online communities as a support function are only part of the overall business model and support the core business of the community operator. For example, this is the case at ebay[8], where the community supports ebay users in using the auction website or building a stronger relationship to the brand.

User Segment. As suggested by Hagel and Armstrong (1997), online communities can be categorized as consumer-oriented and business-to-business. In the B2B case, a firm-initiated social network can be established for its business customers, providing a knowledge sharing platform for users about their products. For example IT-firms may offer communities where the users can exchange information on solving problems with the software. On the other hand, consumer-oriented communities, can, for example, be related to a company's products and brands, or to the mass market which offers social networking to everyone (e.g., Facebook).

Content Focus. Most communities focus on specific themes or target groups, who are differentiated along demographics, geographics, and topics (Hagel and Armstrong 1997). Demographics can relate to age groups or gender, for example, an online community for young women. Geographies can be broken down, for instance, by country, state, region, or city. Topics can range from motherhood to car brands, from sports to health issues, and include any niche topic one can consider. Though there are some communities which are rather of general nature like social networking sites. Facebook for example does not provide a specific topic orientation apart from 'networking with others' and can therefore be considered as a general community.

Needs. As already described above, Armstrong and Hagel (1996) classified different types of communities based on user needs. According to their classification, there are four main user needs that are relevant for customer-oriented online communities: information needs (communities of interest), social needs (communities of relationship building), economic

[8] http://hub.ebay.com/community

needs (communities of transaction), and self-exploration and entertainment needs (communities of fantasy). Besides those needs, there might be other motives which drive users to participate in online communities, for example the need for recognition or self-enhancement (e.g., Dholakia, Bagozzi, and Pearo 2004).[9]

Content. Customer participation in an online social community can be seen as a form of content co-production and can take different forms. Co-production on online platforms can range from firm production to joint production and solely customer production. In this context, the first refers to no customer involvement (e.g., traditional push-information websites without community function) and the last to fully user-generated content websites, with the level of participation depending on the strategy of the platform. Although online communities are mainly built to facilitate user interaction, there are many online platforms that provide a hybrid content offering. This means that users contribute their own content, interact, and share it with other users, but in addition the community operator itself provides editorial content to the platform. There are many examples of such hybrid-content online social communities. For example, news websites with community functionalities to comment and discuss topics. Sports websites providing a community to meet sports mates and make training appointments, but at the same time offering articles on effective training. Or communities for mommies, who discuss educational issues and get expert advice offered by the community operator.

Access. Online communities can be either open for registration to everyone or exclusive for people who receive invitations from other members or the operator. This could have implications on the growth of the platform when not everyone is initially able to subscribe.

The overview of the different dimensions (Figure 2) can help to define and classify a specific community based on its characteristics. The single categories of each dimension are not necessarily exclusive. They can overlap for specific communities, for example, when different needs (like social and informational) are addressed simultaneously by the online community. Further, the overview contains the most relevant dimensions and categories, but does not claim to be completely exhaustive. Nevertheless, the categorization scheme may help define a relevant market position for an online community. Operators can use the categorization to identify attractive niche markets and differentiate themselves in one or several dimensions against their competitors. For a competitor analysis this is an important step. From a user per-

[9] See also chapter 2.4.1 for an overview of relevant motives and needs in past online community research.

spective it may clarify the value proposition of the respective community and what benefits can be achieved by the use of a specific online community.

2.3 The Value of Online Social Communities from a Marketing Perspective

The importance of online social communities for marketers stems from the enormous popularity they have achieved to date. With millions of people active in online social communities, such as Facebook, MySpace, LinkedIn, or Xing, they offer access to a large group of members. Consequently, online social communities have a great potential to bring together customers and organizations. However, online communities changed the way of communication from a one-to-many communication, where marketers just try to reach a mass of customers, to a many-to-many communication, where customers regularly interact with each other (Cothrel 2000; Hofmann and Novak 1997; Rheingold 1993). Online social communities can therefore be seen as social networks with direct interaction between its users (n:n), while the traditional top-down communication (1:n) does not allow for direct interaction between the users on the platform (Weiber and Meyer 2002). Integrating online communities in the marketing activities of a firm innately changes the process of directing marketing measures toward customers. There exists different alternatives to capture the value of online social communities for marketing – some based on more traditional tools, and others opened up by new opportunities that occur from the interaction of firms with their customers and among members. The benefits of online social communities for firms are presented in this chapter; the benefits for the community members through need fulfillment are included in the next chapter.

The interest in online communities is grounded in the opportunity for operators to directly or indirectly generate value for the firm. Thereby, community participation can be translated into monetary and non-monetary benefits for the community operators and third parties involved in them. In general, online communities can gain value through (1) revenues from advertising, membership fees or selling transactions, (2) increased customer loyalty and retention, (3) innovations derived through ethnographic observation of customers, or (4) increasing organizational efficiency and effectiveness by sharing knowledge, solving problems of other customers, or providing self-service to employees, experts, or other interest groups (e.g., Algesheimer, Dholakia, and Herrmann 2005; Clemons 2009; Dholakia et al. 2009;

Armstrong and Hagel 1996; Nambisan and Baron 2007). In order to capture this value, a critical mass of members (Markus 1987; Preece 2000), who show enough participation, interaction and contribution of relevant information, is needed. If there are not enough members contributing, the value of the community is undermined and it cannot survive for lack of content (Markus 1987; Morris and Ogan 1996)[10]. Therefore, online communities can provide value to the community operator and marketers, if enough participation takes place.

(1) Advertising and Fees. Similar to traditional media, one main income stream of the Internet is advertising revenue. In fact, Internet advertising spending increased to US$ 26.0 billion in the US in 2010 (IAB 2011), which emphasizes the importance of advertising in the online medium today. Further, many online communities rely on advertising as their predominant revenue model (Enders et al. 2008). According to eMarketer (2011), Facebook achieved US$ 1.86 billion in advertising revenue worldwide in 2010, where advertising accounted for more then 90% of total revenue. Advertising is a significant revenue stream. Online communities are of specific value for advertisers, because they allow for a more targeted advertising than other media. Advertisers can focus on specific demographics and interests of the users. Many online communities also provide access to a targeted audience per se, because they are directed towards a specific topic. For example, communities on motherhood would provide a good advertising platform for child and pregnancy related topics. Therefore, online communities represent a highly attractive and efficient way to reach customers.[11] Although advertising is frequently used to generate revenue for community operators, banner ads are often ignored or blocked by built-in ad-blockers, so that conventional display advertising is becoming less attractive. Further, targeted advertising and the use of member data evokes privacy issues for the end user. Therefore, advertising concepts need to be selected carefully.

A way to directly capture monetary benefits for community operators is to get usage, content, or transaction fees (Armstrong and Hagel 1996). While usage fees are related to accessing the online community itself, content fees are associated with the access to specific content, like articles, photos, or videos. Transaction fees are realized by selling products and ser-

[10] Interactive media only become useful as more and more people adopt it. The size of the critical mass is dependent on the type of online community, the definition of the target population, and the purpose of the interaction. But all online communities require a critical mass of participants to supply enough content as a public good to survive (Morris and Ogan 1996).
[11] Note that advertising can also take different forms. According to the IAB (2010) there are eight types of advertising formats: search, display/banner, classifieds, lead generation, digital video, rich media, sponsorship and Email. The relevant formats depend on the purpose, functionality and content of the respective online communities. However, display/banner seem to be an important element for most communities.

vices directly or through provisions paid for such transactions between the community users and third parties. For example, online communities can sell products related to the topic they are centered around or provide the platform for other companies to sell their products and services. However, fees are not always favorable for online communities. Usage fees do not encourage members to use all functionality and users do not stay in the community for a long time when the functionality is limited (Armstrong and Hagel 1996). In the same way, membership fees would also hinder the growth of an online community, which is especially important in the growth stage to reach a critical mass of members, who contribute content (Hagel and Armstrong 2006; Rothaermel and Sugiyama 2001).

(2) Loyalty and Retention. Online platforms like brand communities or problem-solving communities relate to consumption topics, products, brands or services. They share the opportunity to increase customer loyalty and commitment to those products and services (e.g., Algesheimer, Dholakia, and Herrmann 2005; Casalo, Flavian, and Guinaliu 2008b). From a customer relationship perspective, the community helps to build a stronger tie between the user and a company and its products. For example, an Apple community can bring together Apple enthusiasts, who exchange information and experiences about Apple products, which in turn increases their loyalty to the brand (Shang, Chen, and Liao 2006). In addition, users expressing their loyalty and positive feelings towards a certain product or brand can potentially spread positive word-of-mouth recommendation to other consumers in the online community, thereby increasing their interest and buying intentions (e.g., Chevalier and Mayzlin 2006; Hennig-Thurau and Walsh 2004). A potential threat to firms is the occurrence of negative evaluations of products and services. The producers of products and providers of services might profit from getting deeper knowledge about failures of products, as this helps to find improvement potential. However, the emergence of negative word-of-mouth communication or even hate communities can have a significant impact on sales. For example, Chevalier and Mayzlin (2006) show that the impact of one-star reviews on online book review sites is greater than the impact of five-star ratings, suggesting that negative WOM has a higher effect on buying behavior.

(3) Customer Insights. Online communities are valuable sources for customer information, because they provide an unobtrusive way of getting access to customer opinions by observing public conversations. This can help to better understand how customers use products and services, what factors they take into consideration when buying products or potential problems that might occur with the usage of products. Particularly, online communities involving product evaluations and ratings can provide interesting information for producers. In fact, online community members can act as a sort of focus group, which discusses specific product related topics (Hagel 1999). Kozinets (2002) calls this type of market research "netnography", which is ethnography adapted to the study of online communities, but faster, simpler, and less expensive than traditional ethnography, and more naturalistic and unobtrusive than focus groups or interviews. A careful evaluation of user discussions can also reveal innovative ideas and consumer insights, leading to novel product concepts by investigations that begin with netnography (Kozinets 2002). Some firms observe public discussions for innovation management, with examples coming from computer-controlled music instruments or car communities, which make use of this information provided by its users (e.g., Henkel and Sander 2007; Jeppesen and Frederiksen 2006). Although, netnography and user information in online communities provide valuable insights, it has some limitations of which market researchers must be aware. Qualitative investigation of online communities narrows the focus on its members and is dependent on the interpretive skills of the researcher, which might lead to difficulty in generalizing results to groups outside the online community sample (Kozinets 2002). However, it is a cheap and easy way to gain valuable information from the customers.

(4) Synergies. Online communities can also capture synergies, such as reducing service costs (Hagel and Armstrong 2006). Online problem-solving communities can provide peer-to-peer customer service, where users exchange answers on how to deal with specific problems or how to make better use of a product or service. Examples include forums on statistical software packages, where the users help each other to find the right solution for their statistical problems. Also knowledge sharing communities can help to enhance organizational processes by providing access to the experience and know-how of other users, thereby reducing search costs.

The discussed benefits show that the value of online communities is versatile and depends on the business model and objectives. Although some online communities do not have an intention of generating monetary benefits (e.g., member-initiated communities), the majority of large public online communities have a commercial interest. The commercial role of online communities gained importance for community operators and sponsors in order to ensure its sustainability and generate value (Hagel and Armstrong 2006). As previously mentioned, it is critical to have a sufficient number of users taking part and contributing to the online community to achieve the described benefits. If a user signs up for an online community, and there are no other members interacting or sharing information, it is not very valuable for that user. Community operators need to consider the advantages and disadvantages of each form of benefits, the type and context of the particular community, and its objectives in order to make the right choice about how to capture the value of their online community. Eventually, marketers should use online communities as a platform for collaboration marketing, that is to listen to the customers, integrating their input and developing a broad relationship (Hagel 1999).

2.4 User Benefits and User Types in Online Social Communities

2.4.1 User Benefits of Online Social Community Usage

From a member perspective, the value of participating in an online community comes from the fulfillment of specific member needs. According to the different types of communities defined by Hagel and Armstrong (2006), which are based on certain member needs (i.e. communities of relationship building, of interest, of transaction, and of fantasy), the members essentially profit from (1) social support and relationships, (2) information exchange, (3) economic benefits, and (4) entertainment. (1) Participants of online communities can fulfill social needs by connecting with other community members, finding new friends and like-minded people, as well as getting social support when experiencing problems in their lives. For example, online communities about illness and severe diseases (like cancer) can help people to express their fears and hopes and exchange their experience with others. (2) The information provided by other members on certain topics can help to solve specific tasks and gain additional know-how. Knowledge sharing and problem solving communities take a vital role in this respect (e.g., Chen and Hung 2010; Dholakia et al. 2009). The social capital that lies in the connection to other members in the community makes it possible to achieve benefits from knowledge and information exchange. (3) Economic benefits can arise when exchanging information and experiences about the purchase and consumption of products and services. This can help in making the right buying decision, for example getting a better deal when planning to purchase a product. Furthermore, the information users share with each other on products and services can also be seen as word-of-mouth that takes place in an online community. Product recommendation and rating platforms are popular places to pursue this activity. Because word-of-mouth is perceived as more trustworthy and risk reducing than other marketing communication channels, information from other customers can take an important role in purchase decision making (Arndt 1967; Murray 1991). (4) Entertainment and recreation can also be achieved in online communities. Here, users can potentially take on roles of virtual characters and play games together or just consume text, photos and videos to have fun.

Ridings and Gefen (2004) basically confirm that these categories cover the most important motives to join online communities. They identified information exchange, friendship, social support, and recreation as the main reasons for joining. They also demonstrated that the reasons differ depending on the community type. Consequently, there might be other needs relevant for the users to participate in an online social community, depending on the purpose of the community. Dholakia, Bagozzi, and Pearo (2004) determined purposive value (as a form of information needs), self-discovery, social enhancement, maintaining interpersonal connectivity, and entertainment value as the main motives for using online communities. In that respect, they add to the proposed categories of Hagel and Armstrong (2006) the social enhancement motive, which is an extrinsic motive associated with the user's social status in the community. Self-discovery might be related to the communities of fantasy and interest, because those communities might specifically help "to form, clearly define and elaborate on one's own preferences, tastes, and values" (Dholakia, Bagozzi, and Pearo 2004, p. 244). Overall, information, social, entertainment, and extrinsic benefits (e.g., economic benefits or status seeking) are considered the most important benefits for online community users, because they are frequently mentioned in recent online community studies (e.g., Dholakia, Bagozzi, and Pearo 2004; Park, Kee, and Valenzuela 2009; Sangwan 2005; Wasko and Faraj 2000).[12]

2.4.2 Identification of Different User Types in Online Social Communities

Motivation-Based Typology. To identify different types of users, one approach is to categorize them based on their needs. Hennig-Thurau et al. (2004) used the users' motives to participate in web-based opinion platforms in order to come up with a user segmentation. Four different clusters were identified based on the level of motivation in each category[13]: a) 'self-

[12] Chapter 2.5 includes a review of the effects of user motivation on user participation in recent empirical online community studies.

[13] The eight motivational factors are: platform assistance, venting negative feelings, concern for other consumers, extraversion/positive self-enhancement, social benefits, economic incentives, helping the company, and advice seeking (Hennig-Thurau et al. 2004).

interested helpers' appear to be strongly driven by economic incentives, b) 'multiple-motive consumers' are motivated by a large number of different factors, c) 'consumer advocates' are motivated predominantly by the concern for other consumers, and d) 'true altruists' appeared to be both strongly motivated by helping other consumers as well as helping companies (Hennig-Thurau et al. 2004). Overall, it is evident that users can gain different benefits from participation in online communities. Several motivations may work at the same time at different levels. Therefore, it is important for online community operators to address multiple user needs to be successful. Further, the clusters also show different participation levels. For example, multiple-motive consumers, scoring highest on almost all dimensions, showed the highest contribution activity and visit frequency on the platform (Hennig-Thurau et al. 2004). For community operators, such a segmentation suggests that they may need to develop different strategies for encouraging participation behavior among their users. Although this segmentation provides valuable information for this specific community, relying on an existing typology can be misleading. Therefore, operators should run their own segmentation as the motives and their respective levels differ across communities.

Participation-based Typology. As is evident from the results of Hennig-Thurau et al. (2004), not all users participate equally. Because participation is critical for the existence of online communities, a more popular differentiation of member types is based on the member's participation behavior. A first distinction can be made between members and non-members of the online community. ***Members*** are people who registered for the online community. Accordingly, ***non-members*** are not registered. The reason for not being a member might be twofold: either they have never been with the community or they unsubscribed with the online service. One important goal of community operators is to turn non-members into members and retain existing members in order to maintain a sufficient number of users on the platform. For example, marketing efforts must be made to gain recognition of the online community and to communicate its benefits to attract new members.

It is important to note, that members can also become ***inactive members*** when they stop using the platform without de-registration. Such members are still members in the broader sense, but they are not active anymore, that means they do not log in the online community, neither to post nor to consume content or surf the site. The definition of inactivity is most often based on the customers' last activities (e.g., Wübben and Wangenheim 2008); for example, users who have not logged in for more than three months could be classified as inactive (cf. Chapter 7 for a description of user defection in an empirical study of this dissertation). ***Active members***, on the other hand, regularly log in and use the platform.

When looking at the group of active members, the most common generic typology is to differentiate between posters and lurkers. ***Lurkers*** are still active, but they merely browse the platform and consume content (Madupu and Cooley 2010). Lurkers visit and use the community, and they participate by reading the posts of others, spending significant amounts of time doing so, but they do not post messages (Ridings, Gefen, and Arinze 2006). Lurkers are mostly invisible for the rest of the community members, because they do not share any information and do not interact. On the other hand, ***posters*** are those people who contribute content and interact with the community and its members. A poster adds to the discussion and actively invests time and effort in the online community (Ridings, Gefen, and Arinze 2006). They are important to bring new content to the community and keep the community alive. The share of lurkers to posters depends on the type of community, but there are indications that there are much higher numbers of lurkers in many online communities (e.g., Nonnecke and Preece 2000). Although a large amount of active community participation stems from only few users of the total customer base in many cases (e.g., de Valck, van Bruggen, and Wierenga 2009), it is important to note that not only contributors build the community, as lurkers can become contributors over the time. Further, lurkers browse the site and consume information, which can also lead to increased page impressions, and consequently higher advertising revenues.

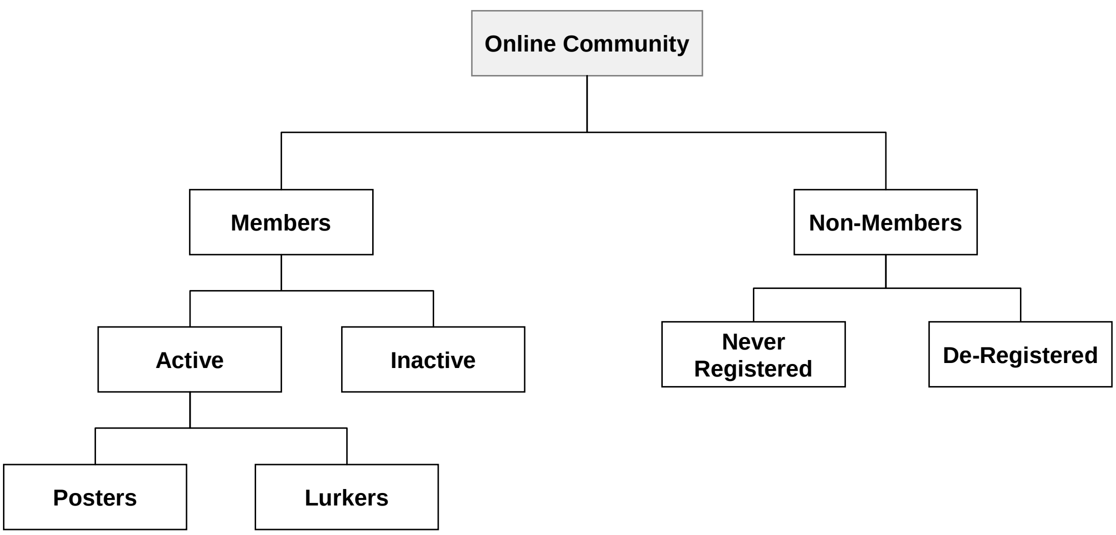


Figure 3: Typology of Online Community Users based on Activity and Participation

Figure 3 provides an overview of the generic typology discussed. There are many possible ways to segment users and to describe different types of users, which depend on the purpose and the functionality of the online community. Community operators should define criteria relevant to their specific community in order to get an appropriate classification of users in their online network. However, the basic distinction between members and non-members, active and inactive users, and lurkers and posters already provides a useful basis for using dedicated marketing communication to these different user groups.

Further Differentiation of Posters and Lurkers. The poster-lurker dichotomy is useful, but a finer granularity may provide community operators and researchers with additional understanding when looking at those two user groups. Ridings, Gefen, and Arinze (2006) define three groups of users instead of two: lurker, poster, and infrequent poster. In particular, they further differentiate posters. Infrequent posters are defined as users who mostly show passive participation, such as lurkers, but also contribute to and interact on the platform in irregular intervals. Consequently, posters contribute more regularly, for instance, in the case of Ridings, Gefen, and Arinze's (2006) study this is four or more times per month. In fact, the authors demonstrate that the two poster groups differ in their levels of trust and motivation towards the online community, thereby justifying a further breakdown of posters.

In another empirical study, de Valck, van Bruggen, and Wierenga (2009) present a classification based on the participation of the users in terms of their frequency of visits, duration of visits, retrieved information, supplied information, and discussed information. They identified six distinct clusters: a) 'Core members' represent the most active participants within the community, who score far above the mean on all variables (~6% of respondents). b) 'Conversationalists' make frequent, but short visits, and they participate to a relative high degree in supplying and discussing information. In particular, they show a relative high level of engagement in forum discussions and chat sessions (~10% of respondents). c) 'Informationalists' show relatively high participation in retrieving and supplying information, but they score low on discussing information; their visit frequency and duration is comparable to that of the conversationalists (~14% of respondents). d) 'Hobbyists' visit the community frequently for an extended time, but they show low levels of information retrieval, supply, and discussion (~17% of respondents). e) 'Functionalists' only retrieve information to a larger degree, but show low participation with regards to visit frequency and duration, information supply and discussion (~28% of respondents). f) 'Opportunists' score far below the mean on all five clustering variables. Therefore, they are the least active and least regular participants (~25% of respondents). Overall, this classification demonstrates that there are only few users, particularly the core members and partly the conversationalists and informationalists, who contribute to a larger extent. Most users are less actively involved in the community.

An even greater granularity of activities is used by Alarcon-del-Armo, Lorenzo-Romero, and Gomez-Borja (2011) to come up with different user types of a social networking site. Their segmentation is based on the frequency by which users perform 20 different activities on the platform. Four different segments have been obtained. “Introvert users’’ are the least active users, using the social networking site mainly to send E-Mails. “Novel users” occasionally contribute to the platform, mainly by communicating with friends, sharing comments and messages, and spend more time on the site than introverts. “Versatile users’’ perform many different activities, although occasionally. They do not only use the community for communication with friends. Finally, “expert-communicators’’ pursue a great variety of activities with a higher frequency. They are the most active group of people. Again, these clusters differ in their usage intensity. But users from different clusters also show distinct motivations to use the site as well as different numbers of contacts in the community (Alarcon-del-Armo, Lorenzo-Romero, and Gomez-Borja 2011). For example, expert-communicators are not only most active, but they also have the highest number of contacts.

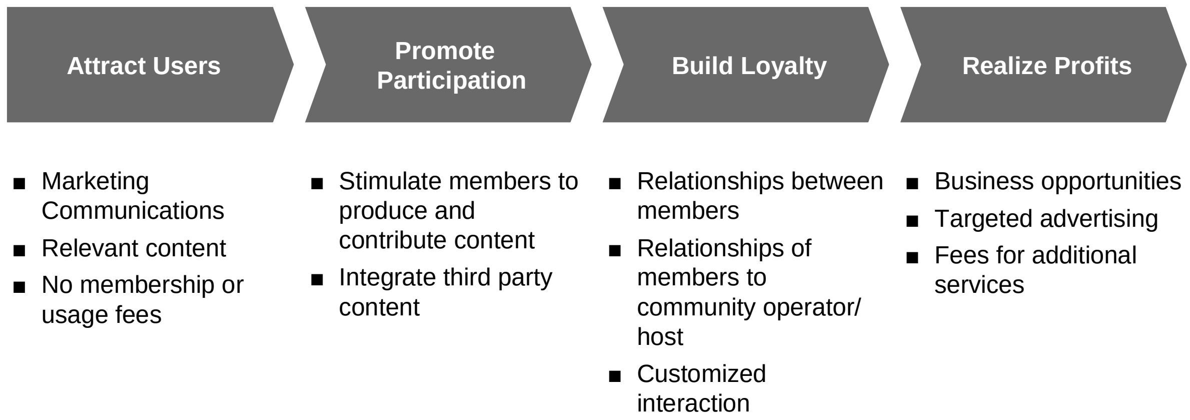
These three recent empirical studies showed different user segments based on more granular activity categories. The results indicate that in most online communities lurkers and members who barely contribute make up the largest part of users in the community. In contrast, the most active members with the highest levels of contribution and interaction make up smaller proportions of the online community population. This group of people is of specific interest for operators to ensure a sufficient number of information is exchanged on the platform. In between are users who show moderate participation patterns, and who can show high intensity in a few specific activities. Community operators might address each customer segment differently, according to their participation patterns and their needs for interaction and information exchange in the online community. For marketers and scholars it is therefore important to understand what influences user participation, and how users of different segments can be stimulated. Thus, it is the aim of this dissertation to gain further knowledge on the influencing factors of user participation, which is addressed in the empirical investigations in chapters 5, 6 and 7.

2.4.3 The Development Process of Users in Online Social Communities

In this section, a more dynamic perspective on user types is presented along the life cycle of online community members. This provides further insights on how different user types develop over time. An understanding of this dynamic perspective should also help to set the research focus of this dissertation in context, where the empirical studies of this work are targeted to get insights at specific stages of the membership development (cf. chapter 4.1).

As already shown, online community users differ in their participation levels. In the course of their relationship with the online community, users can evolve to take on different roles and represent a different member type. Therefore, belonging to a specific user type is a rather dynamic than a static classification, because over their life cycle users can fall in different categories of user roles. “From a firm’s perspective, customer life cycle can be best understood as a series of transactions between the firm and its customer over the entire time period the customer remains in business with the firm. Customer life cycle varies from business to business and customer to customer and could be short or long depending on the nature of business of the firm, the profile of its customers, and the interaction between the firm and its customer” (Jain and Singh 2002, p. 35). In an online community, the user life cycle refers to the entire time the user is in a relationship with the community. Considering the life cycle of the members provides a more dynamic perspective on how the members might develop over time. The classical poster-lurker dichotomy and also more granular classifications of user-types describe rather static typologies of online community users. However, user participation behavior in online communities is more diverse and flexible, and users can also dynamically switch between different roles. Thereby, they can be lurkers at one point in time, and posters at another point. Depending on the interaction and participation with the online community, individual users belong to varying user types. Instead of solely using active and passive participation behavior, some researchers take a more process oriented view and include the users’ involvement in the online community to distinguish different user roles. In the following, different concepts are presented, which are based on a dynamic interpretation and evolution of the member roles.

The typology of users often follows different stages they take in order to develop in the online community (Hagel and Armstrong 2006). The “typical” membership development process from an operator perspective contains four stages. (1) Attract members: community operators need to get attention for the online community. Similar to other products and services potential customers need to be convinced to try the service, i.e. become members by registering and try out the online community. (2) Promote active participation: the next step is to make the members participate more actively; they should visit the community more often and spend more time using the platform. (3) Increase loyalty: facilitate relationships to other members and to the community operator. Ideally, no member should be lost due to decreasing interest or competition. Therefore, user retention is of high importance. (4) Generate profit: the commercially oriented operator wants to profit economically from their members, for example through advertising or fees (cf. chapter 2.3 for different options to generate value for the operator). The last stage is often realized in the course of user participation through page impressions in advertising-based business models. Nevertheless, other business models would, for example, thrive on converting the members into premium members, paying extra fees to have access to more functionality (e.g., Xing.com) or make them purchase products and services via the online community. Figure 4 illustrates these four stages.



Source: Adapted from Hagel and Armstrong (1997, p. 106)

Figure 4: Membership Development Process

Within this membership development process, users can take different roles that represent different levels of value contribution to the community. Hagel and Armstrong (2006) describe four types of users: browsers, builders, lurkers, and buyers. Browsers are new members, who explore the platform without obligation and commitment, and who show low usage intensity. Browsers are of low economic value, but they can develop to lurkers or builders. Lurkers spend more time in the community than browsers, they consume content and use information, but are also of lower economic value. Builders are engaged in the community, they spend a lot of time in the community and contribute to a high degree. Buyers are those members, who actively purchase products and services, generate provision fees, and provide high economic value. Nevertheless, the role of buyers is not existent in every online community, and depends strongly on the purpose and objectives of the community. If the community does not provide the sale of products, buyers would not be present. Hagel and Armstrong (2006) emphasize, that most users will most likely not be assigned to one single role but take on different roles in different situations.

Another approach of classification is based on the users' level of involvement with the online community and the consumption activity. Kozinets (1999) determines four member types in communities of consumption: “Tourists” have weak social ties to the community and only superficial interest in the consumption activity. “Minglers” maintain strong social ties, but have minimal interest in the consumption activity. “Devotees” have weak ties, but show strong interest and enthusiasm in the consumption activities. Finally, “insiders” have strong ties to the community and to the consumption activity. The core segment is represented by insiders and devotees, who tend to be the most important targets for marketing, because they appear to be loyal heavy users (Kozinets 1999). Users can progress from a visitor towards the insider over time by gaining experience and getting to know the group of members.

Kim (2000) describes five stages of community membership: visitors, novices, regulars, leaders, and elders. While visitors have not registered yet on the platform, novices are registered members, but they still have to be introduced to the online community and need to learn how the community works and what the rules and roles are. Regulars are more experienced members, who know how to find information, how to use the functionality of the online community and how to interact with other people efficiently. Leaders are user who take important roles, help to integrate novices and support the operations of the online community. Elders are less active than leaders, but they are long-term users of the community who “give the place a sense of history, depth and soul” (Kim 2000, p. 119). This five-stage life cycle helps to understand different user roles in established online communities. However, it does not give objective criteria to identify the different user roles. It only implicitly relates to the intensity of user participation. Over the course of the membership, users evolve to become more and more active, until they reach the stage of the elder, where they reduce their participation intensity.

One limitation of the life cycle model of Kim (2000) is that the membership develops through the different stages. Sonnenbichler (2010) builds upon Kim's (2000) generic community membership life cycle model, but includes different development paths between the different roles, which makes the model more appealing. He suggests six roles in the user life cycle: visitors, novices, actives, passives, leaders, and trolls. Other than Kim's (2000) roles of visitors, novices, and leaders, he defines active, passive, and troll users instead of one role of regulars. Essentially, actives and passives refer to posters and lurkers/infrequent posters, as defined above. Trolls take a more negative role, and are defined as users who want to disturb the community and cause trouble. The advantage of Sonnenbichler's (2010) model is that it allows for a more flexible development of users to other roles, not only forward development to become more active and involved, but also backward development to reduce involvement and activity in certain time spans. The paths from one user type to another are more flexible. While visitors can only become novices, novices can develop to take the role of trolls, actives, or passives. Actives can either become trolls, passives or leaders. In this respect, actives can either show higher commitment to the community and be rule and opinion makers in the role of leaders. But they also can show negative behavior and become trolls, or they show no active participation and become passive users. Passive users can become active, thereby actively contributing. And finally, leaders can take a step down to become actives or passives, or also become trolls. The different development paths show, that each user can move between roles and be active, passive, or a leader. The roles are not static and individual users do not necessarily have to evolve only in one direction, they can also become more passive over time again. Overall, the generic roles provide a useful understanding for community operators on how the user base of their community is structured, and what activities could be taken to improve the community.

Table 2 summarizes the different typologies presented in this chapter. Some typologies are conceptually developed, while others are gained from empirical research. The different typologies help community operators to better understand their members' participation behavior and needs in the community. The life cycle perspective shows that user roles can evolve and change over time, but not only in one direction. Although usually involvement increases with the duration in the online community, users can become more or less active in the community over time, until the user ends the relationship with the online community.

Authors	Criteria for Classification	User Types
Hagel and Armstrong (1997)	Online community user types; static perspective; generic classification	• Browser • Builder • Lurker • Buyer
Kozinets (1999)	Online community of consumption member types; based on social and topical involvement; dynamic perspective; generic classification	• Tourists • Minglers • Devotees • Insiders
Kim (2000)	Online community member roles; dynamic perspective; generic classification	• Visitors • Novices • Regulars • Leaders • Elders
Hennig-Thurau et al. (2004)	User motives to participate in web-based opinion platforms; static perspective; empirical classification	• Self-interested helpers (34%) • Multiple-motive consumers (21%) • Consumer advocates (17%) • True altruists (27%)
Ridings, Gefen, and Arinze (2006)	Contribution to the online community; static perspective; empirical classification	• Lurkers • Infrequent Posters • Posters
de Valck, van Bruggen, and Wierenga (2009)	Participation in the online community, including visit frequency, visit duration, information retrieval, supply and discussion; static perspective; empirical classification	• Core members (6%) • Conversationalists (10%) • Informationalists (14%) • Hobbyists (17%) • Functionalists (28%) • Opportunists (25%)
Alarcon-del-Armo, Lorenzo-Romero, and Gomez-Borja (2011)	Types of social networking site users, based on frequency of different activities; static perspective; empirical classification	• Introvert (19%) • Novel (25%) • Versatile (36%) • Expert-communicator (20%)
Sonnenbichler (2010)	Online community member roles; dynamic perspective; generic classification	• Visitors • Novices • Actives • Passives • Trolls • Leaders

Table 2: Selected Overview of Different User Typologies and Roles

Particular types of members and different stages in the membership life cycle require specific marketing activities accomplished by the community operator to make the platform more successful. It is important not to treat every member the same as they obviously differ significantly in their relationship to and participation in the community. What is common in most typologies is that there are specific groups of core users (also called leaders, experts, heavy users, builders, etc.), who are of specific importance, because they increase the amount of contributions and content on the platform (e.g., Hagel and Armstrong 2006). Those most active users must be retained and kept active. Passive members might be activated to become active members, so that more content is generated and the interaction in the community is increased. Further, communities should also be interested in winning new members, because the loss of some passive and active members cannot be avoided. The objective of the community operators is to attract a high number of active users, make passive users active, and retain the active users on the platform in order to ensure that the overall number of members and the proportion of active members are on a sufficient level. Thereby, user classification helps to better understand the structure of the online community, and consequently identifies measures that should be taken to keep the participation at a healthy level.

2.5 Online Community Participation – A Literature Review

Traditionally, marketers are interested in customer behavior and why certain behavior occurs. Online communities have become an interesting research object in academia because of their ability to investigate the dynamics of customer behavior in these communities and the interaction between their members. After the publication of the work of Rheingold (1993) and Hagel and Armstrong (1997), different research streams started to elaborate on the importance of user participation in online communities. Although marketers have been quite effective at developing an understanding of online communities and their usage, a systematic overview of the antecedents, constituents and consequences of online community participation is still missing. This section provides a broad overview of past research studies. It further outlines some important research gaps and approaches which should be addressed in research on online community participation, and which in fact will be addressed in the empirical studies of this dissertation. Therefore, this review is also meant to provide a basis for the empirical studies presented in chapters 5, 6 and 7.

Because online community research is multi-disciplinary, studies from different disciplines are considered to cover a versatile and rich set of relevant influence factors. This includes for example studies from the areas of marketing and management, information systems, knowledge management, organizational research, and social-psychology. Thereby, different perspectives can provide valuable insights on how online community participation works and what its marketing-relevant causes and consequences are.

Although conceptual as well as empirical studies contributed to online community participation research, this literature overview focuses on empirical studies which provide evidence for causal relationships between user participation and its antecedents and consequences. Empirical studies test the theoretically developed hypotheses on the relationships between different variables. Such statistical tests provide concrete insights on these relations and are therefore of high value. Further, empirical studies have already been widely used in online community participation research, which allows for a broad review of relevant studies. Following that, those research studies are taken into account, which used a construct or measure that can be directly referred to participation in an online community.[14] Because the focus of this dissertation is related to the individual-level of user behavior, rather than an aggregated or macro-level of overall participation in online communities, the literature overview is concentrated on the user-level perspective.

First, a description and definition of online community participation is given, followed by an overview of research studies from different disciplines, which investigated important factors associated with user participation. From this overview a framework of antecedents and consequences of online community participation is derived and the gains and gaps of recent empirical studies are discussed.

2.5.1 Description and Measurement of Online Community Participation

In order to establish a successful online community it is important to reach and maintain a sufficient level of user participation. User participation and engagement have been acknowledged to take on a central role in online communities in order to keep the community attractive, retain members and create value for the firm (e.g., Hagel and Armstrong 2006; Bagozzi and Dholakia 2002; Butler 2001; Casalo, Flavian, and Guinaliu 2007; Woisetschlaeger, Hartleb, and Blut 2008). In a general sense, online community participation is described as the behavior of members within the online community of interest. Thereby, it is necessary for users to register and log in the online community in order to take advantage of its functionality, published information and social interactions. Membership alone does not imply participation. In fact, online communities do not only accommodate active users, but also inactive 'ghost' accounts. Such inactive members do not participate on the community website anymore, although they still might be members. Consequently, users do not provide value to the

[14] Only 'online' community research is considered in this literature review.

community and the operator when inactive.[15] Therefore, in this review community participation does not simply mean community membership, it rather requires active community membership, and is associated with users who log in and show onsite activity.

It is the ongoing participation of its members, their willingness to use the community, as well as to contribute to the community that creates value for other members and keeps the community alive. If there is too little social interaction, because of an insufficient number of members and contributors, this can lead to the downfall of the online community (e.g., Preece 2001). Because of the interplay between community size and communication activity (which is a form of benefit creation for the members and the community operator), it is critical to hold a certain level of activity to sustain the community (e.g., Butler 2001).

Different Types of Participation. Participation can take on different forms in online communities. Relating to the already described distinction between lurkers and posters (see chapter 2.4.2), participation can be either passive or active. Lurkers visit and use the online community and are active members, but they merely show passive participation, browsing the platform and consuming content instead of posting messages (Madupu and Cooley 2010; Ridings, Gefen, and Arinze 2006). Therefore, passive participation is characterized by activities like reading posts and articles, browsing profiles of other members, watching photos and videos, or searching for relevant information and knowledge. The value of lurking lies in the consumption of information. For community operators, passive participation is important because it can generate revenue, for example, in advertising-based business models through exposure to online ads and the selling of the members' attention to advertisers (Ridings, Gefen, and Arinze 2006). Further, passive participation helps the users to understand the community better, and lurkers can potentially turn into contributors (e.g., Nonnecke and Preece 2001; Ridings, Gefen, and Arinze 2006).

On the other hand, posters show not only passive, but also active participation. They contribute content and interact with the community and its members. Thus, active participation is the main element to keep the online community alive and provide it with endurance (Hagel and Armstrong 1997; Rheingold 1993). Without active users producing public content and interacting with other users, the online community would lose its capital and its members because they are not stimulated by new and dynamic content. For example, social networking sites like Facebook are dependent on their users' contributions, because the posts of one's

[15] Because inactive users do not contribute anymore to the value of the community, they are not of interest from a customer behavior perspective. Though, it has to be noted that wining back such users can be an important activity for community operators.

friends build the core information provided to its members. An empty newsfeed of the friends' activities would undermine the value of the online community platform because of the lack of relevant content and information. The same is true for other kinds of online communities. When there is no valuable information and interaction on the site, the community members would lose their interest. Therefore, active participation and social interactions are at the heart of online social communities.[16]

Operationalization and Measurement of User Participation. Past research has investigated active and passive participation in online communities, as both types of participation are important for the success of the community. Translating key user behavior into a consistent operational definition has proven to be challenging. Because various research disciplines studied online communities from different perspectives in the past, it is not surprising that this led to versatile operationalization of online community participation. For example, some studies on knowledge sharing communities build upon the volume and quality of knowledge contributions by community members (e.g., Chen and Hung 2010; Wiertz and de Ruyter 2007), while others use more general participation indicators like frequency of visits (e.g., Dholakia, Bagozzi, and Pearo 2004; de Valck et al. 2007; Hennig-Thurau et al. 2004) or willingness to use the community (e.g., Chen 2007; Tiwana and Bush 2005). Because online community participation has been operationalized, tested and applied in numerous ways, an overview of different operationalization and measurement alternatives is provided. This should help to better understand the different facets of community participation investigated in past studies.

Frequently, data on user participation is collected employing user surveys and observing user behavior. In questionnaires, participation is either measured on Likert-type scales (e.g., Casalo, Flavian, and Guinaliu 2008a; Hsu et al. 2007) or by directly asking for the quantity of contributions, visits, or time spent using the community (e.g., Chen and Hung 2010; Dholakia, Bagozzi, and Pearo 2004; Hennig-Thurau et al. 2004). Likert-scales are used to ask about the user's perceptions or intentions of participation in the respective online community. While some studies use items and constructs to measure the current behavior (e.g., "I usually actively share my knowledge with others"), others measure the intentions of participation

[16] One might argue that the community operator could publish content in the community to keep it alive. This can help stimulate the participation of the users. However, if the users do not talk about the published content and interact with each other, but rather consume only the content provided by the community operator, the community would turn more into a one-way communication mode, providing its users with content and not forming relationships between users. Consequently, the online community would loose its character of being a community without active participation and interaction between users.

(e.g., “I intend to continue sharing knowledge in this discussion forum”). According to the theory of planned behavior, behavior is directly related to and follows behavioral intentions (Ajzen 1991). In fact, it has been demonstrated that community participation is affected by intentions to participate (Algesheimer, Dholakia, and Herrmann 2005; Dholakia, Bagozzi, and Pearo 2004). Therefore, behavioral intentions can also provide insights on participation behavior in online communities.

Alternately, objective data uses member activity that is either tracked by the community operator or collected through content analysis on the website (e.g., Wasko and Faraj 2005; Wiertz and de Ruyter 2007). Objective data has several advantages. It can help to address some relevant issues of self-reported marketing research such as common-method bias (Podsakoff et al. 2003) or recall of actual behavioral actions. It provides an accurate and unbiased measure of participation. Consequently, it would be the preferred method to get data about user behavior. Nevertheless, due to limited access to objective data sources, in past studies the prevalent method used to get information about community participation has been the deployment of self-reported survey measures.

Independent from the source of the data, be it self-reported or objective, both active and passive participation have been measured and used. Although some researchers investigated both active and passive participation in one study (e.g., Kang et al. 2007; Chen and Hung 2010; Koh et al. 2007), most studies measure either one of the two participation types. Passive participation is often measured in a more general form, asking for frequency of use, time spent in the community, intention to keep on using, or how many posts have been read (e.g., Dholakia, Bagozzi, and Pearo 2004; de Valck et al. 2007; Lin 2007; Wang and Fesenmaier 2004b). General measures of participation, like usage frequency and time spent on the platform, cover the overall behavior using the platform. Thereby, they can also include active participation behavior. However, because passive behavior often occurs to a larger extent than active behavior (Nonnecke and Preece 2000), general participation measures are considered to capture rather passive participation. On the other hand, active participation is mainly described by the number of contributions and willingness to actively contribute content (e.g., Koh and Kim 2004; Wasko and Faraj 2005). The concrete form of active user participation depends on the functionality and purpose of the online community, which are determined by the community operator. With the advancement of digital technology and social software, functionality and interactivity within online communities increased over the past decade and resulted in more and more possibilities for users to actively participate. Today’s online social communities offer a wide variety of functionality and applications, for example to write messages, guestbooks and status updates, to post profile information, comments in discussion groups, pictures and videos, to chat, to play games, or to use social apps. Basically, the participation in the online community is predetermined and also limited by the functionality provided on the platform. However, most online communities provide similar basic functionality which enable active participation. Thus, past research often refers to writing messages, posting in discussion groups or forums, uploading content, or more generally sharing knowledge, or providing information.

Table 3 shows a categorization of the different measurement approaches used in the past and examples of operationalization. The first dimension differentiates between self-reported quantitative, self-reported Likert-scaled perceptual and intentional, and objective measurement. The second dimension distinguishes between active and passive participation. Although many alternative operationalizations of user participation exist, the common objective of recent studies is to identify how participation is influenced and what consequences it has. Different perspectives on participation can even help to develop deeper insights into online community participation. As both active and passive participation can create value for the community operator and the community members, an overview of studies which employed active participation, passive participation, or both, should provide a better understanding of its antecedents and consequences. Nevertheless, when interpreting the results of empirical studies, one should consider how participation is measured.

	Active Online Community Participation	**Passive Online Community Participation**
Self-Reported, Likert-Scaled Data of User Participation Perceptions and Intentions	"I post messages with excitement and very frequently on the Online Community Site" (strongly agree – strongly disagree) "How likely are you to contribute content to the Online Community in the future" (very likely - very unlikely)	"I intend to read a lot of posts in the future" (strongly agree – strongly disagree) "I intend to visit XXX in the future" (strongly agree – strongly disagree)
Self-Reported, Quantitative Data on User Participation	"How many messages do you post per week?" „Number of comments published" (1-10; 11-35; 36-99; 100+)	"How many posts do you read per week?" "How many times do you visit the Online Community per week?"
Objective Data of User Participation	Number of knowledge contributions Number of photos uploaded	Number of posts read Number of logins Hours spent on the platform

Table 3: Examples of Online Community Participation Measurement

2.5.2 *Different Research Perspectives on Online Community Participation*

Given that online community research is multi-disciplinary, factors associated with user participation have been studied from various angles. While some aspects are of general importance across disciplines, other elements are of specific interest only for a certain research stream. However, most studies are related to the area of marketing and management or information systems. Therefore, a closer look is taken on these two research areas.

In marketing and management studies, the importance of user participation is underlined by its influence on loyalty, recommendation behavior and transactional behavior (e.g., Algesheimer et al. 2010; Kim, Lee, and Hiemstra 2004; Woisetschlaeger, Hartleb, and Blut 2008). In fact, research in the area of marketing and management has mainly investigated behavioral and transactional outcomes of online community participation. This is probably based on its specific interest in the value for the firm. Although different types of communities are determined as research objects, brand and product-related communities take an important role in marketing research. Because brand communities are centered around specific brands and products, antecedents and consequences of user participation do not only refer to the community itself, but also to these brands (e.g., Casalo, Flavian, and Guinaliu 2008b; Shang, Chen, and Liao 2006). Typical concepts related to marketing and customer relationship management, like satisfaction, loyalty, WOM, trust, and commitment, are well established when researching user participation from a marketing perspective. Nevertheless, marketing and management studies often take a multidisciplinary approach in observing community behavior. Often, early online community studies approached the question why users participate and investigated the effect of different motivations, needs, and benefits on the users' behavior (e.g., Bagozzi and Dholakia 2002; Hennig-Thurau et al. 2004; Wang and Fesenmaier 2004a). Many marketing studies are also based upon social and psychological theories to understand the influences of user participation.

Information systems research is often associated with the acceptance of the technology (the online social community) and the sustainability of these systems. Research in this area deals with the question of what drives people to use and continue using the system in order to understand the success factors of technology. Thereby, researchers tested effects on user behavior, relating for example to the technology acceptance model (Davis 1989; Davis, Bagozzi, and Warshaw 1989) and the information systems continuance model (Bhattacherjee 2001). This includes the interest of several information systems studies in the effects of system quality, information quality, and perceived usefulness (e.g., Jin et al. 2009; Lin 2007;

Yoo, Suh, and Lee 2002). Also of particular importance from this perspective are measures of satisfaction and outcome expectations (e.g., Chen 2007; Hsu et al. 2007; Ma and Agarwal 2007). Online community systems with different focus topics and purposes are studied in information systems research, but often knowledge sharing platforms are used as the research object. Here, the critical success factor is information sharing and retrieving as the relevant form of user participation (e.g., Chen and Hung 2010; Chiu, Hsu, and Wang 2006). Specific aspects of virtual knowledge communities might be accentuated compared to other types of communities, such as social networking sites. For example, it has been shown that an important prerequisite for participation in knowledge sharing communities is the user's self-efficacy and necessary expertise (e.g., Chen and Hung 2010; Hsu et al. 2007; Kankanhalli, Tan, and Wei 2005).

Although, different research disciplines focus on certain aspects more than others, many studies take a multidisciplinary perspective, and consequently overlap to some degree. For example, scholars across disciplines regularly apply social concepts and theories, like social identity theory (e.g., Dholakia, Bagozzi, and Pearo 2004; Ma and Agarwal 2007) or social capital theory (e.g., Kankanhalli, Tan, and Wei 2005; Wiertz and de Ruyter 2007) to explain participation behavior. Because of the nature of online social communities and the interaction and interconnection of its members, it is not surprising that social theories play an important role in many empirical studies. It can be seen that the concepts employed in social theories can be applied to different perspectives regarding online communities.

In order to derive a systematic overview of relevant factors associated with online community participation, specific antecedents and consequences of user participation are discussed in more detail in the following sections. The focus is on the most relevant and most often employed constructs which are directly associated with online community participation.

2.5.3 The Consequences of Participation

From a marketing perspective, user participation is one of the most relevant factors for a successful online community, because it influences user attitudes, intentions and behaviors. Depending on the study context, attitudes and intentions can be related to the community itself or the topics the community is centered around (e.g., brands, products). On an individual user level, several studies confirm, that a higher participation in online communities leads to more favorable attitudes of the community members towards the community and its associated products and services. In brand communities, higher participation has been demonstrated to impact commitment (Casalo, Flavian, and Guinaliu 2008b; Casalo et al. 2009) and loyalty intentions (Algesheimer, Dholakia, and Herrmann 2005; Casalo, Flavian, and Guinaliu 2007; Shang, Chen, and Liao 2006) towards a brand, as well as the image of a brand (Woisetschlaeger, Hartleb, and Blut 2008). In addition, it has been shown that user participation is crucial to ensure the continuity of the community, as it strengthens the relationship to the community and positively affects the loyalty and continuance intentions of its members. Particularly, higher participation leads to higher loyalty intentions (Lin and Lee 2006; Pajuniemi 2009; Woisetschlaeger, Hartleb, and Blut 2008), WOM intentions (Chen and Hung 2010; Koh and Kim 2004; Woisetschlaeger, Hartleb, and Blut 2008), trust (Casalo, Flavian, and Guinaliu 2007), and bonding (Pajuniemi 2009; Yoo, Suh, and Lee 2002) to the community itself.

In addition to attitudes and intentions, the behavior of the community members is also affected by their level of participation. Algesheimer et al. (2010) found that participation in the online community of an online auction site affected the buyers' and sellers' behavior. Another research on members of travel-related online communities revealed that the number of purchased travel products was affected by the users' participation in the online community (Kim, Lee, and Hiemstra 2004). Further, a study by de Valck, van Bruggen, and Wierenga (2009) in a community on culinary matters demonstrated that community interaction, particularly visit frequency and retrieval of information, influenced the members cooking frequency, recipe knowledge, recipe choice, and satisfaction with cooking results. In the context of knowledge communities, it is also shown that participation can help to solve problems at work and improve professional know-how through utilization of knowledge from the platform (Chen and Hung 2010). Altogether, these results provide evidence that behavior and perceptions outside the community are affected by participation in the community.

Although not directly examining the link between users' participation and individual outcomes, further studies underline the importance of online community participation. For example in brand communities, Adjei, Noble, and Noble (2010) found that communications between customers influence purchase behavior through uncertainty reduction. In another case, community participation is found to improve overall product sales through customer ratings on online platforms (Moe and Trusov 2011). This indicates that participation in the online community can lead to relevant transactional behavior and therefore monetary outcomes for firms. In addition, Trusov, Bodapati, and Bucklin (2010) demonstrated that the behavior of some users also affects the behavior of others in the community, showing that logins of the most influential users increased the logins of their contacts. This, in turn, indirectly affects advertising revenues through page impressions.

Theses research studies illustrate the value of online community participation for marketers and firms, as it can produce attitudinal, intentional, and behavioral consequences. Participation affects customer behavior outside and inside the community, which can lead to higher revenues for the involved firms. Additionally, participation also impacts the size, endurance, and attractiveness of the online community itself as it generates loyalty, recommendations and interactivity. Altogether, this underlines the importance of user participation in online social communities.

2.5.4 The Antecedents of Participation

Because of the impact of online community participation on user attitudes, intentions, and behavior, it is critical for marketers and community operators to understand what influences user behavior: what makes the members participate in the online community? Consequently, a plethora of research from different disciplines emerged to explain the main drivers of participation in online communities. A wide range of factors, depending on the research perspective and the type of the online community, have been demonstrated to impact online community participation. Basically, from a marketing perspective the most relevant factors investigated in past research can be subsumed into four broad groups: (1) individual attributes, (2) attitudes and perceptions toward the community, (3) structural aspects, and (4) other factors. Individual attributes include motivations and personal characteristics; attitudes and perceptions are related to the experience with the community and the relationship to the community, which comprises factors like identification, commitment, or satisfaction; structural aspects can be described by the dyadic structures between the users; and the other factors include that of information systems aspects, community characteristics, or external factors. Figure 5 summarizes all relevant direct effects associated with online community participation that are described in this chapter. Appendix 1 provides a detailed overview of all relevant studies by the author, including antecedents and consequences of participation, as well as the research object.

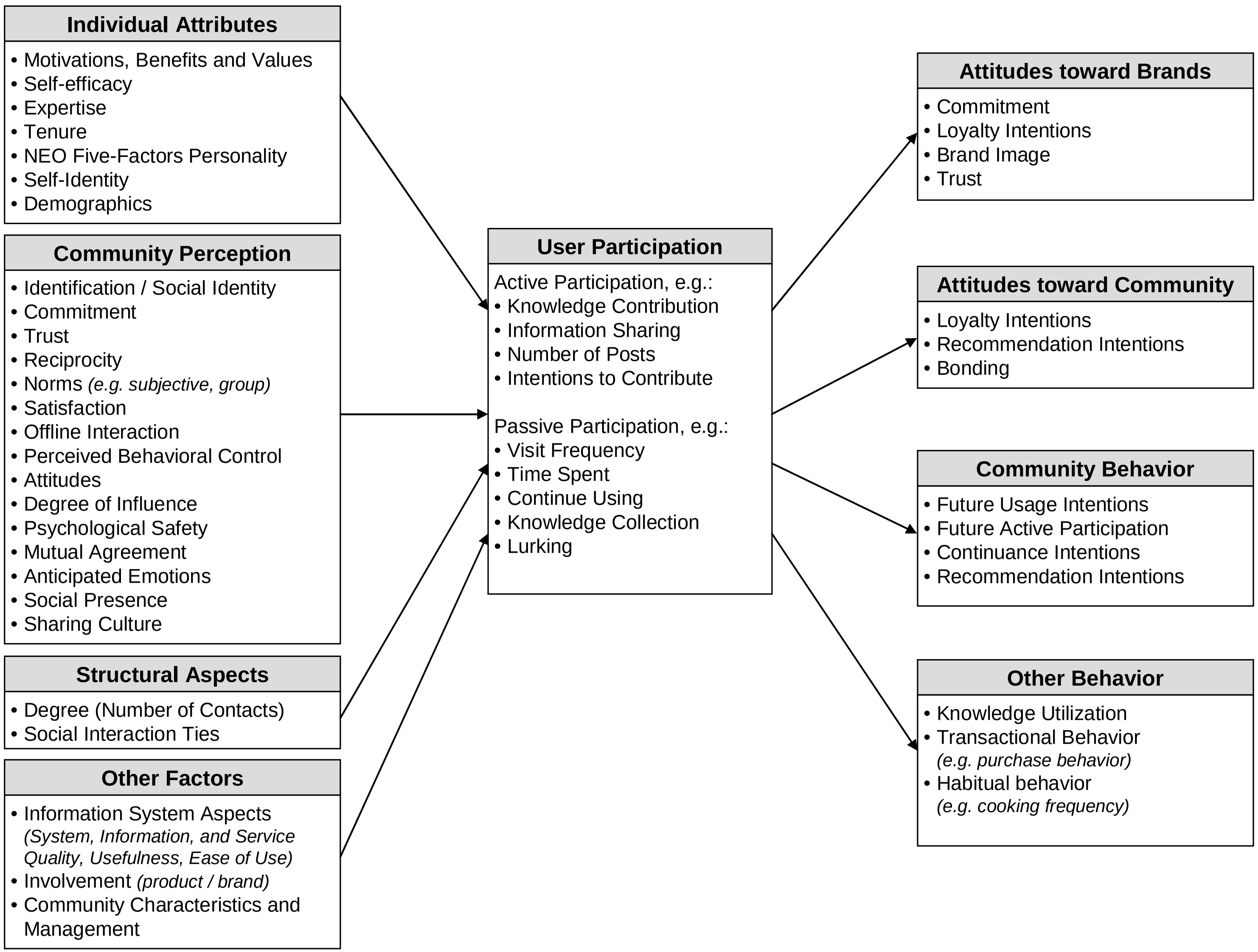


Figure 5: Summary of Relevant Antecedents and Consequences of User Participation

2.5.4.1 Individual Attributes

Benefits and Motivation. Because community participation is predominantly voluntary, there is a need to understand what motivates people to spend time and effort to participate. Exploratory studies gained first insights and found that different motives are at place when users become members of online communities (e.g., Ridings and Gefen 2004; Wasko and Faraj 2000). Motives are the “general drivers that direct a consumer’s behavior toward attaining his or her needs” (Assael 1998, p. 78). Therefore, recent research is interested in its impact on user participation. Also labeled differently as motives, needs, benefits or values, these factors are often interpreted in a similar way as some kind of user motivation. Empirical studies employed a wide range of motives and benefits which are demonstrated to influence behavioral intentions and behavior.

Two factors are of particular importance in online social communities: information benefits and social benefits (Dholakia, Bagozzi, and Pearo 2004; Dholakia et al. 2009). Benefits from information exchange play a key role in influencing participation behavior, especially in knowledge and information-intensive online communities (e.g., Chung and Buhalis 2008; Dholakia et al. 2009; Lampe et al. 2010; Wiertz and de Ruyter 2007). If the user values the information on the platform and is motivated to give or get information his participation is higher. On the other hand, social benefits occur through interaction with other members of the online community and are not solely associated to information exchange. They are based on a more personal level and provide social support and relationships to the community. Several studies highlight that higher social benefits and motivation for connection and interaction with other users lead to higher participation in the online community (e.g., Dholakia et al. 2009; Han, Zheng and Xu 2007; Hennig-Thurau et al. 2004; Lampe et al. 2010; Nambisan and Baron 2007; Wang and Fesenmaier 2004a).

Other factors regularly used to explain user participation include hedonic benefits/ entertainment, enjoyment (to help others), or self-related factors like self-enhancement and self-discovery (e.g., Chung and Buhalis 2008; Hennig-Thurau et al. 2004; Kankanhalli, Tan, and Wei 2005; Lampe et al. 2010; Nambisan and Baron 2007; Wang and Fesenmaier 2004a; Wasko and Faraj 2005; Yu, Lu, and Liu 2010). For some communities, extrinsic benefits like rewards, reputation, or advantages associated with the job may also work as motivators (e.g., Chen and Hung 2010; Kankanhalli, Tan, and Wei 2005; Tiwana and Bush 2005). All these factors were predominantly confirmed to affect participation. In recent studies, motives have also been starkly adapted to the functionality of the community; for example, relating to status updates, social browsing, or using applications, games and quizzes in social networking sites (Joinson 2008). In contrast, Hennig-Thurau et al. (2004) identified platform assistance, venting negative feelings, concern for other consumers, extraversion, social benefits, economic incentives, helping the company, and advice seeking as motives for contribution in product recommendation platforms. This shows that because of the specific context, different sets of measures are defined for some studies. In fact, there are various motives that are unique to certain studies, which are not discussed here because of their specificity (please refer to Appendix 1 which includes all motives covered by studies within this literature review).

Overall, a multitude of studies confirm that motivations drive participation, but the context of the study and the online community of interest must be considered when looking at the results. Results often differ depending on the type and objectives of the community. Dholakia, Bagozzi, and Pearo (2004) demonstrate that in larger network-based communities giving and getting information is a key driver of participation. On the other hand, the authors provide evidence that social benefits are more important in small group-based communities. Further, Dholakia et al. (2009) show that functional benefits have a larger effect on passive participation (helping oneself) than on active participation (helping others), while for social benefits the effect is larger on active participation. This underlines that the different types of participation can also be impacted differently by user motivation. In addition, in every community there are various types of users, driven by their individual motives. Therefore, it can hardly be generalized that all motivational factors have the same effects in every community and on all kinds of participation behavior. Motivational factors need to be viewed in a differentiated way. For this reason, it is important for community operators to know who their customers are, why they are members, and what needs they expect to fulfill in their community. Nevertheless, recent research emphasizes that information and social benefits predominantly take on central roles in explaining user participation, independent of the type of community. In order to succeed, community operators need to address multiple needs, with emphasis on social and information needs, because the focus on one specific need to the exclusion of the others would undermine the value of the online community (Hagel and Armstrong 2006).

Self-Efficacy and Expertise. Besides motivation, the individuals' ability to participate and provide content is also important. Knowledge sharing requires the knowledge to do so. "Perceived self-efficacy is concerned with judgments of how well one can execute courses of action required to deal with prospective situations" (Bandura 1982, p. 122). In knowledge sharing communities it is related to the perception of being able to contribute relevant knowledge to other members (Chen and Hung 2010). Several studies provide evidence that higher levels of self-efficacy lead to higher active participation intentions and behavior (e.g., Hsu et al. 2007; Kankanhalli, Tan, and Wei 2005; Zhang et al. 2010) In the same way, Lampe et al. (2010) showed that the self-efficacy with regards to the community usage has a positive effect on future contributions. Further, Wasko and Faraj (2005) explored the effect of tenure in the field, which means being a member of a professional association, and found that it has a positive effect on the volume of knowledge contribution. This underlines, that expertise is an important factor, at least in knowledge sharing communities.

Tenure. The tenure of membership is also related to participation. In this regard, tenure reflects the user's experience with the online community. Across different types of communities it has been demonstrated that higher tenure is associated with higher contributions and participation intentions (Han, Zheng and Xu 2007; Nambisan and Baron 2007; Nov and Ye 2008; Tiwana and Bush 2005; Wang and Fesenmaier 2004b). De Valck et al. (2007) even show that tenure has an increasing effect on members' visit frequency (quadratic effect).

Other Individual Attributes. Research has also investigated other attributes, but to a lesser degree. For example, Wilson, Fornasier, and White (2010) applied the NEO five-factor personality inventory and found that conscientiousness and extroversion had significant impact on user participation. In particular, participants scoring lower on conscientiousness and higher on extroversion reported spending more time using a social networking site. Further, self-identity – a concept that reflects the extent to which engaging in a behavior is important to an individual's self-concept – was found to positively impact online social networking intentions and behavior (Pelling and White 2009). In addition, demographic information is often used as control variables in the empirical online community studies. Therefore, the variables of age, gender, education, occupation, or nationality are used.

2.5.4.2 Community Perceptions and Attitudes

Identification and Sense of Belonging. Identification with the community and social identity are among the most frequently studied constructs associated with online community participation. According to social identity theory, social identification is the perception of belonging to a group with the result that a person identifies with that group (Bhattacharya, Rao, and Glynn 1995). It therefore represents the association of oneself with a group of members or the online community as a whole. Bagozzi and Dholakia (2002) and Dholakia, Bagozzi, and Pearo (2004) find that an individual's social identity leads to higher intentions to interact again with a group of people in a virtual community. The positive effect of identification on participation is also found across different types of communities, be it either knowledge sharing communities (Chiu, Hsu, and Wang 2006), brand communities (Woisetschlaeger, Hartleb, and Blut 2008), social networking sites (Cheung and Lee 2010; Han, Zheng, and Xu 2007), or consumer communities (Kim, Lee, and Hiemstra 2004).

Closely related to social identity and identification is the sense of belonging. It is often directed and operationalized more in the sense of the affective component of social identity. Even so, it has a positive effect on active and passive participation behavior as well (Lampe et al. 2010; Lin 2007). The feelings of membership, as part of the sense of community concept, are also positively related to user participation (Kim, Lee, and Hiemstra 2004; Yoo, Suh, and Lee 2002; Zhang 2010). Altogether, this emphasizes the usefulness of social identity to explain participation and participation intentions in online communities.

Commitment. Commitment to the relationship is defined as "an enduring desire to maintain a valued relationship" (Moorman, Zaltman, and Deshpande 1992, p. 316). It is a central construct in relationship marketing literature, as exchange partners believe that an ongoing relationship is important and that it is worth working at maintaining it to ensure long-term endurance (Morgan and Hunt 1994). Customers with high commitment to the service provider are more loyal and show favorable behavior towards firms and organizations (e.g., Bettencourt 1997; Garbarino and Johnson 1999; Gruen et al. 2000; Hennig-Thurau, Gwinner, and Gremler 2002).

In recent studies of online communities, commitment to the community is often related to the members' identification with the community and a sense of belonging. This refers to the affective commitment and overlaps to a certain degree with other constructs, when it is defined and operationalized with elements like belonging, attachment, and willingness to interact with the exchange partner (e.g., Kim, Choi, and Han 2004; Lampe et al. 2010; Kang et al. 2007). Nonetheless, researchers revealed a positive relationship between commitment and participation in online communities (Cheung and Lee 2009; Nov and Ye 2008; Wiertz and de Ruyter 2007; Xie, Chen and Wu 2008). To some extent, this confirms the findings of the positive effect of identification, but also shows that it is a relevant driver of online community participation per se. If a user is committed to an online community, he wants to stay in that relationship with the community and is willing to put forth effort, for example in the form of active participation, to maintain it.

Xie, Chen and Wu (2008) differentiated between different types of commitment, namely affective, continuance and normative commitment, and demonstrated that all aspects had a positive effect on user participation, though on different types of participation.[17] Affective commitment was found to influence participation intention and contribution intention, normative commitment affects contribution intention, and continuance commitment is positively related to advocacy intention. Although this confirms the impact of commitment on participation in general, a more differentiated view on different types of participation reveals, in this case, that affective and normative commitment are important for active participation, and affective and continuance commitment for passive participation. In addition, Wiertz and de Ruyter (2007) even found that the direct effect of commitment on knowledge contribution is moderated by online interaction propensity, suggesting that commitment builds over repeated interactions with other users, being more important when the need for interaction is higher.

Trust. Trust has been proposed as an important concept in relationship marketing (e.g., Morgan and Hunt 1994), also taking a central role in the online context (e.g., Shankar, Urban, and Sultan 2002). Moorman, Zaltman, and Deshpande (1992, p. 315) define trust as “a willingness to rely on an exchange partner in whom one has confidence”. In relationship marketing, trust has been acknowledged to be strongly related to commitment, so that trusted relationships are found to lead to higher loyalty intentions (e.g., Garbarino and Johnson 1999; Morgan and Hunt 1994). Research on the effects of trust in online communities adapted and proposed different types of trust which work in the relationship to the community. For example, Chen and Hung (2010) found that interpersonal trust, which relates to trustworthiness and honesty of all members in the community, positively affects knowledge contribution and collection. Hsu et al. (2007) demonstrated the impact of identification-based trust, which is defined as members’ trust due to emotional interaction among members. Another conceptualization of trust regards it as a multidimensional construct including the elements of ability, benevolence, and integrity, which altogether showed a significant positive effect on active participation (Casalo, Flavian, and Guinaliu 2008a; Casalo, Flavian, and Guinaliu 2008b),

[17] The three forms of commitment are adapted from Meyer and Allen’s (1991) organizational commitment conceptualization. Affective commitment refers to an individual’s emotion attachment to, identification with, and involvement in an organization. Normative commitment reflects an individual’s felt sense of obligation to continue employment. Continuance commitment focuses on the consequences of an individual’s awareness of the costs associated with leaving an organization.

and on the desire to give and to get information (Ridings, Gefen, and Arinze 2002).[18] Overall, trust measurement in online communities is mainly directed towards the community or the group of members as a whole. Thus, recent studies found this generalized trust to affect user behavior (Han, Zheng and Xu 2007; Zhang et al. 2010).

Trust is regarded with specific significance in online communities, because of the anonymity and the lack of face-to-face contact between unknown members (Ridings, Gefen, and Arinze 2002). Trust is especially of high relevance in information-intensive communities like knowledge sharing communities. If the user cannot trust the received information and the information sources, there is no value in the exchange of know-how. Nevertheless, research also provided evidence for its importance in other types of communities, like brand communities (e.g., Casalo, Flavian, and Guinaliu 2008b; Shang, Chen, and Liao 2006).

Reciprocity. Reciprocity describes the process of give and take. It generally refers to the expectation of one party, who provided some kind of support or information to another party, that the other party returns the favor (Wu et al. 2006). Reciprocity is thereby an obligation to repay the benefits received (Gouldner 1960). Ridings, Gefen, and Arinze (2002) argued that online communities will hardly survive without reciprocity, because contributions are most important to keep the community attractive, and individuals who post content mainly expect some type of response. Surprisingly, most research did not find a significant positive relationship between the norm of reciprocity and online community participation (Chen and Hung 2010; Kankanhalli, Tan, and Wei 2005; Wiertz and de Ruyter 2007). The only study that supports the positive effect of reciprocity on the quantity of knowledge sharing was conducted by Chiu, Hsu, and Wang (2006). Chen and Hung (2010) and Wasko and Faraj (2005) even found negative effects on knowledge collection and knowledge contributions, respectively. They argued that knowledge seekers have no control over who responds to their question and that there is a possibility that in online communities a generalized reciprocity exists, which assumes no direct return by the recipient, but by the community as a whole. In another knowledge sharing community, Kankanhalli, Tan and Wei (2005) examined no significant direct effect, but a moderated effect of reciprocity contingent on pro-sharing norms. When having high pro-sharing norms, indicating a climate of collaboration and cooperation, knowledge contributors do not look for reciprocity when contributing their knowledge, but low lev-

[18] Ability is described by the skills or competencies that an individual has in a certain area; benevolence refers to the expectation that others will have a positive orientation to do good to the trustee; integrity is the expectation that others will act in accordance with socially accepted standards of honesty or a set of principles that the trustor accepts (Ridings, Gefen, and Arinze 2002).

els of pro-sharing norms emphasize the reciprocity benefit as a motivator for knowledge contribution. Further, Wiertz and de Ruyter (2007) observed in their study that the positive effect of reciprocity on the quantity of knowledge sharing was suppressed by the users' propensity for online interaction, i.e. their overall positive attitude towards interaction. Altogether, the mixed results of past research indicate that there might be a more complex relationship between reciprocity and other factors and that reciprocity alone can hardly explain higher user participation in many online communities.

Norms. Basically, a norm represents a degree of consensus in the social system (Coleman 1990). Such norms are related to a relevant group of other people and are a potential source of social influence for an individual. Different types of norms have been researched in conjunction with online community participation. Among them, subjective norms and group norms are frequently investigated. Subjective norms reflect social pressure from significant others to perform or not perform a certain behavior, while group norms represent the shared values or goals perceived by the individual between oneself and other members of the online community (Bagozzi and Dholakia 2002). Overall, the effects of norms on user participation revealed inconsistent results across different studies. While subjective norms were found to have a negative effect on participation intentions of the group in one study (Cheung and Lee 2010), another study showed that subjective norms positively predict usage intentions and do not significantly affect participation behavior (Pelling and White 2009).[19] In other communities, subjective norm did not show a significant impact on user participation intentions (Bagozzi and Dholakia 2002; Lin 2006). These mixed results might be associated with the different online community types researched.

Apart from subjective norms, group norms show a tendency of positively affecting participation intentions, although not all studies reveal significant effects (Cheung and Lee 2009; Cheung and Lee 2010; Dholakia, Bagozzi, and Pearo 2004). Higher congruence between one's goals and other members' goals lead to higher participation intentions. In addition to the discussion of subjective and group norms, Nambisan and Baron (2007) revealed that general community norms, i.e. a strong value of interaction in the community, positively influence future participation. Further, pro-sharing norms, defined by Kankanhalli, Tan, and Wei

[19] Similarly, normative community pressure refers to the customer's perceptions of the community's extrinsic demands on a person to interact and cooperate with the community (Algesheimer, Dholakia, and Herrmann 2005). This pressure to conform to the community's norms and objectives has been found to have a negative effect on participation intentions and through reactance on continuance intention in (offline) brand communities (Algesheimer, Dholakia, and Herrmann 2005). If the user is too heavily restricted and coerced by the community, the user can be pressured to become inactive or de-register

(2005) as norms to cooperate, collaborate and share opinions and ideas openly, did not show direct effects on knowledge sharing and played only a moderating role in conjunction with reciprocity (see above). Given the inconsistent results in studies to date, it is suggested that norms do not offer a congruent explanation of participation behavior.

Satisfaction. In marketing research, satisfaction has developed to one of the most frequently researched constructs, and it is an important antecedent of consumer behavior like WOM provision, complaint behavior, retention and repurchase (e.g., Gustafsson, Johnson, and Roos 2005; Szymanski and Henard 2001; Wangenheim and Bayón 2007). In online community studies, satisfaction is also regularly deployed to explain user participation, where it is considered a key driver of members' loyalty to the online community. De Valck et al. (2007) show that satisfaction as an overall evaluation of the performance of the online community has a positive effect on visit frequency. They also suggest that satisfaction in online communities is related to four different interaction dimensions: satisfaction with member-to-member interactions, organizer-to-member interactions, organizer-to-community interactions, and with the community site. Though, organizer-to-community interactions do not significantly impact visit frequency. From an information systems perspective, Lin and Lee (2006) confirmed that satisfied needs positively impact community participation and loyalty intentions.

Several other studies investigated the relationship between user satisfaction and user participation in online communities. Basically, most research confirms that highly satisfied users show higher levels of participation. The positive effect of overall satisfaction is found for active participation like knowledge contributions and perception of active participation (Casalo et al. 2009; Ma and Agarwal 2007; Zhang et al. 2010), as well as passive participation like visit frequency and intention to continue using the community (Chen 2007; Cheung and Lee 2009; Jin et al. 2009; Lampe et al. 2010; Tiwana and Bush 2005; Woisetschlaeger, Hartleb, and Blut 2008). Altogether, the majority of studies across different types of communities, like knowledge sharing communities, bulletin boards, or open source software communities, confirm the positive impact of satisfaction.

Offline Interaction. Because online and offline environments can intertwine, some researchers have investigated the impact of offline interactions on online participation. Ma and Agarwal (2007) found a positive relationship between offline interactions and knowledge contribution. Koh et al.'s (2007) results reveal that offline interaction has a positive effect on posting behavior, but not on viewing behavior, suggesting that contact in the offline world can stimulate active participation online. In another study, Lin (2007) demonstrated that offline interaction has an indirect and positive effect on behavioral intention mediated through sense of belonging.

Further Attitudes and Perceptions. In addition to these well researched antecedents, certain studies also used other factors to explain user participation. Although they have been less frequently researched, some relevant factors are described here briefly. ***Perceived behavioral control*** has been tested in a few studies (Bagozzi and Dholakia 2002; Lin 2006; Pelling and White 2009). It is defined as individual perceptions of the ease of participating in the online community and an individual's control to do so. Only one study revealed a significant positive effect on participation (Lin 2006). In line with the theory of planned behavior, more positive ***attitudes*** towards the participation in the online community resulted in higher participation intentions (Lin 2006; Pelling and White 2009). Additionally, Nambisan and Baron (2007) found that the attitude towards the host firm can also have a positive effect on participation. ***Degree of influence*** is another factor researched. Woisetschlaeger, Hartleb, and Blut (2008) showed that a higher degree of influence, that is the perception of being able to influence and shape the community, leads to higher perceptions of participation.

Additional factors found to have an impact on user participation include mutual agreement (Dholakia, Bagozzi, and Pearo 2004), anticipated emotions (Bagozzi and Dholakia 2002), psychological safety (Zhang et al. 2010), social presence (Shen and Khalifa 2008), and sharing culture (Yu, Lu, and Liu 2010) These factors are not described in greater detail here. Please refer to Appendix 1 and the work of the respective authors to get more information on these constructs.

2.5.4.3 Network Structure

As the interconnection between users is more visible in online social networks of today, there is a growing interest in the impact of the online network structure on behavior. Though, only a few studies investigated the effect of structural components on online community participation. Social theorists posit that the position of a user in a social network, in this case in the online social community, can have an important impact on the user's behavior because of his access to certain resources and information (e.g., Freeman 1978/79; Nahapiet and Ghoshal 1998).

Degree. The degree of an individual is described as the number of contacts he has. Especially in management literature, it has been shown that degree centrality is a strong predictor of individual performance in workgroups (e.g., Ahuja, Galletta, and Carley 2003; Sparrowe et al. 2001). Similarly, community researchers demonstrated that a higher number of contacts (the degree of the user) is associated with higher knowledge contribution and a higher number of posted content like photos (Nov and Ye 2008; Wasko and Faraj 2005). This indicates that more friends lead to more participation.

Social Interaction Ties. Another structural component is the social interaction ties the users maintain in the online community. Such ties consider a bond between two people or an overall group of people in the community and are based on one or more relations between them. Chen (2007) measures the users' perceptions of their social interaction ties to other members of their community to explain user behavior. Thereby, she operationalized social interaction ties as a self-reported assessment of the users' relationship and interaction with other users. She found that social interaction ties significantly influence the users' continuance intentions to use the online community. Using a similar measure of social interaction ties, Chiu, Hsu, and Wang (2006) provided evidence for its positive impact on the quantity of knowledge sharing.[20]

Although other authors contributed additional insights on how network structure affects the success of online communities (Toral et al. 2009), they took an aggregated perspective rather than focusing on individual user behavior. Overall, this low number of relevant studies dealing with network structure as a factor influencing user participation suggests that there is a need for further research in this area.

2.5.4.4 Other Factors

Other factors include those that are related to the community information system, community external factors like involvement in a product or brand (the topic of the community), or community characteristics.

[20] In these two studies, social interaction ties are operationalized as an evaluation of the different aspects of the relationship to other members in the online community. Thereby, the measure is of more qualitative nature compared to objective data on user positions within the online social network, which can be calculated by social network analysis. Social network analytical measure that can be used to evaluate the network structure are presented in chapter 3.2.2.

Information Systems Aspects. Specific to information systems research are factors addressing the quality and usability of the system and the associated information provided on the platform. Therefore, several studies from this area have incorporated such technical aspects. Here, researchers often base their empirical work on concepts like the technology acceptance model, the information systems continuance model, or the updated information systems success model (Davis 1989; DeLone and McLean 2003; Bhattacherjee 2001). Lin and Lee (2006) demonstrated that three dimensions of information systems quality, namely system quality, information quality, and service quality, are all positively associated with behavioral intentions of community use.[21] Yoo, Suh, and Lee (2002) confirm the positive impact of information system quality, and especially of information quality, on online community visit frequency and usage time. The usefulness of the information provided on the platform is also documented to significantly affect passive participation like viewing activity (Koh et al. 2007), continuance intention to use the information on the platform (Jin et al. 2009), as well as active knowledge sharing (Yu, Lu, and Liu 2010). Several other researchers provided evidence that information system aspects, like information and system quality, usefulness and ease of use are also mediated through other factors, such as satisfaction, attitude or sense of belonging (e.g., Chen 2007; Lin 2006; Lin 2007). Overall this emphasizes the key role quality aspects take in online communities to establish favorable attitudes and perceptions toward the community, which result in user participation intentions and behavior.

Involvement. Further, involvement of the user in the topic of the online community can have a significant impact on user behavior. One study in the context of an Apple computer community provided evidence that cognitive involvement in Apple products led to higher lurking behavior (Shang, Chen, and Liao 2006). Additionally, another study within open source software (OSS) communities showed that involvement in the respective OSS project, which is the topic of the community, positively influenced the contribution of knowledge to the projects in the community (Xu, Jones and Shao 2009). This means, if the users' interest in the community topic is very high, this can have an additional impact on their participation in the online community.

[21] System quality includes, system reliability, convenient access, ease of use and system flexibility; information quality describes accuracy, timeliness, usefulness, completeness and customised information; and service quality is measured by interface design presented to members, trust mechanisms provided by the online community, and willingness to help members and provide prompt service (Lin and Lee 2006).

Community Characteristics and Management. Community characteristics can include very different elements like size, age, type of community, composition of member base, or culture. For example, Chu (2009) included community characteristics in the user survey, including questions on size, diversity of members, if the community provides ancillary resources, and role of members (peripheral or central), and found that all four dimensions positively affect members' helping behavior. In another study, Yoo, Suh, and Lee (2002) provided evidence that the community's managing strategy in the form of its purpose, rules, events, and subgroups, partly influences the visit frequency. In a series of experiments, Ling et al. (2005) demonstrated how community management can affect participation. They found that reminding community users about the uniqueness of their contributions leads to increased participation. Further, they also showed that users contribute more when they are given group goals compared to individual goals in the community. The results emphasize that proper community management is important and can result in higher participation.

2.5.5 Discussion of Past Research on Online Community Participation

The presentation of the empirical findings of recent online community research provides valuable insights on the constituents, antecedents and consequences of user participation. Figure 5 summarizes all relevant direct effects associated with participation that are described in the previous sections and Appendix 1 provides a detailed overview of all relevant studies.

Overall, factors of different dimensions have been found to affect user participation. Members' behavior is influenced by 1) the user's attributes, 2) attitudes and perceptions of the relationship to the community, 3) structural aspects, and 4) other factors like system aspects, involvement and community characteristics. For the most relevant factors, investigated in several studies, the majority of results are rather consistent across different study contexts and different online communities. This emphasizes their impact on user participation. Within the group of factors concerned with individual attributes, motivations are most often researched, because they are of high relevance for understanding why users participate in online communities. Social and information benefits have been identified to be of particular importance. From a practical perspective, community operators need a deeper understanding of the users' motivations and needs in order to stimulate a more active behavior in the community. Moreover, variables related to the users' perception of his relationship to the community have been found to take a key role in determining the users' participation. Particularly, identification, trust, reciprocity, and norms are all elements of the relational dimension of social capital, which is a well established theory for explaining performance and behavioral outcomes (Nahapiet and Ghoshal 1998; for an introduction to social capital please refer to chapter 3.3). Altogether, these elements are positively related to participation, which is also confirmed by studies which used aggregated measures of relational and social capital (Chu 2009; Tiwana and Bush 2005). Although structural aspects have been researched to a lesser extent, taking the position and relationships to other members into consideration yields further explanation of user behavior. Overall, online community operators benefit from a thorough knowledge of the drivers and consequences of online community participation, so that they can more effectively manage user behavior in their communities. A deeper understanding of the relevant factors can help them better allocate their resources to stimulate participation and manage user behavior.

Despite the usefulness of these findings from past research, some issues of these research studies should be discussed at this point:

Measurement of Participation. A large share of the reviewed studies used self-reported instead of objective data for the investigation of user participation. Although self-reported data can provide some useful insights, objective data is preferred because it accurately reflects the true user behavior. Moreover, the studies use different operationalizations and types of participation, for example active vs. passive participation. These different perspectives on participation provide valuable insights by employing distinct types of participation. Nevertheless, the inconsistency in operational definitions and measurement of online community participation within academic literature may hinder the comparison across studies. For example, the results of different studies utilizing various participation measures are not completely comparable due to the diverse contexts and purposes of the online communities observed. But studies that include both kinds of participation can provide valuable insights on how certain factors affect each type of participation. Recent studies found that in many cases active and passive participation are influenced differently or at least to a different extent (e.g., Dholakia et al. 2009; Hennig-Thurau et al. 2004; Koh et al. 2007; Lampe et al. 2010; Shang, Chen, and Liao 2006). For example, Dholakia et al. (2009) found that the effects of functional and social benefits on both participation variables, helping oneself and helping others, differ in their significance and strength. In particular, social benefits show a much stronger effect on helping others than on helping oneself in the consumer sample, and in the B2B sample the effect of social benefits on helping oneself is even insignificant. This suggests that social benefits have a much stronger impact on helping other members (i.e. on contributing) than on consuming information. Other studies show further differences for the effects of certain variables. Koh et al. (2007) even argue that “in any given community, the posting activity stimulant is not the same as the viewing activity stimulant”. The comparison of the effects on different participation behavior is therefore very useful for community operators when planning to take marketing measures to increase participation.

Another issue is that participation can consist of different elements. Because of the dynamic development of Internet technology, online communities steadily advance in their functionalities. The core elements of online communities remained similar with the interaction between users at the heart of these online services. However, the types of possible interaction increased, so that today many different ways are available to exchange information and social support. The reviewed online community studies often take a more general view on user participation or focus on knowledge contribution and posts. Because basically all communities offer functionality to communicate with others, most of the past studies seem to be rather comparable. The question emerges if new functionality leads to changes in user behavior. New functions such as playing games or using apps could be affected by different factors than communicating and posting. For example, playing games would potentially be more related to entertainment needs than to information needs, while consuming and contributing knowledge is more associated with information than entertainment needs.

Because of the diverse measurement of online community participation used by different researchers, multiple measures should be used in empirical investigations. This would allow the researcher to analyze the data in multiple ways, compare the results with other empirical studies and can test for the consistency of the findings. Different operationalizations can also help to test for the robustness of the results. Therefore, objective data should be used whenever possible. Preferably, both active and passive participation should be compared, if the study focus is directed towards factors associated with both types of participation. For example, when comparing user groups in their participation behavior, this could be done using the number of visits, the time spent on the platform, and active contributions. Further, multidimensional measures can be broken down in more constituents, and can be aggregated to different levels.

Antecedent Effects. Although using different operationalizations for user participation, several different studies show similar results for certain effects, i.e. similar significant effects of the antecedents on user participation. This finding emphasizes the impact of such factors like identification, commitment, trust, satisfaction, attitudes, system aspects, or involvement.

However, some of the reviewed studies reveal significant inconsistencies in their results of certain antecedents across studies. For example, the effects of reciprocity and norms on user participation are partly significant, partly insignificant, and in some cases even in the other direction (negative sign). This is rather confusing. Such results might suggest that the relationship between those predictors and participation are less stable and most likely are dependent on other factors, like the context of the community. In case of more complex relationships further investigation of those factors is needed including additional variables and interactions with other factors to gain a better understanding about how they work. For example, Kankanhalli, Tan, and Wei (2005) found that reciprocity does not have a direct effect, but a moderated effect contingent on pro-sharing norms. In specific situations, moderating and non-linear effects might therefore play a more important role in future research.

Gaps in Researching Structural Aspects. The literature review demonstrates that specifically structural components are merely underrepresented in past research. With regards to individual participation in the community, most studies focused on the users' number of contacts or the overall strength of social ties to other users in the community as the only structural measures (Chen 2007; Chiu, Hsu, and Wang 2006; Nov and Ye 2008; Wasko and Faraj 2005). Some scholars studied the effect of network structure on community success on a network level (Toral et al. 2009; Toral, Martínez-Torres, and Barrero 2009). However, the explicit constitution of the users' networks of contacts in conjunction with their attitude towards the community has not been studied as predictors of active user participation. This is rather surprising, as the very nature of online communities are the interconnections among their members. Thus, there is a need to incorporate objective social network analytical measures, like centrality, ego-network density, or the offline-online configuration, in studies on online community behavior. It is expected that a central position in the network will have a significant effect on a user's behavior. In fact, in this dissertation the value of studying such social network factors is demonstrated in two empirical studies (cf. chapters 6 and 7).

Performance Outcomes of User Participation. Antecedents of user participation have been more popular in academic literature. With respect to the different business models presented in chapter 2.3, future research could focus on the consequences of online community participation. Recent studies rarely investigated the financial outcomes of participation. Because many communities incorporated advertising based business models, there is a direct link between participation and advertising revenues, because more active users also generate more page impressions and thereby their exposure to online ads.[22] Nevertheless, in other business models, community operators need to convert participation into transactions, clicks, membership sign-ups, or innovation contribution. A direct link of participation or outcomes of participation to objective financial returns is still missing in the online community context.

2.5.6 Research Prospects for the Empirical Study

From a marketing perspective, online social communities provide a rich context for studying user behavior in the online domain. As online communities have gained more importance in the users' lives, understanding how and why they behave in certain ways is one of the main tasks for practitioners and scholars. It is pointed out, that online community participation is the most critical and most focused aspect of online community management and research. In this comprehensive literature review, the most relevant factors associated with online community participation are explained. This already provides valuable insights to marketers and community operators on how participation can be stimulated.

Nevertheless, this overview also helped to identify some issues that can be addressed in future research. In the empirical part of this dissertation, several of the limitations of past studies are addressed. It is not possible to address all research gaps at once, however four issues of high relevance will be considered:

(1) Objective measurement of participation: Many recent studies are based on self-reported measures of user participation. As already suggested, objective data is more accurate and unbiased. Therefore, objective measurement of active participation is used in all empirical studies of this dissertation (cf. chapters 5, 6, and 7). In addition, active and passive participation are taken into consideration in the course of the three studies, which helps to compare certain results and identify potential differences.

(2) The effect of customer acquisition: The literature overview clearly shows that the effects of user acquisition on user participation have not been studied so far. Study 1 of this dissertation provides insights on how the user acquisition channel affects the attitudes and participation of members in their post-adoption phase (cf. chapter 5).

(3) The impact of social network structure: It is shown that recent research only superficially studied the effects of social network positions of users in online communities as an in-

---

[22] One example of the financial impact of user participation is provided by Trusov, Bodapati, and Bucklin (2010). They estimate the value of an average user per year based on advertising revenues to be around US$ 1.50. However, they also argue that the value of the users depends also on their influence on the participation of other users.

fluencing factor of user participation. Using objective data, which is easy to retrieve for community operators can help to further explain how user participation is affected. Therefore, a more comprehensive set of network measures is used to investigate active user participation in this dissertation (cf. chapter 6).

(4) Investigation of user defection: The loyalty of users in online communities is mainly measured through intentions to continue using the platform. Objective measures for loyalty are not considered in recent research. Because community operators need to focus on customer retention, an interesting research question would be the reasons why users leave the online community. Therefore, studying the loyalty of online community users incorporating behavioral data would provide additional insights. In study 3 of this dissertation (cf. chapter 7), longitudinal data is used to explore the influence factors of user defection, i.e. that users do not return to the platform. In study 3, the network factors used in study 2 are also included, which helps to compare their effects on different types of user participation – active participation and usage participation in the sense of staying with the community.

This dissertation is proposed to address some of the identified research gaps within the context of online communities and more specifically online community participation. In the following chapters, relevant theories are discussed which build the basis for the three empirical studies that investigate the users' network positions, attitudes and behavior. After the theoretical underpinnings the empirical studies are presented in detail.

3 The Social Context of Communication and Relationships in Online Social Communities – Theoretical Background

In order to lay out the theoretical basis for the research questions and the hypotheses development in the empirical study, this section introduces different theories and concepts related to the social context of online communities. First, interpersonal communication modes as means to acquire new members are presented. Second, social network concepts are introduced. These concepts play a significant role to determine the network structure and position of individual users in the online community. Third, social capital theory is described, which builds an important framework for studying influencing factors of online community participation. Last, additional concepts that are related to the connection and interaction of actors in the social context of the community are briefly discussed. These include the concepts of social identity, social exchange, social presence and collective behavior.

3.1 Interpersonal Communication Channels for User Acquisition

Marketing literature promotes various marketing mix elements which may help to transmit marketing messages to existing and prospective customers (e.g., Borden 1964). It ranges from mass market advertising (e.g., TV, radio, print) to personal selling, and from information on websites to consumer-initiated word-of-mouth recommendations on the Internet (Chen and Xie 2008; Duncan and Moriarty 1998). Interpersonal communications play an important role in attracting new customers as it has been demonstrated that they are effective channels for influencing customers (e.g., Herr, Kardes, and Kim 1991; Price and Feick 1984). Many online communities and social networking sites like Facebook, LinkedIn, or StudiVZ grow their customer base through these channels. Two interpersonal communication channels are of specific interest: personal selling, as a form of employee-to-customer communication, and word-of-mouth recommendations, as a form of customer-to-customer communication. Both forms work in a similar fashion. The sender of the marketing information approaches the potential customer and transmits a message containing relevant information for the receiver. Thereby, sender and receiver have a relationship between each other concerning the information about the promoted offering and the fulfillment of needs. Because these two forms are regarded explicit attention in the empirical study of this dissertation they are described here to gain a general understanding of these communication channels.

3.1.1 Personal Selling as a Marketing Channel

For long, personal selling has been an essential part of the firm's marketing mix (Borden 1964) and can be viewed as an important element in promoting products and services. Traditionally, personal selling includes a presentation of arguments within a conversation between the sales representative and one or more potential buyers with the goal of selling products and services (Meffert 2000). Therefore, selling "is a process whose success depends on the salesperson properly identifying and satisfying the needs of the customer" (Szymanski 1988, p. 65). Personal selling can play an important role in relationship building with the goal of creating satisfied and long-term customers (e.g., Reynolds and Beatty 1999; Solomon et al. 1985). Person-to-person communications with sales representatives are advantageous because the sales person can provide information instantly, address individual needs, and build rapport (Barlow, Siddiqui, and Mannion 2004). Additionally, Palmatier et al. (2006) found that relationship marketing is more effective when relationships are built between the customer and an individual person like, the sales representative, rather than a selling firm, which favors personal selling over mass media communication.

Marketing research has mainly observed explanatory factors related to salesperson performance (e.g., Churchill et al. 1985; Szymanski 1988; Webster 1968), that is the selling of products and services. But the effectiveness of personal selling cannot only be measured by the fact that the customer adopts a service or product. With the development of the salesperson from a role of selling products to the customer to a role of partnering with the customer and building relationships, both attitudes and behaviors of the customer are important components of the relationship quality with the selling firm (Weitz and Bradford 1999). In the context of online social communities, this means that post-adoption attitudes and behavior are relevant indicators for a satisfactory relation of new users with the community because higher participation constitutes the value that is brought by the user to the platform. In the empirical study of this dissertation, these attitudes and behaviors are evaluated for users coming from the personal selling channel and compared to word-of-mouth (WOM) referred users in order to understand the effectiveness of personal selling.

For local online communities, personal selling can play an important role for attracting new customers to use the service. For example, communities offering thematic information on sports could contact prospective users in local sport stores or at games of the local soccer teams in order to promote the online service. Frequent and repeated contacts may thereby foster the relationships with the users, which can help to motivate them to use the platform more intensively and keep the interest in the platform alive. Therefore, interpersonal communication between the service operator and the customer might be an effective tool for value generation.

3.1.2 Word-of-Mouth Recommendations as a Marketing Channel

In contrast to commercially driven, firm-initiated communication WOM recommendations are customer-initiated and in less control of the firm. WOM communication has received broad attention from researchers and practitioners because of its effectiveness in acquiring new customers (e.g., Engel, Kegerreis, and Blackwell 1969; Herr, Kardes, and Kim 1991). Harrison-Walker (2001, p. 63) defines WOM as "informal, person-to-person communication between a perceived noncommercial communicator and a receiver regarding a brand, a product, an organization, or a service“. The main difference to personal selling is the non-commercial purpose of the communication. It has already been studied, that in the decision making process firm-initiated communication is usually seen as less credible than WOM (e.g., Arndt 1967; Murray 1991). Therefore, WOM communication is an effective alternative to firm-initiated marketing efforts. It has developed into an important tool for acquiring new customers and increasing value to the firm (Villanueva, Yoo, and Hanssens 2008; Wangenheim and Bayón 2007). Here, three perspectives on the WOM communication process are identified: (1) factors influencing the WOM sender to give recommendations, (2) factors influencing the attitudes and behavior of the WOM receiver, and (3) the importance of the social relationships between WOM senders and receivers.

3.1.2.1 Different Perspectives on the Word-of-Mouth Communication Process

The Sender Perspective – how WOM giving is influenced. Usually, WOM sending is voluntary and without any initial profit motivation. The importance of WOM senders is obvious; they contribute directly to the acquisition of new users to the platform. One main question addressed in marketing literature is what drives WOM provision, i.e. what factors influence the proactive spreading of the word about products and services (e.g., de Matos and Rossi 2008). Several factors have been identified that have significant impact on WOM intentions and recommendation behavior. In a meta-study of antecedents of WOM activity, de Matos and Rossi (2008) found significant influences of satisfaction, loyalty, service quality, commitment, trust and perceived value on WOM activity, using data from multiple research papers.[23] Other relevant factors affecting WOM activity and its influence include the characteristics of the WOM sender (e.g., Bansal and Voyer 2000; Gilly et al. 1998; Wangenheim and Bayón 2004b) and motivational factors (e.g., Hennig-Thurau et al. 2004). In online community research, it has been demonstrated that active user participation is a significant antecedent of WOM activity (e.g., Chen and Hung 2010; Woisetschlaeger, Hartleb, and Blut 2008). Overall, these studies provide insights on what influences the provision of WOM. In order to stimulate WOM the mentioned factors should be addressed to increase WOM communications about the firm's services.

The Receiver Perspective – how WOM giving influences customers. Other research has concentrated more on the effects of WOM reception. Here, positive WOM has been found to have a significant impact on product adoption as well as customer attitudes and perceptions (e.g., Arndt 1967; Bone 1995). Thereby, customers benefit from retrieving more trustworthy and reliable information that can be used in decision making and usage of products and services (e.g., Arndt 1967; Murry 1991). It is an effective means to acquire new customers (Wangenheim and Bayón 2007); some studies even show that WOM communication is more effective than traditional advertising or media messages (e.g., Engel, Kegerreis, and Blackwell 1969; Hennig-Thurau and Walsh 2004; Herr, Kardes, and Kim 1991; Katz and Lazarsfeld 1955). Therefore, it represents a cost-efficient way to grow the customer base of a company. Further, WOM referred users are found to be more loyal and they generate WOM referrals themselves (e.g., Gilly et al. 1998; Wangenheim and Bayón 2004a).

Overall, WOM can affect the customer in different phases of the decision making process. In order to better understand how WOM reception can influence the receiver's attitudes and behaviors, an overview of the adoption process of the receiver is helpful. De Bruyn and Lilien (2008) use three stages to describe possible effects of WOM in this process: (1) the awareness stage, (2) the interest stage, and (3) the final decision. In the case of online social communities, the user would become aware of the service by getting WOM from other users. In the interest stage, the prospective user might get more information about the online social community through WOM which helps to evaluate "if it is worth" using the platform. In the

[23] To get an overview of research dealing with the antecedents of WOM provision please refer to the meta-study of de Matos and Rossi (2008).

last stage the new user decides to adopt the service. Thus, WOM recommendations may help the receiver to make an adoption decision. Adoption is only the start of a customer relationship, in which the user will form attitudes and behavior towards the service or product. Some research has already observed such post-adoption effects. Studies by Villanueva, Yoo, and Hanssens (2008) and Trusov, Bucklin, and Pauwels (2009) show that compared to traditional marketing channels WOM is more effective over time, and customers acquired through WOM add more long-term value to the firm. On an individual level, Wangenheim and Bayón (2004a) observed that referral switchers are more satisfied and loyal after switching to the service provider. Therefore, the post-adoption phase is of specific interest for online communities, as continuous user participation is needed to keep the service attractive for the users.

The Relationship Perspective – the role of social factors in the WOM process. Besides the perspectives on WOM senders and receivers, the relationship between them adds further insights on how WOM communication works. WOM builds upon social relationships, so that senders use their social network to give recommendations to other people (Bansal and Voyer 2000; Reingen and Kernan 1986). This is of specific importance for online social communities because the existing offline social structures can directly be transformed into online social relationships. Researchers have identified relational factors impacting the influence of WOM communication on receivers. Relationships with higher levels of demographic similarity between the sender and receiver of WOM are activated more likely for the flow of the referral (Brown and Reingen 1987). Similarity on general preferences and values affects the strength of WOM influence (Gilly et al. 1998; Wangenheim and Bayón 2004b; Wangenheim and Bayón 2007). Further, with regards to the strength of the tie between the WOM actors, strong ties are more likely to be activated than weak ties for referral, and they show a higher impact on the perceived influence of the referred information (Bansal and Voyer 2000; Brown and Reingen 1987).[24]

3.1.2.2 Offline and Online Word-of-Mouth

More recently, WOM occurs not only offline (traditional WOM), but also takes place in the virtual environment (online WOM or eWOM). Chen and Xie (2008) even identified online reviews as a new element of marketing communications. Various channels of online

[24] Tie strength is indicated by the importance of a social relation, frequency of contact, and type of social relation (e.g., close friend, acquaintance) (Granovetter 1973). It is described in more detail in chapter 3.2.1.1.

WOM have been studied, including posted comments and recommendations on online platforms (e.g., Godes and Mayzlin 2004; Hennig-Thurau et al. 2004; Liu 2006), online conversations in weblogs (Kozinets et al. 2010), pass-along E-Mails (De Bruyn and Lilien 2008; Huang, Lin, and Lin 2009), or directly recommending products or services found on the web by using "tell-a-friend" functionalities (Trusov, Bucklin, and Pauwels 2009). Although offline and online WOM share similarities in their purpose, the main difference is the separation of the WOM sender and the receiver by space and/or time. In the online domain senders and receivers might hardly know each other, for example when strangers write product reviews, which nevertheless can lead to high valuation of provided information (Weiss et al. 2008). Further, WOM via product rating platforms or weblogs can reach a much larger audience on the web compared to offline conversations, so that more people can potentially profit from recommendations or critics than in one-to-one conversations (Dellarocas 2003). Online technologies also make it much easier to reach a high number of people simultaneously than in the offline world. The online channel provides much faster and instant sending mechanisms, where the WOM sender can immediately and easily give recommendation about a product or service, rather than waiting until he meets the receiver in the offline world. Table 4 summarizes different channels and communication modes, in which WOM communication can take place.

	One-to-One Communication	**One-to-Many Communication**
Offline Word-of-mouth	■ Face-to-face conversation ■ Telephone conversation	■ Customer recommendations in mass media ■ Social Circles
Online Word-of-mouth	■ E-Mail ■ Chat ■ Tell-a-friend functionality	■ Online recommendation and rating platforms ■ Weblogs ■ Online Forum ■ E-Mail lists

Table 4: Examples of Different Word-of-Mouth Types and Channels

Despite the new channels for WOM referrals on the Internet, the predominant volume of recommendations still takes place offline (Keller and Berry 2006). Nevertheless, consumer platforms and online communication have become important tools for consumers to share recommendations, as underscored by the growing number of users. In addition, most online community operators rely on tell-a-friend functionality to facilitate new customer acquisition through online recommendations of their users.

3.2 Social Networks

This section provides an introduction to social networks and social network analysis. First, the theoretical foundation for understanding the main components and visualizations of social networks are presented. Thereafter, specific social network elements and the measurement of social network indicators are discussed. Overall, this section builds the basis for the application of social network concepts in the empirical studies of this dissertation.

3.2.1 Notations and Basic Concepts of Social Network Components

The concept of social networks has a long history. In fact, research on social networks dates back to the early twentieth century, with Moreno (1934) as the most prominent researcher on social structure at that time.[25] Moreno is regarded as the 'inventor' of the sociogram and founder of the field of sociometry, which is dedicated to the measurement of interpersonal relationships in small groups, and is considered the precursor to social network analysis (Wasserman and Faust 1994). Moreno's major contribution was the sociogram as a way to represent the formal properties of social configurations, though its innovative character is hardly appreciated today, as its use is nowadays taken for granted (Scott 2000). "A sociogram is a picture in which people (or more generally, any social units) are represented as points in two-dimensional space, and relationships among pairs of people are represented by lines linking the corresponding points" (Wasserman and Faust 1994, p. 12). Figure 6 illustrates an example of a sociogram (also called a graph).

[25] For a more detailed overview of the historical development of social network theory and social network analysis refer for example to Freeman (2004), Scott (2000), or Wasserman and Faust (1994).

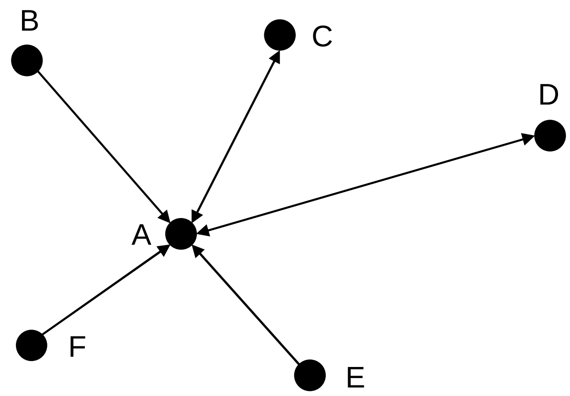


Figure 6: Example of a Sociogram

The sociogram thereby reflects the benefits of a social network[26]. In contrast to determining and analyzing individuals as isolated cases, the use of social networks takes the relationships between individuals into consideration. While, for example, classical and neoclassical economics operates with an atomized and undersocialized conception of human action (Granovetter 1985), social network concepts provide a theoretical alternative to the prevailing perspective of independent social actors (Wassermann and Faust 1994). Social network theory assumes that human behavior is affected by the social network of an individual, where social structures are in the center of human behavior (Borgatti et al. 2009). Thereby, the presence of regular patterns of relationship among actors constitutes structure (Wasserman and Faust 1994). The incorporation of the positions of single actors within their social networks contributes to explain behavioral outcomes. For example, individuals with exactly the same characteristics and knowledge differ in their outcomes according to different positions in the network structure. An individual with many relationships to other individuals can provide and receive more support than an individual without any contacts. Thereby, social network analysis is developed to understand the relationships between individuals as well as the implications of these relationships.

There are several key elements, which provide the basis for the discussion of social networks. Actors (also called points, nodes, or vertices) represent discrete individual, corporate, or collective social units (Wasserman and Faust 1994). Actors are the smallest units in a network. In the case of this dissertation, the actors are the users in an online social community,

[26] Wasserman and Faust (1994) note, that many people attribute the first use of the term "social network" to Barnes (1954).

but actors could also represent other groups of people like employees, organizations, or nations.[27] The relationship between the actors, that is their linkage to one another, is also called a tie (or line). Ties can represent many different social relations. Examples of ties between actors are transfer of material resources, interaction, communication, biological relationship, or friendships. Ties can be either directed (called arcs), for example when actor B sends a message to actor A, or bidirectional/ undirected (called edges), when there is a mutual relationship like a friendship between A and C (as illustrated in Figure 6). The structure of the network is then represented through a set of actors and a set of ties, which connect the pairs of actors (Freeman 1978/79; de Nooy, Mrvar, and Batagalj 2005).

There are two concepts, which are related to the characteristics of the actors and their relationships to each others: tie strength and homophily. Both are described in the following.

3.2.1.1 Tie Strength

As described above, the concept of social networks is based on the actors within a network and the social structure of relational ties between them (Burt 1980; Wasserman and Faust 1994). Ties can be described by their strength, which represents a "[…] combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" (Granovetter 1973, p. 1361). Ties may range from strong primary, such as spouse or close friends, to weak secondary, such as seldom-contacted acquaintances (Reingen and Kernan 1986).

Strong ties are consequently characterized by "(a) a sense that the relationship is intimate and special, with a voluntary investment in the tie and a desire for companionship with the partner; (b) an interest in frequent interactions in multiple contexts; and (c) a sense of mutuality of the relationship, with the partner's needs known and supported" (Walker, Wasserman, and Wellman 1993, p. 76). Strong ties can be very useful in cases of change and uncertainty, as they provide a base of trust, and are perceived as more credible than weak tie sources (Krackhardt 1992). They are more easily accessible through close contacts, and thereby have the advantage of allowing faster information flows.

On the other hand, in weak tie relationships actors are more loosely bound to one another. Weak ties are generally established between distant friends and acquaintances, who do not have an intimate and frequent relationship. Granovetter (1973; 1983) emphasizes the

[27] The use of the term "actor" does not imply that these entities necessarily have volition or the ability to "act" (Wasserman and Faust 1994).

"strength of weak ties". He argues that if an actor A has a strong tie to actor B and also to actor C, then B and C will be more likely brought together by A, compared to the situation in that A only has weak ties to both B and C (under the premise that B and C did not have a relationship before). This means, that an actor's close friends are more likely to socially interact with each other than are the actor's weak tie acquaintances. Therefore, groups of close friends build a dense social structure, where acquaintances of a focal actor mainly do not know each other, thus being not densely connected. However, weak tie contacts are likely to have close friends in their respective social circles and are thus able to connect two densely knit friendship circles. Consequently, Granovetter (1973) claims that weak ties can play an important role in linking members from one densely knit clump of close friends to members of different other densely knit groups. Having a connection to people outside of one's own clique can facilitate access to non-redundant flows of information. In this sense, weak ties can build a bridge between different social circles (Burt 1992). These bridging weak ties are of particular value to individuals. Although not all weak ties build bridges to distant groups of actors, weak ties are by far more likely to be bridges than strong ties (Burt 1992; Granovetter 1983).

In summary, weaker and stronger ties both provide significant advantages for the actor, depending on the specific situation and context in which the relationship is needed. "Weak ties provide people with access to information and resources beyond those available in their own social circle; but strong ties have greater motivation to be of assistance and are typically more easily available" Granovetter (1983, p. 209). Related to the discussion of strong and weak ties are the concepts of network closure and brokerage (Burt 2000), which are discussed in more detail in chapter 3.3.2 as part of social capital theory.

3.2.1.2 Homophily

Although related to the strength of a tie, homophily is a conceptually distinct concept that refers to the degree to which individuals are similar in terms of certain attributes, such as demographics, social status, or lifestyle (Brown and Reingen 1987).[28] Lazarsfeld and Merton (1954) differentiate between status homophily and value homophily: status homophily bases similarity on informal, formal, or ascribed status, including socio-demographics (age, gender, race, religion, occupation, education, etc.), while value homophily is based on similarities with regards to values, attitudes, and beliefs. Homophily differs from tie strength in that it

[28] For an overview of past studies on homophily and its dimensions refer to McPherson, Smith-Lovin, and Cook (2001).

reflects similar attributes that individuals possess who are in a relation, whereas tie strength describes the characteristics of the social relationship between the individuals itself (e.g., types of relationships such as close friends or acquaintances) (Brown and Reingen 1987).

Homophily implies that a contact between similar (more homophilous) people occurs at a higher rate than among dissimilar people, and that differences in social characteristics relate to network distance, which describes the number of relationships through which information flows to connect two actors (McPherson, Smith-Lovin, and Cook 2001). There is a tendency that socially similar people connect with each other more closely than dissimilar people from separate social worlds (Burt 1992; Granovetter 1973). Also known as the like-me principle, this suggests that people tend to interact with others who are like themselves (Laumann 1966). Further, when individuals are similar, they are more likely to develop greater levels of interpersonal attraction, trust, and understanding, than would be expected among dissimilar individuals (Ruef, Aldrich, and Carter 2003). On the other side, ties between non-similar individuals also dissolve at a higher rate (McPherson, Smith-Lovin, and Cook 2001).

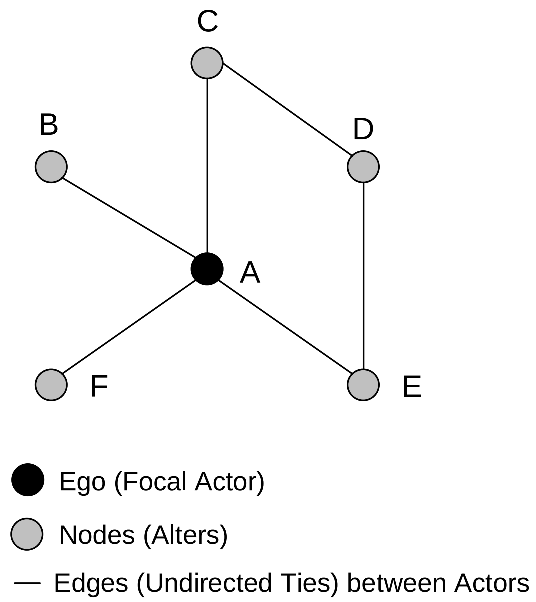
Overall, homophily leads to the establishment of certain relationships and the dissolution of others. Thereby, McPherson, Smith-Lovin, and Cook (2001) argue, that homophily also limits people's social worlds in their access and reception of information, their formation of attitudes, and their interactions. They observe that homophily in race and ethnicity most strongly structures individuals' social environments and relationships, followed by age, religion, education, occupation, and gender (McPherson, Smith-Lovin, and Cook 2001).[29]

3.2.2 Relevant Analytical Tools and Network Measures

Social network analysis is a technique that builds upon the importance of social networks for human behavior, and includes different tools and measures to describe social networks analytically. In order to evaluate the structure of a network and compare different ego- or sub-networks, different network indicators can be calculated. Network indicators have the advantage of providing a formal measure, which can be related to other relevant variables of individuals' or network characteristics. It can roughly be differentiated between two different levels of observation. First, the network as a whole can be analyzed, including all its nodes and their interconnections. Second, the researcher can take smaller networks as the unit of

[29] McPherson, Smith-Lovin, and Cook (2001) find that the literature on homophily is remarkably consistent across different relationships and dimensions of homophily, in that social networks are characterized by homophily and homogeneity in personal networks.

analysis, for example focusing on one actor and the directly connected contacts of this actor. The later is also called an ego-centered network. It consists of the focal actor, called ego, a set of other actors who have direct linkages to the actor (the neighbors of the ego), and the ties among the neighbors (Wasserman and Faust 1994). Such data on ego-centered networks are often referred to as personal network data. For example, the ego-centered network of actor A in Figure 7 consists of actors B, C, E and F.



Node	$C_d(n_i)$	$C_d'(n_i)$	$C_b(n_i)$	$C_b'(n_i)$	$C_c(n_i)$	$C_c'(n_i)$
A	4	0.80	7.50	0.75	0.17	0.83
B	1	0.20	0.00	0.00	0.10	0.50
C	2	0.40	1.50	0.15	0.13	0.63
D	2	0.40	0.50	0.05	0.10	0.50
E	2	0.40	1.50	0.15	0.13	0.63
F	1	0.20	0.00	0.00	0.10	0.50

Figure 7: Example of Network Measures in a Small Social Network

There is a multitude of different network measures in the literature on social network analysis (for an overview see for example Wasserman and Faust 1994; Freeman 1978/79). Here, the most relevant indicators are introduced. The focus is on indicators on the ego-level, because these are of specific interest for the empirical analyses presented in chapters 6 and 7.[30] The relevant indicators include centrality measures as well as ego-network density. Figure 7 provides an example of a social network with six actors, their relationships to each other, and the results of calculating the presented network measures.

[30] Because this dissertation concentrates on the investigation of network effects on an individual user level, the main interest is to obtain ego-based network measures, which can be analyzed in conjunction with other ego-level variables.

3.2.2.1 Centrality

Centrality is a key measure in social network analysis. Actor-based measures are especially useful for the analysis on an individual level. Basically, the centrality of an individual actor is the extent to which an individual is linked to others in the group (Ahuja, Galletta, and Carley 2003). Centrality represents the ego's position in the network, and indicates the access to resources and information from other actors. Thereby, central actors can control and influence the information flow between other actors. Further, centrality can also be viewed as a source of informal power because central actors can have broader access to various resources (Brass and Burkhardt 1993). The origins of the idea of centrality lie in the sociometric concept of the star-network, with the person in the middle being the most 'popular' in the group, who stands at the center of attention (Scott 2000). Wasserman and Faust (1994) describe central actors as those who are strongly involved in relationships with other actors, which results in increased visibility to the others. Several approaches and methods to measure centrality exist in network research (e.g., Freeman 1978/79). The most popular and relevant measures are degree, betweenness, and closeness centrality.[31]

The concept of **degree centrality** accounts for the number of ties that are adjacent, and thus directly related to an actor (Freeman 1978/79).[32] The degree centrality of a node is the number of lines incident with it and is expressed by (Wasserman and Faust 1994, p. 178):

$$(3.1) \quad C_d(n_i) = d(n_i) = \sum_j x_{ij}$$

with *d* as the degree of node n_i, which is the sum of all ties (or lines) *x* between the focal actor *i* and all adjacent actors *j*. Figure 7 indicates the calculated values for the degree centrality of each node: point A has four direct neighbors, resulting in a degree of 4, point B has a degree centrality of 1, and so on.

A standardization of degree centrality helps to compare this measure between networks of different sizes. $d(n_i)$ denotes the degree of the focal actor and *g* the number of nodes within the network.

[31] According to Wasserman and Faust (1994), centrality seems most appropriate for nondirectional relations. Because friendship relationships between online social community users are mutual, and therefore nondirectional, the indicators are described for the nondirectional case. Nevertheless, there also exist indicators for directional relations (see for example Wasserman and Faust (1994).

[32] de Nooy, Mrvar, and Batagelj (2005) use the term degree centrality also for the degree of an actor.

(3.2) $C'_d(n_i) = \frac{d(n_i)}{g-1}$

An actor who is in direct contact with many other actors of the network is potentially a major channel of information. This actor is likely to develop a sense of being "in the mainstream of information flow in the network" (Freeman 1978/79, p. 220).

Betweenness centrality takes the position of the actor in the entire network into account, and not only the ties to his direct neighbors. As interactions between two non-adjacent actors might depend on other actors who lie on the paths between the two, this form of centrality can be included by the idea of betweenness. It measures the interrelationships or communication flows within a network and calculates how often an actor is located on the shortest path (called geodesic) between all other pairs of actors in the network (Wasserman and Faust 1994; de Nooy, Mrvar, and Batagelj 2005).[33] When an actor lies either on the only geodesic or on all geodesics linking a given pair of other actors, he stands between those actors. An actor has a high betweenness centrality if he is often located on the geodesics. Thereby, betweenness centrality measures the amount of flow in the network, that is in the "control" of a certain actor, as he is able to cut that flow of information (Borgatti 2005). Having control over the communication between other actors is an advantage, but more notably actors with high betweenness are important intermediaries in the communication network, that means they are involved in the interaction of the other actors (de Nooy, Mrvar, and Batagelj 2005; Freeman 1978/79). A critical assumption for the basic betweenness centrality is that all lines have equal weight, and that communications will travel along the shortest route, regardless of the actors along the route (Wasserman and Faust 1994). Then betweenness centrality is denoted as the sum of the probabilities of actor n_i lying on the shortest paths between all other pairs of actors (Freeman 1978/79, p. 226-7):

(3.3) $C_b(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk}$

with g_{jk} being the number of shortest distances for every pair of actors *jk*. $g_{jk}(n_i)$ is the number of locations on shortest paths of the focal actor n_i between *j* and *k*. Consequently, the betweenness centrality measures the probability that a path from *j* to *k* takes a particular

[33] A path is a sequence of one or more edges.

route. If there is more than one geodesic between *j* and *k*, all geodesics are equally likely to be used. The ratio of these two values is aggregated for each pair of actors in the network.

Again, standardization of the betweenness centrality makes the measure comparable and independent of the network size:

(3.4) $$C'_b(n_i) = C_B(n_i) / [(g-1)(g-2)/2]$$

The betweenness centrality of the nodes in Figure 7 is calculated based on these formulae. For example, the betweenness centrality of point C is 1.5, i.e. C lies in each case on one of the two shortest paths between A and D, B and D, and D and F. Consequently, the standardized betweenness centrality equals .15 for point C.

Closeness centrality again takes the whole network into consideration, not only the ego-network, and therefore depends on direct and indirect ties. It measures the distance of the focal actor to all other actors within the network. The idea of this measure is that an actor is central, if he can quickly interact with all others (Wassserman and Faust 1994). According to Freeman (1978/79), the simplest measure is that of Sabidussi (1966), in which actor closeness is a function of geodesic distances. Closeness centrality is calculated by the sum of the shortest distances of a node to all other nodes.

(3.5) $$C_c(n_i) = \left[\sum_{j=1}^{g} d(n_i, n_j) \right]^{-1}$$

$d(n_i, n_j)$ is the number of lines in the geodesic linking actors *i* and *j* (i.e. $d(\cdot,\cdot)$ is a distances function). The total distance of *i* to all other actors is $\sum_{j=1}^{g} d(n_i, n_j)$. The closeness centrality measure takes the inverse value of the sum of distances. This leads to the interpretation, that a small value of distances (or a larger value of the inverse value) indicates that the node is “close” to a large number of actors in the network, and therefore more central (e.g., Scott 2000). At the maximum, the index equals $(g-1)^{-1}$, which arises when the actor is adjacent to all other actors (Wasserman and Faust 1994). This measure of closeness centrality is only meaningful in a connected graph, because in an unconnected graph every point is at an infinite distance from at least one other point (Freeman 1978/79).[34]

Standardization makes the closeness measure comparable across networks (Wasserman and Faust 1994, p. 185):

$$(3.6)\quad C_c'(n_i) = \frac{g-1}{\sum_{j=1}^{g} d(n_i, n_j)} = (g-1)C_c(n_i)$$

With respect to the example in Figure 7, the closeness centrality of A is 0.17 and the standardized closeness centrality is 0.83.

3.2.2.2 Ego-Network Density

The density of the ego's network (expressed by the clustering coefficient) describes the density among its neighbors, i.e. the number of relationships between direct neighbors of the focal actor in relation to the maximum number of relations between these neighbors (Batagalj and Mrvar 2010). D_E is the ego-network density of focal actor *n* with *l* being the number of lines between adjacent neighbors and *d* being the degree (centrality) of the ego *i*:

$$(3.7)\quad D_E(n_i) = \frac{2l_i}{d_i(d_i-1)}$$

Density takes on values between 0 (empty graph) and 1 (complete graph). As it measures the fraction of the contacts of a given actor who are connected with each other, the clustering coefficient describes the cliquishness of a network (Mayer 2009). A network being cliquish means that if *i* is connected to *j* and *j* is connected to *k*, there is a relatively high probability that *i* is connected to *k* as well (e.g., Granovetter 1973). The density of the ego-network of actor A in Figure 7 has a value of 0, because the direct neighbors of A (B, C, E, and F) are not connected to each other. If there would be an additional connection between A and D, then the ego-density of A would be 0.20, because there are 2 lines between the neighbors of A out of a maximum of 10 lines.

[34] "A graph is connected if there is a paths between every pair of nodes in that graph." (Wasserman and Faust 1994, p. 109). Because the social network of the research object in the empirical study of this dissertation is not a connected graph, closeness centrality is not included in the empirical study. However, it is described here to give a complete overview of the most relevant centrality measures.

The applications of social network analysis and the presented indicators are broad. They are used, for instance, in sociology, biology, physics, management, psychology, or economics (for examples see Borgatti et al. 2009; Wasserman and Faust 1994). Also computer networks can constitute social networks, when they connect people with each other (Wellman et al. 1996). Recently, different studies investigated the effects of user interactions in the online domain on user behavior (e.g., Dholakia, Bagozzi, and Pearo 2004; Trusov, Bodapati, and Bucklin 2010). Thereby, online communities gained attention of network researchers (e.g., Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Toral et al. 2009). Also in marketing, it has been postulated to take not only individual perspectives into account, but rather take a more holistic view on social and networked environments (e.g., Achrol and Kotler 1999; Algesheimer and Wangenheim 2006). Therefore, social network analysis is increasingly important to investigate social relationships and human behavior in online communities. For example, measures on the individual's position in the network help to explain user participation in online communities (e.g., Wasko and Faraj 2005).

3.3 Social Capital Theory

After the introduction to social networks, its basic concepts and analytical measures in the previous chapter, a useful theoretical framework for understanding the value of such networks for the actors is provided by the concept of social capital. Social network research emphasizes the importance of an actor's position in and configuration of his network because it determines in part the actor's opportunities and constraints, and therefore plays a significant role for the actor's outcomes (Borgatti et al. 2009). There is a link between the social network of an actor and the performance of the actor, in that the performance does not only depend on the human capital an actor has developed but also on the social capital that lies in his connection to other actors (Burt 1992). Nahapiet and Ghoshal (1998, p. 243) define social capital 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." Thereby, the social capital emphasizes the value of social structures, where relationships are an important resource for social action (e.g., Bourdieu 1986; Burt 1992; Coleman 1988).

Like other forms of capital, social capital helps to achieve certain outcomes that would not be possible without it. The key difference between social capital and financial, physical or human capital is that it is embedded in the social realm (Wasko and Faraj 2005). While physical capital is embodied in observable material form, and human capital is embodied in the skills and knowledge acquired by an individual, social capital exists in the relations among persons (Coleman 1988). Thus, social capital is least tangible of the different forms of capital. Burt (1992) highlights two major differences of social capital compared to financial and human capital: First, social capital is not owned by a single entity (be it an individual or organization), but is jointly owned by the parties of a relationship, and is lost in the absence of that relationship. Second, while financial and human capital get invested into production capabilities, i.e. acquiring raw materials and crafting the raw materials into products, social capital determines the rate of return on the invested capital through relations with colleagues, friends, and clients. Therefore, social capital is inherent in the social structure and offers advantages through the access of resources from other actors. Different outcomes are explained by human capital through the ability of people (intelligence, skills, knowledge), and by social capital though the relationships with people (Burt 2000). For example, when two individuals own similar human capital, that is similar knowledge and skills in a certain field, they would perform differently when one has many trustful relationships with different circles of colleagues and clients, and the other actor only has few superficial relationships. Sales persons might have equal skills of selling products, but it is the sales person with more and diverse contacts to potential clients and trustful relationships to these clients who is more successful (all other things being equal). “The social capital metaphor is that the people who do better are somehow better connected” (Burt 2000, p. 347).

3.3.1 Three Dimensions of Social Capital

In order to better understand what constitutes social capital, Nahapiet and Ghoshal (1998) propose three dimensions: structural, relational and cognitive capital. The ***structural dimension*** refers to the overall pattern of connections between actors and includes social interaction and the location of actors and their contacts (Tsai and Ghoshal 1998). Important elements of structural capital are network ties between actors, network configuration, and ‘appropriable’ organization (Nahapiet and Ghoshal 1998). Network ties provide three forms of information benefits: the access to resources and information, the potential to get information sooner than others, and the opportunities that arise through referrals in the future (Burt 1992). Network configuration describes the overall network structure, how the ties, through which the information flows, are arranged and how the network ‘looks like’. And through an appropriable organization, social capital can (in certain contexts) be transferred from one social setting to another (Coleman 1988; Nahapiet and Ghoshal 1998). On an individual level, structural capital constitutes the properties of the focal user's position in the network, which influences the user's behavior and can be operationalized by measures such as density, connectivity, and centrality (Nahapiet and Ghoshal 1998; Tsai and Ghoshal 1998; Wasko and Faraj 2005). Thereby, the structural dimension directly relates to the social network measures discussed in chapter 3.2.

Relational capital focuses on the particular relations people have and which influence their behavior. Its main elements are obligations and expectation, identification, trust, and norms (Nahapiet and Ghoshal 1998). The distinction of structural and relational capital is that the former focuses on structure per se, while the later concentrates on the quality of the relation. For example, in online social communities two users may take equivalent positions in similar ego-network configurations (i.e. they both have the same number of friends who are not connected too each other), but if they differ in their personal and emotional attachment to other users (i.e. one user shares high identification and trust with that group and the other user does not), their actions are also likely to differ. As already shown in the literature overview in chapter 2.5, there is evidence that factors like trust, identification, reciprocity, and norms can facilitate the information exchange through participation in online communities (e.g., Cheung and Lee 2009; Chiu, Hsu, and Wang 2006; Dholakia, Bagozzi, and Pearo 2004; Ridings, Gefen, and Arinze 2002). Therefore, relational capital does not only provide access to the information, but also enhance the willingness for interaction in relationships.

The third dimension is the ***cognitive dimension***. It refers to those resources providing shared representations, interpretations, and systems of meaning within groups and includes facets like shared language and codes and shared narratives (Nahapiet and Ghoshal 1998). The sharing of a social context, in the form of language and narratives, supports the process of meaningful communication between the actors. A common language helps to get access to information and to make sense of that information through common knowledge of its meaning. Shared narratives enable the creation and transfer of interpretations of events, in the form of stories shared between actors (Nahapiet and Ghoshal 1998). Cognitive capital develops through the interaction of actors over time by sharing skills, knowledge, and interpretations. Consequently, cognitive capital consists of individual expertise, which facilitates the understanding of a common language, and experience with applying the expertise (Wasko and Faraj 2005).

These three dimensions all contribute to social capital. They are not completely separated but might be highly interrelated in certain aspects. For example, network structure provides the opportunity for the actor to exchange resources with other actors. If the actor has a favorable position in the network, he has better direct and indirect access to people who can provide resources and to the resources they can reach through their relationships. The quality in the relation to these actors lies in the norms, trust and identification, which facilitates the motivation to exchange resources (Adler and Kwon 2002). Further, the actor needs the ability to interpret and use these resources. Therefore, these dimensions as sources of social capital are linked to each other.

3.3.2 Benefits of Social Capital

The discussion of social capital is related to the actors in the network, but those are not necessarily individual persons. Social capital can be conceptualized on different levels, including individuals (e.g., Wasko and Faraj 2005; Wiertz and de Ruyter 2008), work groups and organizations (e.g., Tsai and Ghoshal 1998), or societies (e.g., Putnam 1995). An important aspect of social capital on all levels is that investment in social relations leads to an expectation of benefits. It can be obtained and used by individuals for their personal benefit, and be a public good that is collectively owned and serves the community as a whole (Bourdieu 1986; Burt 1992; Burt 1997; Coleman 1988). It has been argued that social capital leads to market and non-market returns, including instrumental and social benefits, which are, for instance, efficient transfer and use of knowledge and social support (e.g., Adler and Kwon 2002; Matthwick, Wiertz, and de Ruyter 2008). More concretely, Adler and Kwon (2002) identified three forms of benefits: (1) information benefits are direct benefits that arise from the broad access to information of higher quality, timeliness and relevance; (2) influence, control and power benefits allow the actor to get things done and achieve goals; and (3) solidarity benefits are gained through trust and observed rules that reduce the need for formal controls.

Different forms of social capital contribute to achieve such benefits. Particularly relevant from a structural perspective are two concepts, which support the realization of different benefits: brokerage (Burt (2000) and network closure (Coleman 1988). Both concepts are related to the presence or absence of structural holes and the properties of the network configuration of individual actors and groups. A structural hole is a “relationship on nonredundancy between two contacts” (Burt 1992, p. 65), e.g. a connection between two densely connected clusters. Structural holes are thus an opportunity to broker the flow of information between people, and control the resource flow that bring together people from opposite sides of the hole (Burt 2000). The argument is based on network concepts such as the strength of weak ties (Granovetter 1973; cf. chapter 3.2.1.1) and betweenness centrality (Freeman 1978/79; cf. chapter 3.2.2.1). In a sparse network, actors can build bridges between separate social circles of friends. The focus of this ‘bridging’ social capital is on the relations of the focal actor to actors of other groups. In line with the strength of weak ties argument, it is related to the access of non-redundant information, ideas and knowledge. Weak ties are most likely to take on the function of bridging two groups of people. However, as long as a tie spans the structural hole it can provide access to information and control benefits, independent of the tie’s strength (Burt 1992). Overall, information and power benefits are emphasized through brokerage.

While brokers connect two otherwise weakly connected groups of people, closure refers to strongly connected, i.e. dense networks (Burt 2000; for density measurement in ego-networks see chapter 3.2.2.2). Closure is maximum when all actors in a network are connected to everyone else. It facilitates the access to information and sanctions that make it less risky for people in the network to trust one another (Burt 2000; Coleman 1988). Therefore, closure enables the development of norms, identities and trust (Coleman 1988). These solidarity benefits help to increase cohesiveness within an organization or community (Adler and Kwon 2002) and can enhance social capital in that it provides the basis for more reliable communication between the actors. On the other hand, network closure also has negative effects regarding a reduced tendency to connect and interact with people outside the sub-network. Based on Granovetter’s (1973) work on the strength of weak ties a dense network may result into redundant information within the group of contacts. In a closed network, where all contacts are connected, little new information would be exchanged within the group and the members would not be exposed to new knowledge, ideas and interaction (Burt 1997).

In summary, both concepts provide advantages to the actors. A brokerage role, i.e. spanning structural holes, gives access to non-redundant information and promotes control and power. Network closure does not provide relationships to actors with potentially new and non-redundant information, but facilitates the development of trustful relationships and group attachment. Both perspectives contribute to social capital, and depending on the situation and the current needs of the actors within a group, one or the other network configuration can be more appropriate. Adler and Kwon (2002) suggest that management of organizations and companies should pay attention to both aspects.[35]

Overall, social capital is a useful concept to better understand human action through the consideration of social relationships. Besides its use in social studies of families, education, community life, or democracy, it also gained relevance in organization studies on individual and organizational performance. Issues researched include the influence of social capital on career success, job search, resource exchange between organizational units, supplier relations, and employee turnover (Adler and Kwon (2002) give a short overview on related studies). The newest applications of social capital include the field of online communities. Recent research has shown that the social capital concept does not only provide valuable insights for offline, but also in online structures. Online community studies explored certain aspects of social capital as facilitators of the creation and exchange of information and knowledge (e.g., Chiu, Hsu, and Wang 2006; Kankanhalli, Tan, and Wei 2005; Wasko and Faraj 2005; Wiertz and de Ruyter 2007). Thereby it is demonstrated that different dimensions can influence the actions of users in these online communities (see also chapter 2.5 which includes some dimensions related to social capital in the literature review). However, especially the importance of the structural dimension is underrepresented by past research. Therefore, this thesis will provide a deeper understanding of the influence of social capital on user behavior in this respect.

3.4 Further Social Concepts to Explain Social Relations and Behavior

Other relevant social theories can provide a further understanding of the relationships between the social context and the behavior of individuals within this social context. This section gives a brief overview of additional theoretical concepts, which help to explain social processes and which constitute a basis for the theoretical propositions and hypotheses presented in the empirical study of this dissertation. Among those concepts are social identity, social exchange, social presence, and collective behavior.

---

[35] Adler and Kwon (2002) relate closure and brokerage to the focus on internal and external ties, i.e. ties within a group and ties to other actors outside the group. Thereby, closure is more related to internal ties, while brokerage is related to external ties. However, for all forms of social capital both linkages within and outside the group are important, as both provide benefits relevant to specific situations, even when the focus is on external ties.

3.4.1 Social Identity

Social identity theory is concerned with intergroup relations, group processes, and the social self (e.g., Hogg, Terry, and White 1995; Tajfel 1974). Tajfel (1974, p. 69) defines social identity as “that part of an individual’s self-concept which derives from his knowledge of his membership of a social group (or groups) together with the emotional significance attached to that membership.” Social identity is related to the social groups an individual belongs to, and is thus in contrast to personal identity, which consists of idiosyncratic characteristics and attributes of the individual such as competence, ability and interests (e.g., Bhattacharya, Rao, and Glynn 1995; Lantz and Loeb 1998). Social identification emphasizes the group as “we”, where personal identity construes the individual as “I” or “me” in dyadic relationships with specific other people (Hogg 2006).

Different components support the establishment of social identity, which include a cognitive awareness of one’s membership in a social group (self-categorization), a sense of emotional involvement with the group (affective commitment), and a positive and negative value connotation attached to this group membership (group-based self-esteem) (Ellemers, Kortekaas, and Ouwerkerk 1999). First, social categorization is closely related to the concept of social identity (Hogg 2006; Hogg, Terry, and White 1995; Tajfel and Turner 1979). Individuals fall in social categories (e.g., nationality, sports team, community, club), to which they feel to belong, and that provide a definition of who one is in terms of defining characteristics of the category as a part of the self-concept (Hogg, Terry, and White 1995). Self-categorization defines intergroup boundaries of social identity and distinguishes the own group of people, with which one identifies, from other groups, with which one does not identify. Thereby, individuals want to achieve and maintain a positive social identity, where other groups (outgroups) are evaluated as distinct and less favorable compared to the own group (ingroup), which results in higher self-esteem (Brown 2000; Hogg 2006). Second, the emotional component can be described as the members’ affective commitment, which presents the identification with and the emotional attachment to the group and fosters loyalty and willingness to maintain relationships in group settings (e.g., Bagozzi and Dholakia 2002; Meyer and Allen 1991). Last, the group-based self-esteem is established through evaluations of self-worth based on the belonging to a group or community (Dholakia, Bagozzi, and Pearo 2004).

Overall, the goal of the individual is to reach a positive social identity, which, in turn, leads to a greater attachment to the group and facilitates that the individual will stay with the group and be involved, rather than leave the group (e.g., Brown 2000; Tajfel 1974). Identification is also considered as an important facet that facilitates the motivation of exchanging resources and information (Nahapiet and Ghoshal 1998), and is thus a basis for the functioning of a group. In online social communities, social identity plays an important role because users with high social identity in the community distinguish themselves more from other communities. Within online communities, also sub-groups can identify with each other and distinguish from other people. The user’s identification with the online social community is based on the understanding that membership leads to significant benefits, which help to fulfill the members’ needs (Dholakia, Bagozzi, and Pearo 2004). Thereby, belonging to the group is important for their behavior within that community. Several researchers highlight the relevance of social identity in the online community context (e.g., Algesheimer, Dholakia, and Herrmann 2005; Dholakia, Bagozzi, and Pearo 2004; Woisetschlaeger, Hartleb, and Blut 2008). The identification with the online community affects the engagement and participation with the group in the online community. Therefore, user behavior can be explained based on social identification processes. Altogether, social identity can affect behavior of the individual in the group and towards other groups.

3.4.2 Social Exchange

Social exchange theory focuses on people’s dependence on one another with respect to their needs of social life and on the benefits that people obtain from and contribute to social interaction (Molm 2006). Although social exchange has been described from different perspectives, there is consensus that it generates obligations through interactions (Emerson 1976). Thereby, social exchange theory posits that social interaction is based on the actors’ expectations that benefits will arise from the social exchange for all parties involved. It focuses on relations of some length and endurance because the payback of benefits is unspecified (Blau 1964). In social exchanges, one actor might do another actor a favor with a general expectation of receiving a favor in return in the future without knowing the exact return (Kankanhalli, Tan, and Wei 2005). Consequently, “exchange theory can be seen as an approach to interaction and structure based on two principles: (i) The actor can be modeled as motivated by interests or rewards/punishments - i.e. all behavior can be seen as so motivated; (ii) most interaction consists of the exchange of valued (though not necessarily material) items“ (Cook and Whitmeyer 1992, p. 114). The decision to take an action of exchange is based on cost-benefit expectations. Exchange theories assume that actors are self-interested, seeking to maximize benefits and minimize costs (Molm 2006). Individuals engage in social interactions, when they expect reciprocal benefits such as returned favors, approval, status, and respect (Blau 1964; Molm 2006; Wasko and Faraj 2005). Reciprocity is thereby an obligation to repay the benefits received (Gouldner 1960). Based on the norm of reciprocity long-term exchange relationships can emerge between actors.

Social exchange can also occur in social networks and not necessarily in dyadic relationships (Emerson 1976). This form of the social exchange approach emphasizes social structures within which exchange takes place (Molm 2006). Thereby, exchange theory and network analysis are both concerned with the social structure, the configuration of relationships between actors and the actor's position (Cook and Whitmeyer 1992). Following this, social exchange and social networks can be linked in communities, for example, in knowledge sharing networks where people exchange knowledge and social support leading to benefits from the exchange (e.g., Kankanhalli, Tan, and Wei 2005; Wasko and Faraj 2005). Interaction and exchange of information can be interpreted as generalized social exchange between a group of people and with indirect reciprocal dependence (Kankanhalli, Tan, and Wei 2005). Benefits or rewards can then increase the frequency of future behavior on which they are contingent (Molm 2006). In online communities, this means that expected benefits through reciprocal interactions, including valued items such as information from other users, social support, recognition or status, lead to certain behavior in the group of people. Thus, social exchange and its link to social networks can help to explain interactions and relationships between the members of a group.

3.4.3 Social Presence

The concept of social presence describes the "degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships" (Short, Williams, and Christie 1976, p. 65). It is concerned with the ability of a medium to generate a perception of the communication partner's presence, i.e. that he is a 'real' person. Thereby, social presence depends on verbal as well as non-verbal context cues (Rice 1993). If there is a good match between the characteristics of the medium and the purpose of the communication or interaction between the involved people, the communication process will lead to a better outcome, for example in terms of satisfaction or effectiveness (Rice 1993).

Social presence varies by media, representing the subjective quality of the respective medium, which affects how individuals interact and communicate (Short, Williams, and Christie 1976). In addition, certain activities are likely to be affected by the medium's social presence which includes information and idea exchange, problem and conflict solving, getting to know people and maintaining relationships (Rice 1993; Short, Williams, and Christie 1976). Different communication media show different levels of social presence according to such factors as their sociability, warmth, personalness, and sensitivity (Fulk et al. 1987). Because of these differences in social presence, different activities fit better to different media. When a task is interpersonally involving, individuals will prefer a medium that provides high social presence (Fulk et al. 1987). For example, solving problems, giving important advice or providing social support are activities that would be pursued preferably through media with a richer context. Thereby, face-to-face is regarded as the channel with the highest social presence and highest appropriateness to pursue most of the relevant activities in organizations, where text-based communication is a channel of low social presence (Rice 1993). Thus, individuals would choose the medium that is most adequate to fulfill their situational needs of interaction with other people.

In the online context, reduced cues lead to more impersonal and nonconforming communication than in face-to-face settings (Parks and Floyd 1996; Sproull and Kiesler 1986). Consequently, it can be argued that computer-mediated environments like online communities show lower social presence in the traditional sense. The limitations of communication channels and the lack in social context cues in computer-mediated communication make close relationships more difficult to develop online (Chan and Cheng 2004). The lower social presence of computer-mediated communication environments may be sufficient to maintain strong ties when people already know each other. However, through different functionalities it is possible to increase social presence, for example through visual elements like videos and photos, emotions provided in text (emoticons), descriptions about the person on profiles, or immediate communication in chats. Therefore, information and social support can be provided in online social communities through the users' relationships. Online communities demonstrate that social relationships can develop online, although it may take longer to develop intimacy, identification, and strong relationships than in the offline world (e.g., Walther 1992; Wellman et al. 1996).

In past research, it has been shown that Website social presence influences perceived usefulness, perceived risk and trust, which affects the intention to transact in online shopping (Dash and Saji 2007). Shen and Khalifa (2008) conceptualized social presence in online communities and demonstrated that the perceived level of social presence affects the users' participation behavior. Altogether, this emphasizes the relevance of social presence for online interaction.

3.4.4 Threshold Models and Collective Behavior

Threshold models help to explain the success or failure of collective action and are concerned with the influence of collective behavior on individual behavior. According to Granovetter (1978), collective behavior occurs in situations where people have to decide between two options and their decision is dependent on the number of other individuals that have already chosen a certain option.[36] People have different threshold levels, i.e. individuals vary in their cost and benefit levels required to choose an alternative. A popular example is joining a riot. People only join the riot when enough other people are already involved and the benefits are perceived as higher as the costs. "A person's threshold for joining a riot is defined here as the proportion of the group he would have to see join before he would do so" (Granovetter 1978, p. 1422). Thereby, the concept of collective behavior adds to the knowledge about preferences, motives, and beliefs of participations for explaining outcomes.

Threshold models can be applied to various settings, for example, they are used in innovation diffusion (e.g., Valente 1996), but could also help to understand the decision to stop an activity such as leaving a boring lecture. When considering an online social community, the user has the choice of leaving or staying in the community, where the decision is based on the costs and benefits to do so and depends in part on how many others make which choice (Granovetter 1978). At some point, when the proportion of people is large enough who take one choice, the perceived benefits to the individual of doing the thing in question (staying or leaving) exceed his perceived costs, which leads to the decision for the other alternative. Consequently, the probability of any individual actor's behavior can be described as a function of the number of other actors in the network that took a certain decision. A key concept is that of "threshold", which describes "the number or proportion of others who must make one decision before a given actor does so; this is the point where net benefits begin to exceed net costs for that particular actor" (Granovetter 1978, p. 1420). This can also result in a chain reaction: if users defect from communication media, the benefits to the remaining users will decrease, and the costs increase, thus stimulating further defection (Markus 1987). Collective behavior suggests that individual behavior is dependent on the number of people who already took a decision of pursuing or not pursuing a specific action. Though in reality more complex, collective action helps to explain human behavior.

---

[36] Granovetter (1978) refers to threshold models that treat binary decisions, i.e. two distinct and mutually exclusive behavioral alternatives.

4 Three Empirical Studies on Network Structure, Attitudes and Behavior in Online Social Communities

The objective of this dissertation is to investigate the effects of structural and attitudinal factors on user behavior and user perceptions in online social communities. Three focus areas were introduced in chapter 1: user acquisition, user activation and user retention. Each of the three areas is addressed by one empirical study that enhances the understanding of the influences on user behavior in online social communities. The empirical analyses address different research gaps in the field of online community marketing and have significant implications for community operators on different stages in the membership development process.

This chapter provides an introduction to the methodology of the empirical studies that will be presented in the subsequent chapters. First, an overview of the research approach is given. The three studies are set in context to the membership development process and the main objectives of the studies are outlined. Second, the research site and the data collection approach are introduced. In this regard, the main variables for the data analysis are presented. Chapters 5, 6, and 7 include the details of the three empirical studies: each section presents the link of the study to the theory, the derived hypotheses, a description of the used data, the analyses to test the hypotheses, the results, and a discussion of the results from both a theoretical and a managerial perspective.

4.1 Objectives and Research Approach along the Membership Development Process

Online communities take on important roles in marketing today (cf. chapter 2.3). User participation and favorable attitudes toward the online community are the most important factors to keep the community alive and attractive for its users. Thus, the focus of the empirical studies is to understand what increases a user's participation and positive perceptions towards online communities. The results of the empirical studies are presented in chapters 5, 6, and 7. They answer a couple of purposes. On one hand, beyond past research on online communities, the results facilitate a better understanding of the drivers of user attitudes and behavior in online communities. Several research gaps have been identified, which will be addressed in these studies (see below for the outline of the studies). This provides additional insights for marketing researchers on how user behavior in online communities works. On the other hand, the practical relevance of understanding the influencing factors of user behavior will be identified. In different phases of the membership development process specific issues emerge for the community operators. The generated knowledge can help to enhance the attachment to the community and the user's participation on the platform.

Because the number of contributors (compared to lurkers) is generally rather low in online communities (Nonnecke and Preece 2000), it is very important to stimulate participation and increase the number of active contributions. Community operators must ensure that enough new and actual content is provided on the website. Frequent interaction and contribution is crucial because it builds a stronger relationship to the community and keeps the community attractive for posters and lurkers (e.g., Lin and Lee 2006; Pajuniemi 2009; Woisetschlaeger, Hartleb, and Blut 2008). Figure 8 shows the central research questions and the integration of the three studies in the generic membership development process[37]. If the user roles in the life cycle and the membership development process are brought together, different focus points can be defined for community operators to build a successful online community. The figure illustrates that there are different stages in the relationship of a user with the online community, at which community operators can intervene.

(1) The acquisition of new members is of high priority for community operators. Customer acquisition is an important step to ensure that a sufficient number of members actively use the community. Operators need to attract people who have not yet registered with the online community (called here 'non-members'[38]) and convince them to register for the service. Different marketing communication channels are available to approach prospective users. Community operators ask themselves which marketing channels are most effective in building a strong member base. Of particular relevance for attracting new members to online communities are interpersonal communications, like customer-to-customer and employee-to-customer communications, because they are effective in influencing customers (e.g., Herr, Kardes, and Kim 1991; Palmatier et al. 2006; Price and Feick 1984) and available for communities of all types and sizes. Therefore, a comparison of those communication channels is drawn with respect to their impact on users' post-adoption attitudes and behavior. More specifically, WOM and personal selling, as well as offline- and online-WOM will be com-

---

[37] Figure 8 describes a simplified membership development process with the most relevant user roles. Although additional user roles could be defined, e.g. leaders, this further differentiation does not provide additional value for the discussion of the objectives of the three empirical studies.

[38] Because visitors are not relevant for some (closed) online communities, like Facebook, it is referred to non-members and not to visitors when talking about people who have not yet registered.

pared[39]. This is the first study to investigate these differences in order to answer the question of how the type of acquisition channel influences the users' participation and perceptions of the community after registration. It is also analyzed what types of users are attracted by each channel and how this influences the link between channel and user attitudes and behavior. Thereby, the benefits of certain channels are elaborated and discussed. For community operators this provides valuable insights on how to use different marketing channels to acquire new users.

(2) After registration for the online community service, one important goal is to promote the users' activity on the platform (Hagel and Armstrong 2006). Many users become passive users (or lurkers) and just browse the platform without any active participation[40]. Because active participation is the key element to ensure a sustainable online community service, passive members need to be converted into active members, and active members should be cultivated to be more active. Therefore, it is crucial to know what the main drivers of active user participation are in order to increase participation behavior. Because of the knowledge gap on how an individual's network position in the online community affects user participation, it is a priority of this study to analyze its impact (cf. chapter 2.5.5 for the limited empirical evidence of the impact of the structural dimension). The additional value of attitudinal factors of the online community users is also investigated. The influence of structural and attitudinal factors is tested based on social capital theory and related social concepts (cf. chapters 3.2 to 3.4). Further, as distinct user groups differ, a comparison of two user groups with distinct social motives can help to better understand how to treat users differently.

(3) At the end of the user life cycle the user separates from the online community. In many cases the user does not unsubscribe from the service, but simply does not return. As customer retention is an important theme in marketing (e.g., Bhattacharya 1998; Bolton 1998; Nitzan and Libai 2011), online community operators need to understand why their users leave the community. Because of the relevance for community operators and the lack of knowledge on this issue in online community research, studying user defection is of high interest. As identified in the study on active user participation and in line with social theory, the impact of the network structure and the user's position in the network play a major role in understanding user behavior. Because network measures are easily available their value to explain user

[39] The definition and operationalization of the channels are provided in chapter 5.
[40] Corresponding to chapter 2.5, active participation is defined as all activities in which the user actively interacts with other users or the platform, and thereby contributes in the form of written text (e.g., messages, guestbook entries), providing content (e.g., sending virtual gifts), or submitting ratings (e.g., rating photos or groups). In contrast, passive participation is related to content consumption or the receipt of messages, content or ratings.

defection is assessed. Additional variables comprising the users' community engagement are also considered. The results provide the community operator with insights on what drives user defection, and thus helps to develop measures to retain users.

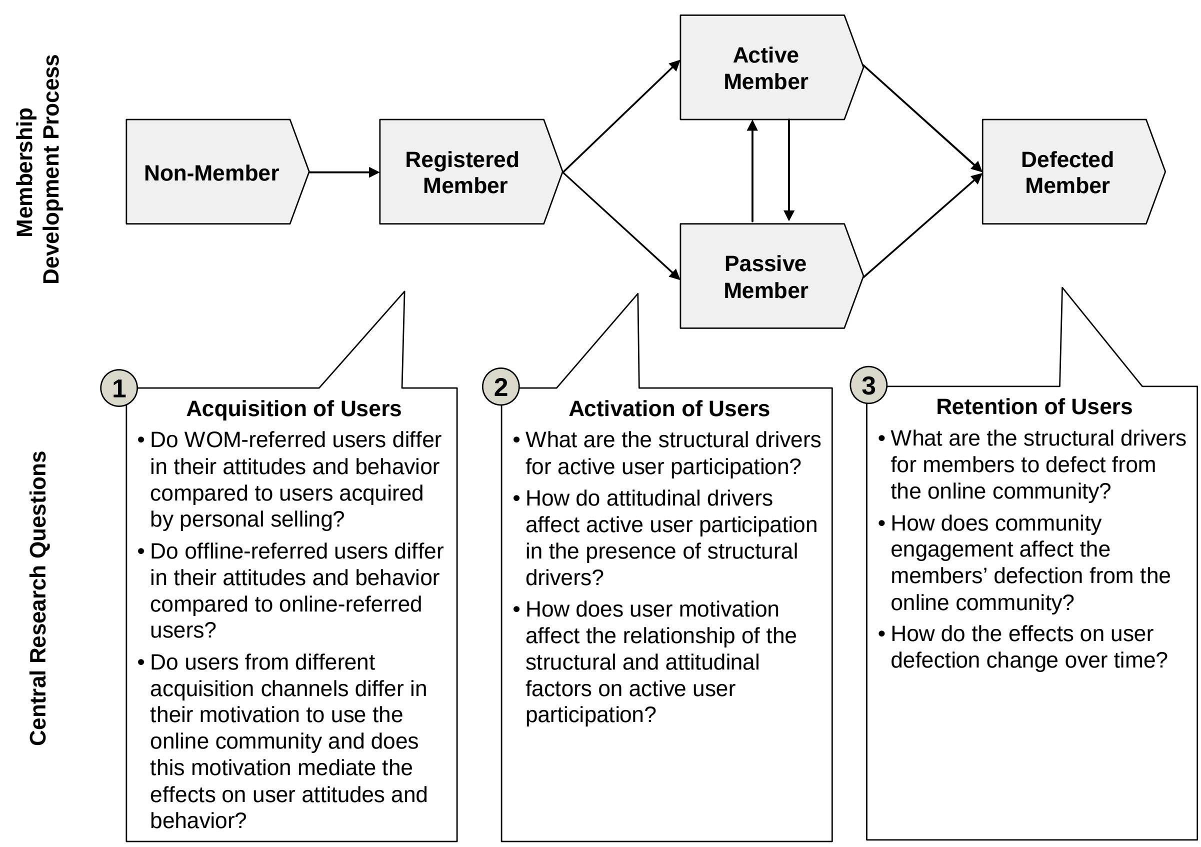


Figure 8: Central Research Questions along the Membership Development Process

One main theme of all three studies is the consideration of the social context in which the relationship between user behavior, user attitudes and its drivers occur. Especially in Studies 2 and 3, the social structure and network positions of the individual users are taken into consideration as an explanation for their behavior on the platform. The membership development process is used as a frame that facilitates the identification of appropriate actions, which community operators can take to ensure the ongoing success of the online community. Taking different perspectives along the membership development process on the users of the online social community might also enhance an understanding of the community operator about the structure of their user base.

Figure 8 provides a brief overview of the conducted research studies. Each empirical study and its research direction are presented in detail in an own dedicated chapter. Study 1 is presented in chapter 5, Study 2 in chapter 6, and Study 3 in chapter 7. Thereby, the three chapters are similarly structured. First, the relevance of the study is discussed and its contribution and research objective are presented. Then, the theoretical context (with reference to the concepts presented in chapter 3) and the link to online community research are described. Thereafter, the relevant factors included in the research model are introduced and hypotheses are developed, which will be tested in the empirical study. A short introduction to the empirical data and the results of the analyses are presented. At the end of each study, the results are discussed from a theoretical and managerial perspective. Some limitations of each study are also addressed. In the following, an overall introduction to the research design, research object and data is given.

4.2 Research Design and Data Collection

For the analyses of the three empirical studies, data from an online social community is used. Because of the specificity of the research questions, it was necessary to collect a set of data through different methods. The research object and data collection process are described in more detail in this section.

4.2.1 Research Object

The research object is an online social community in Europe. The type and purpose of the community can be described by using the classification scheme presented in chapter 2.2. Overall, it is an organization-sponsored online community with a commercial purpose. The community is the core business for the community operator. In that, the revenue model is based to a large part on advertising revenues. Further, it is a consumer-oriented community, which is focused on the private life of its users, in particular on topics concerning leisure time activities and events. An important feature is that community users can maintain profiles and articulate their list of friends, which is a defining characteristic for a social networking community (Boyd and Ellison 2008). The community operator provides its users with a platform to interact with other users by various functionalities, including sending messages, guestbook entries, comments on articles and groups, and ratings of pictures, profiles and groups. In addition, the provider does not rely solely on user-generated content but also integrates firm-produced editorial content. The content and the entire platform are locally organized, i.e. the online social community is structured around several regions or cities with local content from those regions. Consequently, it has a geographical focus as it provides localized content to its users. Although editorial content is provided, user-generated content is still one of the driving forces for success because the amount of firm-generated content is limited and a higher quantity of user-generated content is necessary to ensure the sustainability of the online community. Members need to interact and exchange information on the platform to establish a vivid community. Because it incorporates different functionalities to connect and interact with other users as well as interesting content of leisure related topics, it serves predominantly social and information needs of its users. Moreover, it is a closed community, only open for registered members to participate. Users need to register in order to get full access to the content of the online community.

There are several reasons why this online social community provides a valuable setting for the empirical studies of this dissertation:

- The online community is vivid. There is a sufficient number of users and active user participation behavior on the platform. At the end of 2010, the online community had more than 30,000 active members, who were connected to other users. Those users showed more than 300,000 activities in one month.
- Leisure time activities constitute a broad topic for different kinds of users. In addition, the community provides functionality for user interaction, which is common in many different online social communities. Therefore, the application of the findings to other online community contexts is possible due to its more general nature.
- The platform offers functionality to maintain personal user profiles with general information on the person, profile pictures and articulation of friend lists. Being connected to other users and the information about these connections is a critical element which makes it possible to investigate the explicit network structure and user positions in the online social community.
- The online social community already exists for several years. During the observation period the functionality of the community and the layout in the core areas only changed marginally, which is of particular importance for the longitudinal study (Study 3).
- It provides an appropriate context for the analysis and comparison of different interpersonal communication channels to acquire new users. Based on its hybrid content offering (firm- and user-generated content), commercial and non-commercial person-to-person marketing-communications occur naturally because they are among the most important and cost-efficient marketing means for social media service providers.

- Users of online social communities often meet not only online but also offline, as both worlds can intertwine (Wellman and Hampton 1999). This is especially true for online communities, which offer local content. Thus, the platform also provides an interesting context to investigate factors that are related to the offline world.
- Finally, post-adoption behavior is easy to observe because the community operator is able to track participation data in the underlying database system and hence allows for retrieving actual behavioral data. Because of this fact, online businesses have successfully been used to investigate marketing relevant topics (e.g., Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and Hanssens 2008).

In this thesis, the focus is on the described online social community that served as the research object for the empirical studies. Because of the broad purpose of the community and its comparable functionality to other online social communities, it provides a rich context to test the research questions. A positive aspect of researching one community platform is that the users have the same understanding of the research object and are embedded in the same technical infrastructure and purpose of the community. Observing different online communities simultaneously could lead to undesirable variances between users of different communities with regards to some factors. For example, the participation in the community could differ in general, as the behavior is influenced by the functionality and the context of the respective online community. In addition, setting the focus on one online community is a common practice in online community research (e.g., Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Trusov, Bodapati, and Bucklin 2010).

4.2.2 Data Collection

In order to answer the research questions and enable the test of hypotheses on the relationship between certain variables, data is collected through different channels. The empirical studies are based on two data sources. First, anonymous data on the users' characteristics, behavior, and network data is collected by the community operator. Second, a survey is conducted asking for additional member characteristics, perceptions, and attitudes toward the online social community.

Objective Data. Objective data was collected by the community operator. To protect the users' privacy, each user is represented by an anonymous user ID in a way that still allowed for the observation of user behavior and characteristics. Overall, three types of data are available: user characteristics, participation data, and network data.

User characteristics are constant variables, including demographic data (age, gender and registered region), registration date, the date of their last login on the platform, and type of membership (verified or not verified[41]).

Participation data includes active and passive user participation in the online social community. Active participation is broken down into messages sent to other users, guestbook entries written, virtual gifts sent to other users, comments on articles and groups, and submitted ratings of photos, groups, and other users' profiles. Each time the user pursued an activity it was counted as one contribution. Passive participation includes incoming activity from messages, guestbook entries received, virtual gifts received, and profile ratings received from other users. Participation data is aggregated on a monthly basis.

The *network data* describes the friendship connections of all users in the network. The network is a closed and complete network, consisting of all members registered with the online social community. A friendship tie is a mutual, undirected tie. Both users need to confirm the friendship to establish the online tie between them. When the friendship is confirmed by both, the respective other user appears in the user's friends list on the profile. All friendship ties are included in the entire network data on a monthly basis, representing all ties at the end of the respective month. Therefore, changes in friendship patterns can be tracked over time (which is especially important for Study 3). From this data, the current number of active ties per user is available for each month. Network indicators (as presented in chapter 3.2.2) are calculated on the basis of this friendship network.

Survey Data. The purpose of the questionnaire was to collect data, which is critical for testing the hypotheses in the empirical studies but was not available as objective data. The survey was developed on the basis of the research questions and served the analyses of Studies 1 and 2. The objective of Study 1 is to compare different groups of users according to their acquisition channel in their user behavior and attitudes towards the community. Study 2 includes the users' attitudes as antecedents of their participation in the community. Consequently, the questionnaire consisted of three parts: a) general questions on the online social

[41] Verified membership can be achieved by applying for authentication of one's personal data. The authentication process requires true personal data, specifically real name and address, where the authentication code can be sent. When the code is correctly entered on the community website, the user appears as a "verified" member, which is presented as a symbol on the user's profile page and is visible for all other users.

community membership and community participation, b) attitudes and perceptions towards the online community of interest, and c) demographic data.

The *general questions* include the users' number of contacts in the online community, thereof the number of contacts they regularly meet in the offline world, the tenure in the online community, and the frequency of visits. In addition, the users had to answer questions on the acquisition channel and WOM referrals about the community. First, the users were asked how they initially were attracted to the online social community. Multiple choices were included: referral from friends and acquaintances, contact to an employee (personal selling), or several other media (print media, radio, online search, etc.). Thereafter, all users had to specify how often they provided and received WOM referrals to use the online community via online channels (e.g., sending an E-Mail or an online link) and offline channels (personal conversation with another person). This set of questions was of particular interest for Study 1. The different acquisition channels were based on the answers to these questions. The construction of the group building variables is explained in more detail as part of Study 1 (chapter 5). Table 5 illustrates the operationalization of the questions.

Through an intensive literature review, relevant user *attitudes and perceptions* for the empirical studies had been collected and assessed. The most important factors with respect to the research context were included in the empirical studies. In Study 1, the main interest is to investigate the impact of the acquisition channels on value-driving attitudes and behavior. Therefore, important variables in this respect have been chosen. In fact, satisfaction, identification, and WOM intention are of specific relevance and are frequently used attitudes in online community and marketing research (e.g., Algesheimer, Dholakia, and Herrmann 2005; Dholakia, Bagozzi, and Pearo 2004; Woisetschlaeger, Hartleb, and Blut 2008). In Study 2, the effect of the relational dimension of social capital is represented by self-reported attitudinal measures. Besides identification and satisfaction, reciprocity is additionally included in the survey as a facet of relational capital. Further, the literature review revealed that the most important motivational factors are social and informational needs (cf. chapter 2.5). This is underlined by the context of the research object in this study because it mainly provides functionality and content that targets these two needs. Because of the interest in understanding different groups of users with respect to their motives, social interaction value and information consumption value were chosen to represent user motivation (e.g., Dholakia et al. 2009). Further explanation of the respective variables is given in the hypothesis sections of Study 1 and Study 2 (chapters 5.2.2.1 and 6.3.2) and in chapter 2.5.4.2, which provides a review of research on most of the variables.

Demographic data is included in the last part of the survey where the users needed to fill in data on age, gender, education, and Internet proficiency.

Survey Development and Administration. The survey included the described survey data. The questions comprising the data are illustrated in Table 5. General and demographic questions are operationalized as open questions (e.g., number of friends in the network), or multiple choices for the answers are given. The attitudinal and perceptual variables are measured by one- or multi-item constructs. The items were derived from existing scales of past research and adapted to suit the context of this study on an online social community to ensure face and content validity. Satisfaction was measured by an adjusted scale of the American Customer Satisfaction Index from Fornell et al. (1996). Identification items were taken from a scale of Algesheimer, Dholakia, and Herrmann (2005). WOM Intention is measured as a one-item scale, which was used for instance in the recommendation intention construct of Maxham and Netemeyer (2002). Reciprocity is derived from a scale of Wasko and Faraj (2005). For the motivational constructs, the items are adapted from Dholakia, Bagozzi, and Pearo (2004) for social interaction value and from Mathwick, Wiertz, and de Ruyter (2008) for information consumption value. The answers for all multi-item questions are based on seven-point semantic differentials or Likert-scales.

The survey was pre-tested using different methods. Personally administered interviews with marketing researchers, online community users, and the community operator were conducted to evaluate the clarity, completeness, and interpretation of the questions. Feedback was provided on any lack of clarity or difficulty to fill out the questionnaire. Thereafter, a quantitative pre-test of the survey was conducted in two steps. First, a convenience sample of 221 users of different social networking sites was drawn to test the general applicability of the survey to the online community context. Second, a test-sample of 71 users of the online community of interest answered the questionnaire to verify the suitability of the terminology and the clarity of the questions and scales. According to the results and the feedback of the respondents, adjustments in the scales were made. The constructs, final items and anchors of the answers are presented in Table 5.[42]

---

[42] The items were formulated in the language of origin of the online social community. The items were translated and retranslated to ensure that the understanding remained the same. The original scales (that were used as a basis for survey construction) taken from past research studies can be found in Appendix 2.

Multi-Item Constructs		Measures
Satisfaction[1)]		With Community XYZ as a whole I am...
	1	Very dissatisfied / very satisfied.
		Community XYZ as a whole…
	2	Falls short of expectations / exceeds expectations.
	3	Is not very close to ideal community / is very close to ideal community.
Identification[2)]	1	I am very attached to my Community XYZ contacts.
	2	My Community XYZ contacts and I have many similarities.
	3	The friendships I have with my Community XYZ contacts mean a lot to me.
	4*	I see myself as a part of my Community XYZ community.
Reciprocity[2)]	1	I trust that other members of Community XYZ would contribute something to the community if I contribute something.
	2	I know that other members of Community XYZ would contribute something, so it is only fair for me to contribute something as well.
Information Consumption Value[2)]		Why do you use Community XYZ? I use Community XYZ, …
	1	because I can get relevant Information and photos on Community XYZ.
	2	because I think of Community XYZ as a good resource for information and photos.
	3	because I can get valuable information and photos on Community XYZ.
Social Interaction Value[2)]		Why do you use Community XYZ? I use Community XYZ, …
	1	to have something to do with other Community XYZ users.
	2	to stay in touch with other Community XYZ users.
	3	to keep contact to other Community XYZ users.
Single Questions		**Measures**
WOM Intention[2)]		I would recommend Community XYZ to my friends and acquaintances.
Frequency of Use[3)]		How often do you use Community XYZ?
Marketing Channel[4)]		How was your attention drawn on Community XYZ?
WOM online reception[5)]		How many people recommended, through sending an E-Mail or online link for registration (online), that you should register at Community XYZ?
WOM offline reception[5)]		How many people recommended, in a direct conversation (offline), that you should register at Community XYZ?
WOM online provision[5)]		How many people did you advise, through sending an E-Mail or online link for registration (online), that they should register at Community XYZ?
WOM offline provision[5)]		How many people did you advice, in a direct conversation (offline), that they should register at Community XYZ?
Share of Real Friends[5) 6)]	1	How many friends do you have in your friends list in Community XYZ?
	2	With how many of your friends from your friends list do you have regular contact in the real world?

1) anchored with 7-point semantic differential
2) 7-point Likert scales, anchored 1=strongly disagree and 7=strongly agree
3) scale ranging from „less than once a month" to „7 days per week"
4) multiple-choice answers provided: recommendation of friends; employee at a leisure time event; advertising; online search; editorial article; online article; content on other online communities; other;
5) open question
6) ratio of question 2 divided by question 1 was taken as a measure
** Identification Item 4 has been dropped due to low item-to-total correlation and low content validity in study 1 and 2*
Note: Community XYZ was replaced by the community name

Table 5: Survey Constructs, Items, and Questions

The final survey was constructed with an online tool to provide easy online access to the users. It was sent out by the community operator together with an announcement, which explained the purpose of the survey and guaranteed confidentiality of responses. In order to attract users to participate in the survey, ten gift vouchers were drawn as an incentive for filling out the questionnaire. The survey was online for four weeks in September/October 2010. After two weeks a reminder was sent to all users who had not participated at that time. The overall number of relevant online social community users comprised around 30,000 members at the time of the survey. The relevant user group for the survey-based studies included all registered users who were still active, i.e. who still used the online community at the time of the survey, and who are integrated in the online social network. As pointed out in chapter 2.1, online communities are defined by the relationships between users. Users who do not share any ties to other users are peripheral and not involved in the online social network (Freeman 1978/79). Those users are isolated from direct involvement with other users in the network and cut off from active participation in the ongoing communication process (Freeman 1978/79). For social network analysis, the data should include all social units on which measurements are available (Wasserman and Faust 1994). Because an absence of connections between users often limits the level of interaction (e.g., ability to consume and/or exchange digital content), users with friend connections are of major interest in social community studies (e.g., Trusov, Bodapati, and Bucklin 2010). Therefore, the data collection concentrated on those users being integrated in the online social community and connected to other users.

The final survey sample includes 689 usable responses. Invalid responses were excluded from the sample. Those include responses from employees of the online social community, responses with comments such as ‘I do not know the community’, responses with the same value for more than 90% of all Likert-scale based items, and response times that were dubiously short (e.g., when the user just clicked through the survey to take part in the raffle provided as an incentive to participate). Short response time is defined as being lower than five minutes, with a mean fill out time of 12.4 minutes for completed surveys.[43]

The empirical studies (presented in the following chapters 5, 6 and 7) are based on the survey and the objective data presented above. They use different sets of variables that are appropriate for the respective context of each study. For Study 1, the channel by which a user was acquired by the online community is used to define different groups of users, who are

[43] The original sample included 819 completed surveys from active community users.

compared along different self-reported constructs and participation data. However, not all users can be identified to belong to one of the groups of interest. Study 2 includes important measures of social capital, which is related to the performance of the individual users. While the measures of the structural dimension mainly come from objective data, additional self-reported variables are used to test their contribution to explain user behavior. Active participation is retrieved from the objective data. Study 3 is solely based on objective data. Because longitudinal data is used, survey constructs were not available throughout the entire observation period. Therefore, its analyses are all based on objective behavioral data, network data and user characteristics. Because the analysis applies longitudinal data, a specific group of the user population is observed.

Due to different focuses of the respective studies, the size of the used samples for the analyses varies. This is because not all users show the same characteristics and different groups of users are analyzed based on their general characteristics and responses in the survey. In order to describe the respective samples and variables of each study more accurately, sample characteristics, descriptive statistics, measurement model evaluation, including the tests for reliability and validity of the survey constructs, are provided in the respective chapters for each of the three empirical studies (cf. chapters 5, 6, and 7). In addition, non-response bias tests and verification tests of the models are provided in each study according to the different study populations and constructs in focus.

5 User Acquisition: A Comparison of Interpersonal Acquisition Channels for Online Social Communities

5.1 Interpersonal Communication Channels as Means for User Acquisition

Customer acquisition is an important process that should bring the “right” customers to the firm (Blattberg and Deighton 1996; Hansotia and Wang 1997). Thereby, firms use different marketing communication channels to acquire new customers. Among those channels, interpersonal communications are effective means to achieve this goal (e.g., Reinartz, Thomas, and Kumar 2005; Wangenheim and Bayón 2007). Basically, interpersonal acquisition channels can be differentiated into consumer-initiated communication (word-of-mouth; e.g., Brown and Reingen 1987; Harrison-Walker 2001) and firm-initiated communication (personal selling; e.g., Crosby, Evans, and Cowles 1990; Szymanski 1988). For example, employees of online services, like local news portals, can personally contact potential customers at local events and encourage them to use the service (personal selling). In addition, the customer base of many online firms, like Facebook, grows through word-of-mouth (WOM) recommendation, which is spread between friends in conversations or via some kind of “tell-a-friend” functionality on the website. Those interpersonal communication modes work in a similar fashion but are anticipated to impact customer attitudes and behavior differently, as the relationship between the sender and the receiver of the transmitted message differs, for instance, because customer-to-customer communication is perceived as more trustworthy (Murray 1991). Although interpersonal communication has gained large academic attention in the past (de Matos and Rossi 2008), there is a need for a more differentiated view to better understand different acquisition channels because they are expected to attract different kinds of customers, for example with respect to their value and loyalty (Verhoef and Donkers 2005).

With different interpersonal channels at hand, firms ask themselves which channels to use and what impact a channel has on relevant variables that drive firm value and customer equity. Past research shows that WOM communication is more effective than traditional advertising or media messages (e.g., Engel, Kegerreis, and Blackwell 1969; Herr, Kardes, and Kim 1991; Katz and Lazarsfeld 1955; Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and

Hanssens 2008). Also customers acquired through firm-offered referral programs show higher value to the firm than non-referred customers (Schmitt, Skiera, and van den Bulte 2011). Nevertheless, personal selling can play an important role in relationship building with the goal of creating satisfied and long-term customers (e.g., Crosby, Evans, and Cowles 1990; Reynolds and Beatty 1999; Solomon et al. 1985). Although both communication modes help to acquire new customers, a comparison of the effects of WOM and personal selling on customers' attitudes and behavior has not been tested.

In addition, WOM is not only occurring in the offline world. New technologies emerged that give the customer the opportunity to transmit WOM messages via online tools. For example, Trusov, Bucklin, and Pauwels (2009) use online recommendation behavior to compare the effectiveness of WOM and other marketing channels for social networking sites and reveal that WOM has a strong impact on customer acquisition and is more effective than other marketing channels. One specific shortcoming of this approach is that WOM communication occurs both online and offline. The Keller Fay Group (Keller and Berry 2006) found that, although online WOM has recently gained more popularity, offline is still the predominant WOM channel with around 90% of WOM communication occurring in offline channels. For marketers, the question arises how effective each WOM channel is and on which WOM channel to focus, online or offline.

As various marketing communications can be used, it has been demonstrated that customers acquired through different channels generate different value (Lewis 2006; Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and Hanssens 2008). Thus, different channels attract distinct types of customers according to their characteristics. Although customers show different value and attributes, a balanced customer portfolio is critical to distribute risk and reduce the dependence on one specific customer segment (Tarasi et al. 2011). For example, in online social communities it is important to bring together both contributors and readers of content to increase the total value of the platform (e.g., Ridings, Gefen, and Arinze 2006). In addition, acquisition channels show a different effect on customer value for different customer segments (Schmitt, Skiera, and van den Bulte 2011). Therefore it is important to understand what kinds of customers are attracted by which channel and how those customers behave.

The differences between employee-to-consumer communication (personal selling) and consumer-to-consumer communication (WOM), and between online and offline WOM recommendations are therefore at the heart of this study. The study contributes to the existing literature on marketing communication channels and WOM in three ways. First, it elaborates on the comparison of the two communication channels of personal selling and WOM. It is shown that post-adoption attitudes and behavior of customers acquired by the two channels differ. In the past, only the different effects of WOM and personal selling concerning the decision making process have been investigated (e.g., Katz and Lazarsfeld 1955), and the effectiveness of WOM and other marketing channels in general have been compared (Schmitt, Skiera, and van den Bulte 2011; Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and Hanssens 2008). Second, a distinction of different types of WOM is proposed and post-adoption behavior and attitudes for online- and offline-referred customers are compared. This comparison has received no academic attention so far. This study demonstrates that customers from offline WOM referrals outperform online-referred customers on some of the observed dimensions. Third, customer motivations for participating in the online social community are investigated in order to take channel-specific user characteristics into account. The study compares the motives to use the online social community by interpersonal acquisition channel and tests for the impact of motivation as a mediator.

5.2 Theoretical Background and Hypotheses

5.2.1 The Social Context of Interpersonal Communications

Interpersonal communication modes work in a similar fashion as they describe the transmission of a marketing message between two actors. However, they differ in the quality of the relationship between the sender and the receiver. Traditionally, personal selling includes a presentation of arguments within a conversation between the sales representative and one or more potential buyers with the goal of selling products and services (Meffert 2000). Palmatier et al. (2006) found that relationship marketing is more effective when relationships are built with an individual person (e.g., the sales representative) rather than a selling firm, which favors personal selling over mass media communication. In this study, personal selling takes place in short employee-user interactions in the offline world. The sales representative approaches prospective users at local events by referring to the content of the platform and passing on the online address of the online community website. Thus, there is some short interpersonal exchange between an individual and the sales person. In contrast to firm-initiated, commercially driven interpersonal communications, WOM is an “informal, person-to-person communication between a perceived non-commercial communicator and a receiver regarding a brand, a product, an organization, or a service” (Harrison-Walker 2001, p. 63). WOM occurs not only offline (traditional WOM), but also takes place in the online environment (online WOM or eWOM). There are many different contexts in which online WOM can appear, including recommendation platforms (e.g., Godes and Mayzlin 2004; Hennig-Thurau et al. 2004), weblogs (Kozinets et al. 2010), E-Mails (De Bruyn and Lilien 2008; Huang, Lin, and Lin 2009), or tell-a-friend functionality (Trusov, Bucklin, and Pauwels 2009). The main difference between online and offline WOM is the separation of the WOM sender and the receiver by space and/or time. In this study, offline WOM is related to a personal conversation between the WOM sender and the receiver, while online WOM refers to the recommendation of the online community via E-Mail or sending an online link.

As interpersonal communication is a social phenomenon, Reingen and Kernan (1986) suggest that the social-structural context within which interaction occurs should be accorded explicit recognition. Scholars underlined the importance of investigating the social aspects and dyadic quality of the communication, both from the personal selling perspective (e.g., Crosby, Evans, and Cowles 1990; Webster 1968) and in WOM research (e.g., Bansal and Voyer 2000; Brown and Reingen 1987). In the context of commercial interpersonal relationships, Evans (1963, p. 76) states that “the sale is a social situation involving two persons. The interaction of the two persons, in turn, depends upon the economic, social, and personal characteristics of each of them. To understand the process, however, it is necessary to look at both parts of the sale as a dyad, not individually”. It has already been postulated by marketing scholars not only to take individual perspectives into account but to use a more holistic view on social and networked environments instead (e.g., Achrol and Kotler 1999; Algesheimer and Wangenheim 2006). As already outlined in chapter 3.2, social network theory assumes that human behavior is affected by the social network of an individual where social structures are in the center of human behavior. Two basic concepts from social network literature have received specific attention in studying interpersonal communications from a marketing perspective: tie strength and homophily (e.g., Brown, Broderick, and Lee 2007; Brown and Reingen 1987; De Bruyn and Lilien 2008). These two concepts help to better understand the relationship between the sender and receiver of the marketing message (for a more detailed description of both concepts refer to chapters 3.2.1.1 and 3.2.1.2).

With respect to the dyadic relationship it is expected that WOM communication involves stronger and more homophilous connections between sender and receiver than personal selling. In addition, WOM is perceived by consumers as more credible than firm-initiated communications (Murray 1991). WOM sender and receiver know each other and often share the same interests, while sales representatives in a B2C environment are less, if it all, connected to the prospective customers. With respect to Granovetter's (1973) definition of tie strength, WOM referrals occur in relationships where sender and receiver have more frequent contact, a more intimate relation, and share higher norms of reciprocity, which makes their tie stronger than the relationship with a sales representative. As Hung and Li (2007, p. 486) point out, "although eWOM is similar to personal selling in that it provides explicit information, tailored solutions, interactivity, and empathetic listening, it has a lower distance between the source of communication and the receiver than marketer-induced communications".

In comparison with online WOM, offline WOM referrals are expected to be stronger and more credible, and to be occurring between more homophilous ties. If the WOM sender and the receiver regularly meet in the offline world, a recommendation would most likely be exchanged in a personal conversation. Therefore, using the online channel would indicate a more distant relationship than offline referrals. Two theories support this argument. The concepts of social presence and social context cues share the view that the reduction in cues should cause communication in online settings to be more impersonal and nonconforming than in face-to-face settings (Parks and Floyd 1996; Short, Williams, and Christie 1976; Sproull and Kiesler 1986). Compared to face-to-face interaction, computer-mediated communication is characterized as low in social presence because it typically allows for little exchange of nonverbal cues. Sproull and Kiesler (1986) found evidence that E-Mail reduced social context cues which are usually available in face-to-face communication. In addition, E-Mail includes only limited signs of personal similarity between sender and receiver (Parks and Floyd 1996), i.e. the perceived similarity, and thus the perceived fit to one's own needs, is lower for online WOM receivers. Studies on online and offline friendships indicate that offline friendships show a higher quality of friendship, including higher commitment, understanding, network convergence, and interdependence regarding the relationship (Chan and Cheng 2004; Parks and Roberts 1998). Moreover, Rothaermel and Sugiyama (2001) argue that offline communication strengthens the relationship-building process of virtual communities and might be a stronger factor in explaining community members' embeddedness.

5.2.2 The Effects on User Attitudes and Behavior

The main hypothesis is that users' post-adoption behavior and attitudes are affected by the type of interpersonal communication they are exposed to. Similar to Weitz and Bradford (1999), who propose that relationship quality should be measured by a multidimensional output measure, different value-driving factors are taken into account for assessing the quality of the customers' relationship to the online social community. Thus, attitudinal (satisfaction, identification, WOM intention) and behavioral (community participation, WOM provision) variables are included in this study. The selection of the dependent variables is based on their relevance in marketing and online community research. Indeed, satisfaction and identification are two frequently studied variables that measure the attitudes towards a service provider, which are positively related to customer loyalty to the service (e.g., Bolton 1998; Cheung and Lee 2009; Dholakia, Bagozzi, and Pearo 2004; Gustafsson, Johnson, and Roos 2005; Woisetschlaeger, Hartleb, and Blut 2008). Behavioral factors, such as usage frequency, active contribution and passive consumption, are focused on in various community research studies (e.g., de Valck et al. 2007; Wang and Fesenmaier 2004b; Woisetschlaeger, Hartleb, and Blut 2008; see also the literature review in chapter 2.5). They are of special interest for marketers because participation increases the attractiveness of the community and could yield monetary outcomes for the community operators

In the following, two sets of hypotheses are developed. First, direct effects of the interpersonal communication channels for user acquisition on the attitudinal and behavioral factors are hypothesized. Second, hypotheses on the mediation effect of user motivation between the interpersonal communication channels and the attitudes and behaviors are presented.

5.2.2.1 Direct Effects of the Interpersonal Acquisition Channels

Satisfaction. Satisfaction is regularly based on the expectancy-disconfirmation paradigm and is most often measured as the difference between customer expectations and the actual service or product performance (Oliver 1997). When the product or service exceeds customers' expectations they will be satisfied. In strong tie relationships, senders know more about the needs and preferences of receivers because they are in more frequent and intimate relations with them (Granovetter 1973). They are therefore more concerned about the receivers' needs (Clark, Mills, and Powell 1986). Strong ties are used dependent on the requisite knowledge for understanding the receivers' individual characteristics and needs (Kiecker and Hartman 1994). Accordingly, WOM senders, and more specifically offline WOM senders, should be more familiar with the needs of their contacts. Consequently, the expectations of the prospective users are more likely to be met when information comes from close personal contacts because the differences between needs and service performance are expected to be smaller based on the prior knowledge. Information from sales representatives, who want to bring the individual to the platform, have lower credibility and higher risk that the expectations cannot be met, which potentially leads to less satisfaction after adopting the service. Further, because WOM has been shown to be an effective means of reducing pre-purchase risk (Arndt 1967; Murray 1991), WOM-referred users should be less likely to experience cognitive dissonance than users coming from personal selling. Because of the negative relationship between dissonance and satisfaction (Oliver 1997), WOM and offline-referred users should be more satisfied in the post-adoption phase. Supporting this, Wangenheim and Bayón (2004a) observed that referred customers are more satisfied and loyal after switching to the firm than non-referred customers. Moreover, offline conversations can be richer in information with the presence of social context cues (Markus 1994; Sproull and Kiesler 1986). This results in more social attitudes and better understanding of the offline than the online WOM receiver on what to expect from the service. Therefore, it is hypothesized:

H1a: Compared to users acquired by personal selling, WOM-referred users are more satisfied with the online social community.

H1b: Compared to online WOM-referred users, offline WOM-referred users are more satisfied with the online social community.

Identification. Identification with a group can be considered as the sum of the relationships a person has with other users of an online social community. In this context, social identity describes the "individual's identification with the group in the sense that the person comes to view himself or herself as a member of the community, as 'belonging' to it" (Dholakia, Bagozzi, and Pearo 2004, p. 245; see also chapter 3.4.1 for an introduction to the social identity concept). The integration into the existing community is therefore important to establish a strong identification with the members of the online social community. A better integration of users with strong tie relationships to the sender can be explained by Granovetter's (1973) assumption, that if A (the WOM receiver) and B (the WOM sender) are connected by a strong tie and B and C (an existing contact of A in the community) are connected by a strong tie, then A and C must also be connected by a stronger tie. For weaker sender-receiver ties the connection with other members of the network would be less likely. Users coming from referrals build relationships with other users faster, in higher volume, and with stronger ties to their overall individual sub-networks. When acquired by a sales representative, the integration in the community is expected to be less deep. In addition, as strong relationships and homophily are interrelated (e.g., McPherson, Smith-Lovin, and Cook 2001), higher similarity in the developed network of the new user will support the identification with other users. Because offline WOM referrals are perceived as more trustworthy due to higher social presence and social context cues, the offline-referred user should be more likely to identify with the group of people in the online community because he knows better what to expect. As online WOM receivers will normally not be as integrated as offline WOM receivers, their identification with the community will be lower. Recent studies showed that offline activities increase the belonging to a virtual community, strengthening the links between members (Koh and Kim 2003/04; Lin 2007). This leads to the following hypotheses:

H2a: Compared to users acquired by personal selling, WOM-referred users show higher identification with the online social community.

H2b: Compared to online WOM-referred users, offline WOM-referred users show higher identification with the online social community.

Online community participation. Participation and usage behavior are key variables for generating income for online social communities, especially for advertising based business models. Participation depends on how well the user is integrated in the community, that means how strong his relationship to the community is (e.g., Algesheimer, Dholakia, and Hermann 2005; Dholakia, Bagozzi, and Pearo 2004; Wasko and Faraj 2005). Stronger and more homophilous ties facilitate the integration of users into existing online network structures. Granovetter (1973, p. 1362) supports this view by stating, that “the stronger the tie between A and B, the larger the proportion of individuals in S [the social network of the WOM sender] to whom they will both be tied, that is, connected by a weak or strong tie.” This overlap in their friendship circles is predicted to be least in personal selling relationships, where there are predominantly no initial connections to the community; and it is most in offline referrals, as the WOM sender is already embedded in the community and has a stronger relationship to the WOM receiver. Users exposed to WOM integrate more likely and quickly into the online social network and establish more shared contacts with the WOM sender. Better integration in the community would enhance the interaction with other members, and thus the overall participation (Algesheimer, Dholakia, and Hermann 2005).

Recent research on online services found that users acquired through traditional marketing show lower log-in behavior than WOM-referred customers (e.g., Villanueva, Yoo, and Hanssens 2008). The hypotheses include three different measures of participation because different perspectives to investigate the impact of the communication channel can contribute to a better understanding of the effects on participation (see also the discussion on active vs. passive participation in chapter 2.5). Therefore, it is hypothesized:

H3a-5a: Compared to users acquired by personal selling, WOM-referred users are participating 3a) more actively by contributing content, 4a) more passively by receiving more content, and 5a) more frequently in the online social community.

H3b-5b: Compared to online WOM-referred users, offline WOM-referred users are participating 3b) more actively by contributing content, 4b) more passively by receiving more content, and 5b) more frequently in the online social community.

Word-of-mouth provision. Gilly et al. (1998) found that the consumers' willingness to give referrals correlates with the significance they give to such information themselves, suggesting that WOM receivers will tend to be WOM senders as well. Moreover, people who exhibit high similarity share the same interest in the online social community, be it either the content or the interaction. Being acquired through strong tie relationships would indicate that the receivers would evaluate the service similar to WOM senders, resulting in WOM recommendations from receivers to other people. In WOM recommendation networks it can be seen that WOM messages flow through networks of strong and weak ties (Brown and Reingen 1987), suggesting that WOM is often forwarded to further people. Recent research found, that customers acquired through WOM generate more future WOM than those acquired through other marketing channels (Villanueva, Yoo, and Hanssens 2008). Moreover, Wangenheim and Bayón (2004a) showed that referred customers provided significantly more positive WOM to other people than customers who did not get any referral.

For the comparison of online- and offline-referred users, the view of Gilly et al. (1998) can be extended into the differentiation of the two WOM channels. This means that offline WOM receivers tend to give more offline WOM, and online WOM receivers tend to provide more online WOM, because of the significance they gave themselves to these channels. In a more general perspective, WOM intentions should be higher for offline WOM-referred users, as those users are expected to be better integrated in the online social community. Thus, the online social community is more important to them which results in higher intentions to give positive recommendations about the service. This leads to the following hypotheses:

H6a-8a: Compared to users acquired by personal selling, WOM-referred users are 6a) more willing to give WOM in the future, 7a) more active in giving online WOM, and 8a) more active in giving offline WOM.

H6b-8b: Compared to online WOM-referred users, offline WOM-referred users are 6b) more willing to give WOM in the future, 7b) less active in giving online WOM, and 8b) more active in giving offline WOM.

Acquisition of different customer types. Different acquisition channels are expected to address distinct user groups, emphasizing different needs. Therefore, user motivation is included in this study to reveal potential differences between user groups. The involved individual motives for the use of media are often explained based on the uses and gratifications approach. In recent years, this approach was used to study motives of individuals to use the Internet (Flanagin and Metzger 2001; Papacharissi and Rubin 2000), virtual communities (Dholakia, Bagozzi, and Pearo 2004; Sangwan 2005), and social networking sites (Park, Kee, and Valenzuela 2009). Different reasons to use online media have been identified in these studies, with the value of the single motives depending on the purpose or type of the community (e.g., Dholakia, Bagozzi, and Pearo 2004). Two of the main motivational factors are the value one obtains from the information that is available on the community website (information consumption value) and the value one gets from social relationships with other members (social interaction value).[44] Recent research has already demonstrated the importance of both information and social value in online communities (e.g., Dholakia et al. 2009; Wiertz and de Ruyter 2007). Information seeking is a main objective to visit websites and has been seen in many online communities as one of the key motivations (e.g., Ridings and Gefen 2004; Wasko and Faraj 2000). Social aspects are the predominant factor for the offering of online community services and even important in information- or knowledge-sharing platforms (e.g., Mathwick, Wiertz, and de Ruyter 2008).

Different communication channels may emphasize different values to the prospective user, and therefore address different needs. WOM communication is based on the intimate relationship of the sender and receiver. Social aspects are therefore central for senders to bring their friends and acquaintances to the community. In contrast, personal selling is sup-

[44] While information value is considered as a functional value the user derives from the online social community, maintaining interpersonal connectivity emphasizes the social benefits. In online social communities with a hybrid content offering those two factors can be seen as reflecting well the two main values of the platform – information and community. Although there are many other needs that can be fulfilled by using an online community, these two factors have been identified to be of specific relevance (see literature review in chapter 2.5).

ported by directly promoting the content offerings of the platform, thus pointing out information value. Personal selling is often more information driven, where sales representatives highlight advantages of the service.[45] Having a stronger relationship with the receiver, offline WOM channels should emphasize the social value as WOM senders bring their offline network into the online world. Therefore, the social interaction value is expected to be higher for users referred offline. Online communication stresses the social aspect less, as the distance is greater between sender and receiver. As both, offline and online WOM can profit from information benefits, users coming from the two WOM channels are not expected to differ in their information consumption value. The final set of hypotheses includes motivational factors, differentiated by the acquisition channels:

H9-10a: Compared to users acquired by personal selling, WOM-referred users are 9a) more driven to use the online social community through its social interaction value and are 10a) less driven to use the online social community through its information consumption value.

H9-10b: Compared to online WOM-referred users, offline WOM-referred users are 9b) more driven to use the online social community through its social interaction value and 10b) do not differ in their information consumption value to use the online social community.

5.2.2.2 Mediation Effects of User Motivation

As already noted, social values might be more accentuated in relationships between closer friends and in offline WOM conversations. In contrast, informational value would be rather emphasized in personal selling conversations, where the employee wants to convince the user by emphasizing the quality of content offered on the Website (e.g., articles, reports, photos, or videos). Both, social interaction value and information consumption value are potential drivers for people to use the online community (Dholakia et al. 2009; Wiertz and de Ruyter 2007). Thus, motivations are expected to affect all attitudinal and behavioral variables in this study, i.e. having an impact on satisfaction, identification, WOM intention and provision, as well as on user participation.

[45] Note that there is no pre-selection in the community operator’s personal selling channel with regards to the motivation of the prospective users. Community representatives rather randomly choose people at different events and introduce to them the online community website and its content offering. Personal selling communication is quiet short (in this case) and therefore only some limited information is distributed.

When users perceive high levels of social and informational value of the online social community, through which they are driven to use the platform, their expectations are more likely to be met. Therefore, based on the users' experience with the platform, higher motivational values should lead to higher satisfaction towards the online community. The greater the informational and social usefulness of information and social interaction received in the online social community, the greater will be the perception that one has helped oneself (Dholakia et al. 2009), which potentially influences the assessment of user satisfaction. In the same manner, if social value and intention to interact with other community members is high, identification is expected to be higher. With high social motives to use the community, the user is more interested in other members of the community and can thus identify more with them. If information is highly valued, the user is assumed to identify more with the topic of the community. Thereby, these users are potentially more similar to other users concerning their interest in the content and context of the community, thus leading to higher identification as well. In past research, value perceptions have been shown to impact the users' social identity (Dholakia, Bagozzi, and Pearo 2004).

Users who are highly motivated to socially interact with other users of the online community are expected to participate more in turn; otherwise they would not capture the social value they associate with the platform. In addition, greater perceived social benefits would result in a higher commitment to the community and thereby a stronger feeling help and interact with other users (Mathwick, Wiertz, and de Ruyter 2008). If information value is perceived as high, users would show increased participation on the platform in order to retrieve more of the highly valued information. Last, if the value is high, be it either social or informational value, it would be worth recommending the online community to other people, who are not members yet, because it could also be valuable for them to use the platform (fulfilling their needs). In line with the direct hypotheses on the differences of the perceived social interaction value and information consumption value for users from different channels, the following indirect effects through user motivations are hypothesized:

H11: The effect of the Marketing Channel (WOM vs. personal selling) on the attitudinal and behavioral outcomes is mediated a) by social interaction value and b) information consumption value.

H12: The effect of the WOM Channel (offline vs. online) on the attitudinal and behavioral outcomes is mediated a) by social interaction value and b) not by information consumption value.

5.3 Empirical Study – Methodology

Figure 9 summarizes the hypotheses of the direct effects developed in the previous section (the mediated effects will be addressed subsequently in chapter 5.4.4). To test the hypotheses, data from the online social community is used, which was introduced in chapter 4. There are several reasons why the online social community is an adequate research object to observe interpersonal communication. First, the observed online social community allows for a comparison of commercial and non-commercial person-to-person marketing communications. As both, firm-generated and user-generated content is available, firm-initiated and customer-initiated communications occur naturally. Second, WOM plays an important role for online social communities. They often grow mainly because of WOM recommendations and community operators often facilitate online WOM invitations to join the platform. Third, online social communities mostly cover both worlds, as online and offline social networks regularly get mixed. Thus, both phenomena can be observed (online and offline WOM). Finally, post-adoption behavior can be observed by the means of the objective data that was collected by the community operator. Consequently, both data sources – self-reported survey data and objective data – are used in the empirical analysis.

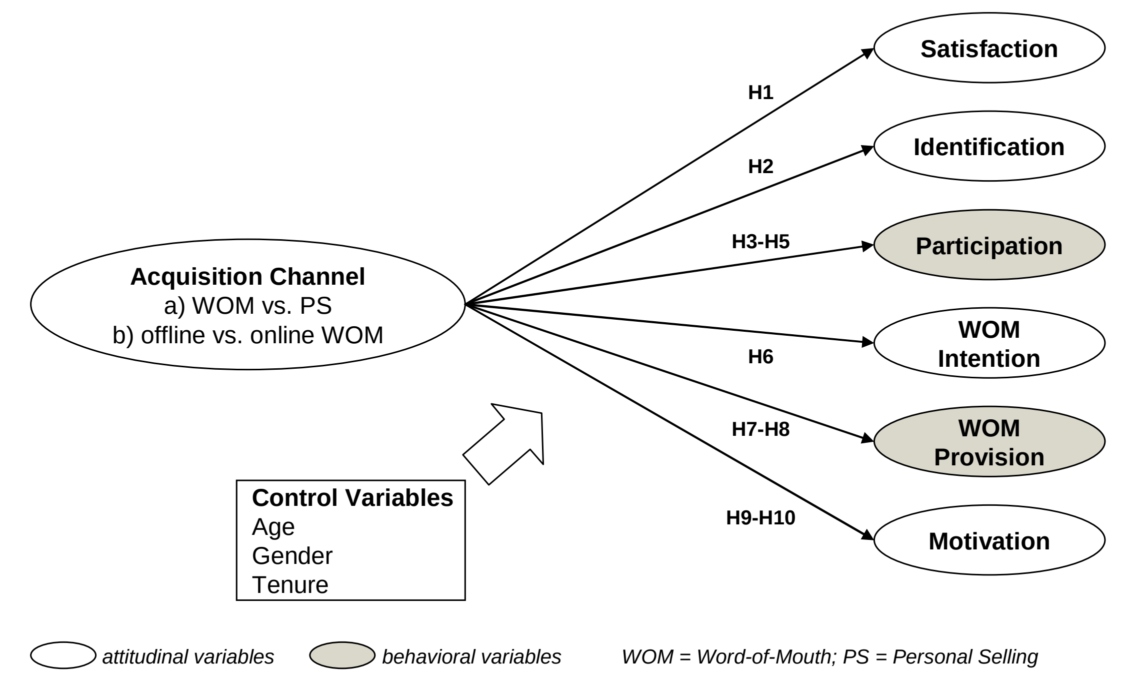


Figure 9: Study 1 – Influence Model of Acquisition Channels

5.3.1 *Data and Sample Characteristics*

Overall, 689 survey responses are used in this analysis (see chapter 4 for more information on the research design). The users in the sample stated that they are still using the online social community and that they are connected to the community, thereby indicating that they are 'true' community users. This ensures the comparability of the various communication channels because people who are not involved in the community do not have the opportunity to be as active as other users. Further, as this study draws upon the social context of the users, it is important to focus on users who are embedded in the online social community.[46]

In order to ensure generalizability of the results, non-response bias should be avoided. Therefore, it is tested for non-response bias using two approaches. First, the responses of early and late respondents are compared by using time-trend extrapolation (Armstrong and Overton 1977). Gender and education are tested by means of χ^2-tests, while all other constructs are tested using t-tests comparing the mean values. No significant differences appear at $p<.05$ level (see Appendix 3). These results suggest that non-response bias is not likely to be a major concern for the study. Second, it is tested for differences between respondents and non-respondents regarding their general characteristics concerning the platform (which are also available for non-respondents), namely age, gender, tenure, number of received profile visits, number of contacts in the users' friend lists, and active participation. A comparison is conducted between the 689 respondents to the survey and active non-respondents of the entire online social community (with a last log-in date of less than two months before the send-out of the survey), who had at least one contact. T-tests of profile visits, number of contacts and gender do not show significant differences. However, t-tests of age, tenure and active participation show significant differences ($p<.05$). Nevertheless, age and tenure reveal only minor differences. The mean age of the respondents is only 2.2 years higher than that of the non-respondents. In addition, non-respondents have a longer tenure (4.3 month longer) than respondents in the online community. In order to account for possible effects of age, gender and tenure, these variables are included in the analyses as covariates to correct for their possible effect when comparing the user groups of interest. Last, active participation is significantly higher for the users in the survey sample. However, the survey sample also includes a large number of users with no or few contributions (80 % of the users made less than 23 ac-

[46] Users in the sample have a self-reported number of friends in the online community of at least one. Further, the sample consists of users who indicated that they are still active.

tive contributions in the two months of interest), which eliminates the risk of having only heavy users in the survey sample.

Definition of Acquisition Channels. Two variables are used to describe the groups of users exposed to different communication channels. The ‘Marketing Channel’ differentiates between WOM and personal selling. The ‘WOM Channel’ differentiates between offline and online WOM. The users were first asked to identify their initial point of contact, which has drawn their attention to the online social community. The answers included WOM, personal selling, and a set of different other channels. Because of the focus on interpersonal communication channels and their relevance for user acquisition, only WOM and personal selling are considered as groups for the Marketing Channel. In fact, these two channels had by far the most answers (see Table 6). Users who answered that they were brought to the community through the recommendation of friends and acquaintances are assigned to WOM. Users who stated that they had contact to an employee at a leisure time event are assigned as personal selling acquired. In the analysis on the different Marketing Channels, only those users from personal selling are taken into account, who did not receive any subsequent WOM because WOM is assumed to have a stronger effect than personal selling and would thereby lead to mixed effects.

Independently from the question on the Marketing Channel, the users were asked to specify the number of WOM recommendations they received from other people online (through E-Mail or an online link) and offline (direct personal conversation). These answers are dichotomized in zero and one or more WOM receptions. They are then used to distinguish users by WOM Channel in users who received only offline WOM and users who received only online WOM. Since the objective of this study is to investigate the effectiveness of offline versus online WOM, users who either received no WOM or both, online and offline, WOM are not considered in the analysis. Because WOM in general is assumed to be more effective than other communication channels (e.g., Engel, Kegerreis, and Blackwell 1969; Herr, Kardes, and Kim 1991; Trusov, Bucklin, and Pauwels 2009; Villanueva, Yoo, and Hanssens 2008), all users who received online or offline WOM respectively are considered here. This also includes users who first heard about the online social community via other channels, but subsequently or simultaneously received referrals from other users. An overview of the sample characteristics and different user groups can be found in Table 6.

	% of users	Mean	Std.Dev.
Gender (female)	54 %	-	-
Age (in years)	-	23.79	9.23
Membership tenure (in months)	-	14.20	12.90
Internet usage per week (in hours)	-	4.25	3.96
Active Participation (log-transformed)	-	1.38	1.88
Passive Participation (log-transformed)	-	2.07	1.53
Marketing Channel *(through which users got to know the online community)*			
WOM	43 %	-	-
Personal Selling	45 %	-	-
Other Channels	12 %	-	-
Word-of-Mouth Channel *(referrals given and received)*			
WOM given offline	57 %	4.20	21.51
WOM given online	33 %	1.86	7.34
WOM received offline	61 %	1.97	3.64
WOM received online	32 %	0.82	2.35
WOM received offline only	37 %	-	-
WOM received online only	8 %	-	-
WOM received both online and offline	24 %	-	-
no WOM received	31 %	-	-

Std.Dev. = Standard Deviation; WOM = word-of-mouth
n=689

Table 6: Study 1 – Sample Characteristics

5.3.2 Measurement of Constructs

All attitudinal scales, frequency of use, WOM reception and provision, as well as the demographic variables are derived from the online survey. The attitudinal factors, i.e. satisfaction, identification, and WOM intention are measured using multi-item scales, from existing research, adapted to the context of this study. All other variables consist of single questions. The survey development and the operationalization of the scales and questions are explained in more detail in chapter 4.

Three behavioral measures are included in this study to investigate the impact of the communication channel from different perspectives. Frequency of site use (measured in days per week) is a self-reported measure taken from the survey. Recent studies have used frequency of use to observe participation in communities (e.g., Algesheimer, Dholakia, and Herrmann 2005; de Valck et al. 2007; Dholakia, Bagozzi, and Pearo 2004). In addition, two behavioral measures were used in form of objective data, collected by the community operator. On the basis of this data, two indices are constructed, which describe the active and passive participation on the platform. Active participation includes out-going activity, i.e. messages to other users, guestbook entries, sending of virtual gifts to other users, comments on articles and groups, as well as submitted ratings on photos, groups and profiles. Passive participation includes incoming activity from other users, including messages, guestbook entries, virtual gifts and profile ratings. Both measures are aggregated activity data from a period of two months which started two weeks before the survey was sent out and ended two weeks after the survey closed. In addition, behavioral variables also include the self-reported number of WOM provided to other people online and offline. The natural log-transformation (ln(x+1)) is used for the objective data and for self-reported data with open ended scales (active and passive participation, number of WOM provision) in order to control for extreme outliers and to reduce skewness (e.g., Wasko and Faraj 2005). Therefore, results must be carefully interpreted in terms of the absolute differences between user groups.

5.3.3 Measurement Model Evaluation of Self-Reported Latent Measures

Standard validity and reliability tests for the survey measures are conducted. All relevant statistics are provided in Table 7. The reliability of the scales was assessed using Cronbach's alpha with a minimum value of 0.7 being satisfactory (Nunnally 1978). To evaluate the unidimensionality of the proposed scales, exploratory factor analyses for each construct are conducted.[47] To test for internal consistency and discriminant validity, confirmatory analyses are accomplished. Asymptotically distribution-free (ADF) is chosen as an estimation method, since it shows more security in samples that might not present multivariate normality (Byrne 2010).[48] The tests are executed for the multi-item constructs used in this study.

Two further measures to evaluate internal consistency of constructs are used – composite reliability (CR) and average variance extracted (AVE). The composite reliability is a measure analogous to the Cronbach α coefficient (Fornell and Larcker 1981) with a suggested cut-off value of 0.6 (Bagozzi and Yi 1988). The average variance extracted estimates the amount of variance captured by a construct's measure relative to random measurement error (Fornell and Larcker 1981). Estimates of AVE above 0.5 are considered supportive of internal consis-

[47] Only one factor was extracted from each scale (criteria used: only factors extracted with eigenvalues higher than 1, factor loadings higher than 0.5, and a significant total explained variance).
[48] To pursue confirmatory factor analysis, the software package AMOS is used, which provides functionality to specify and test structural equation models.

tency (Bagozzi and Yi 1988). The results shown in Table 7 are satisfactory and therefore indicative for good internal consistency.

Self-Reported Measures from Survey	Number of Measures	Mean	Standard Deviation	Cronbach's Alpha	Lowest Item-to-total correlation	Composite Reliability	Average Variance Extracted
Satisfaction	3	4.57	1.20	.83	.66	.86	.67
Identification	3	3.27	1.64	.90	.79	.90	.76
Information Consumption Value	3	4.86	1.72	.87	.74	.88	.71
Social Interaction Value	3	3.49	1.86	.90	.73	.90	.75
WOM Intention	1	4.02	2.07	-	-	-	-
Frequency of Use (in days per week)	1	2.12	2.40	-	-	-	-

n=689; CFA: executed only for multidimensional measures using ADF (asymptotically distribution-free) estimation

Table 7: Study 1 – Descriptive, Reliability and Internal Consistency Statistics for Self-Reported Measures

	Correlations	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Satisfaction	1												
2	Identification	.355**	1											
3	Active Participation1)	.254**	.247**	1										
4	Passive Participation1)	.189**	.227**	.833**	1									
5	Frequency of Use	.329**	.293**	.702**	.635**	1								
6	WOM Intention	.593**	.404**	.256**	.184**	.333**	1							
7	WOM Provision Online1)	.178**	.161**	.058	.073	.166**	.268**	1						
8	WOM Provision Offline1)	.286**	.265**	.270**	.300**	.342**	.399**	.479**	1					
9	Information Cons. Value	.271**	.110**	-.034	-.040	.021	.368**	.126**	.159**	1				
10	Social Interaction Value	.311**	.477**	.433**	.393**	.441**	.380**	.254**	.339**	-.024	1			
11	Age	.036	.007	.047	-.031	-.025	.078*	.017	-.069	-.125**	.107**	1		
12	Gender	-.095*	-.152**	-.125**	-.013	-.118**	-.105**	-.030	-.085*	.037	-.161**	-.274**	1	
13	Tenure	.034	.070	.243**	.294**	.292**	.001	.078*	.268**	-.041	-.172**	-.040	-.032	1

** p<.05; ** p<.01; n=689*
1) natural log-transformed variables

Table 8: Study 1 – Correlations of Variables

Discriminant validity verifies that a construct is significantly distinct from other constructs that are not theoretically related to it. A CFA-model is used, which includes four latent constructs (satisfaction, identification, information consumption value, social interaction value) with three items each. Results show that the model fits the data well (Hu and Bentler 1999). The goodness-of-fit statistics for the model are as follows: $\chi^2(48)=70.90$, $p<.05$, CMIN/DF=1.477, NFI=.958, CFI=.986, TLI=.981, RMSEA=.026, SRMR=.034. Further, the correlations among the latent constructs are lower than 0.8 points (Bagozzi 1994; see Table 8). In addition, a test of discriminant validity was performed as suggested by Fornell and Larcker (1981). Discriminant validity is achieved if the AVE by the underlying construct is larger than the shared variance (i.e. squared correlations) with other latent constructs. This condition is satisfied for all of the cases. In sum, internal consistency and discriminant validity are satisfactory and permitted to include these constructs in the hypotheses tests.

5.4 Results of Main Analysis

To test the hypotheses, two MANCOVAs are first run on all attitudinal and behavioral dependent variables – one with Marketing Channel (WOM vs. personal selling) and one with WOM Channel (offline vs. online WOM) as independent variables. It is controlled for the effects of age, gender and tenure[49] as covariates. Both MANCOVA results show that Marketing Channel (Wilks lambda = .775, F=11.12, p<.01) and WOM Channel (Wilks lambda = .905, F=3.10, p<.01) had significant main effects on the set of dependent variables, suggesting that there is an overall difference between users coming from diverse channels. To further understand the effect of the different Marketing and WOM Channels, separate univariate ANCOVAs on each of the dependent variables are conducted, including the same covariates (age, gender, and tenure). In order to answer the specified hypotheses, only two channels are compared at a time, including an examination of the parameter estimates of the different channels, which is similar to the approach of Lacey, Suh, and Morgan (2007). Figures 10-12 visualize the mean values of all dependent variables for the various communication channels. Table 9 summarizes the results for the marketing communication channels; Table 10 includes the results for the WOM Channels.

[49] Due to missing values in the objective data, tenure is included in form of the self-reported variable as a control variable.

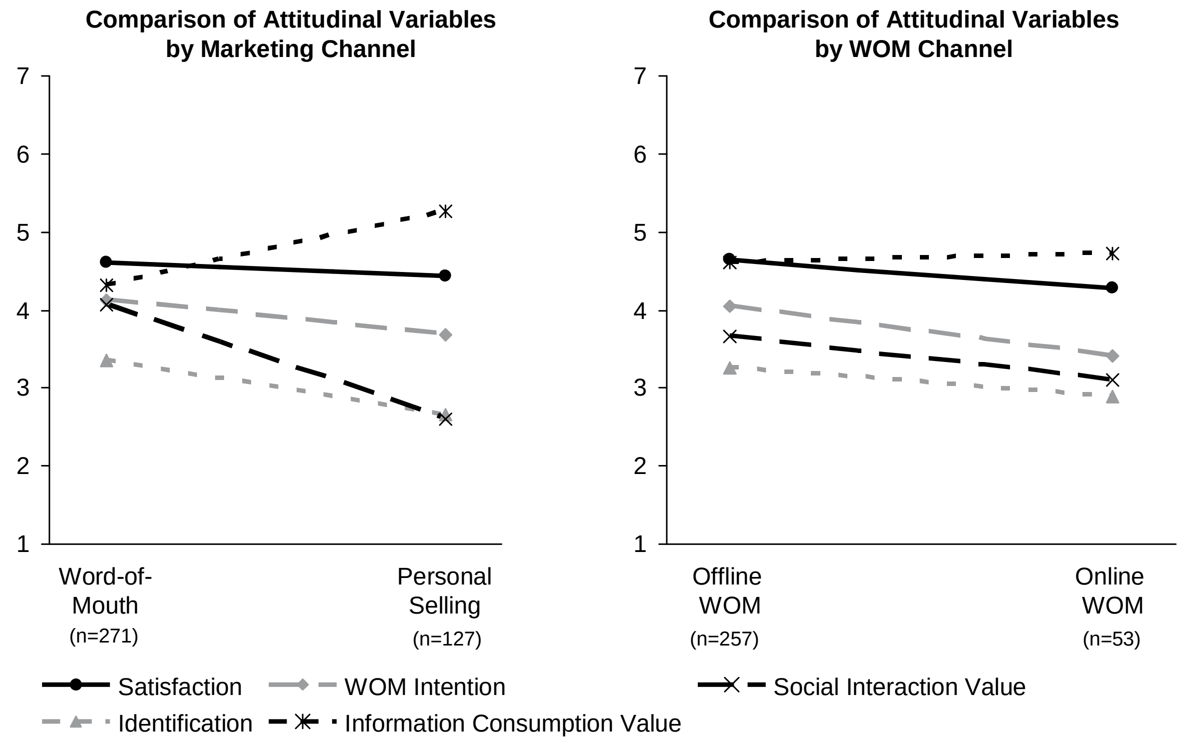


Figure 10: Study 1 – Comparison of Attitudinal Variables by Acquisition Channel

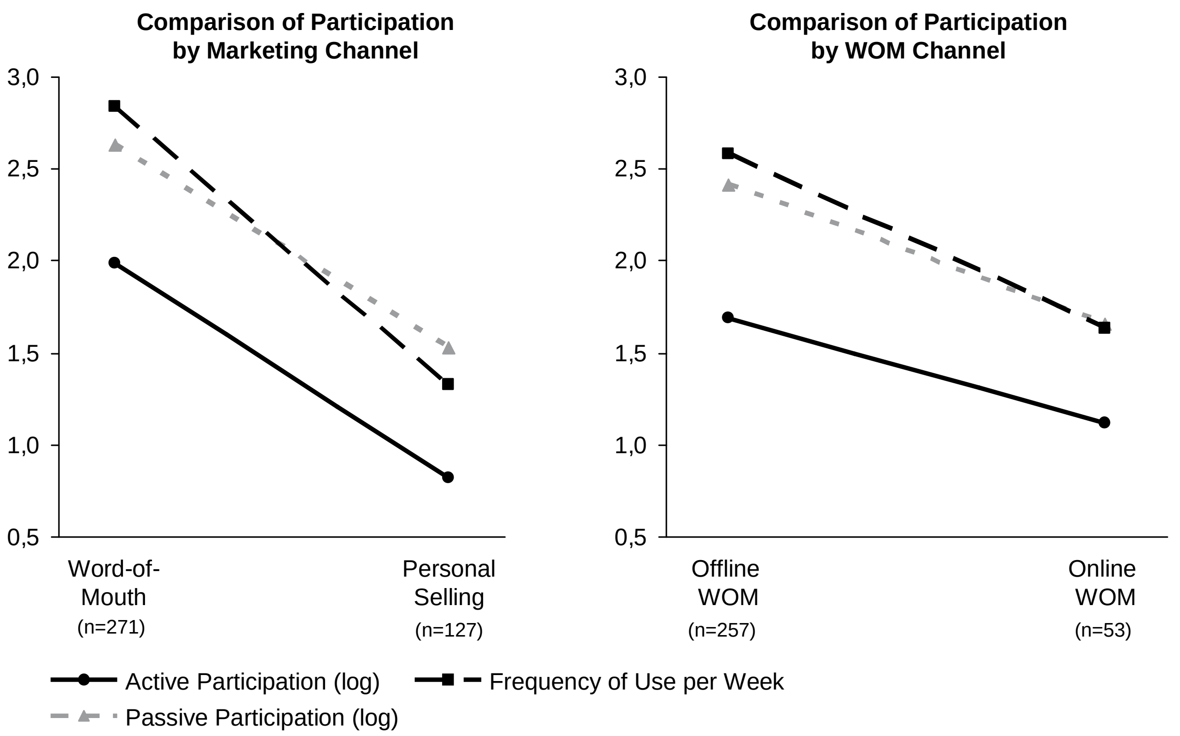


Figure 11: Study 1 – Comparison of User Participation by Acquisition Channel

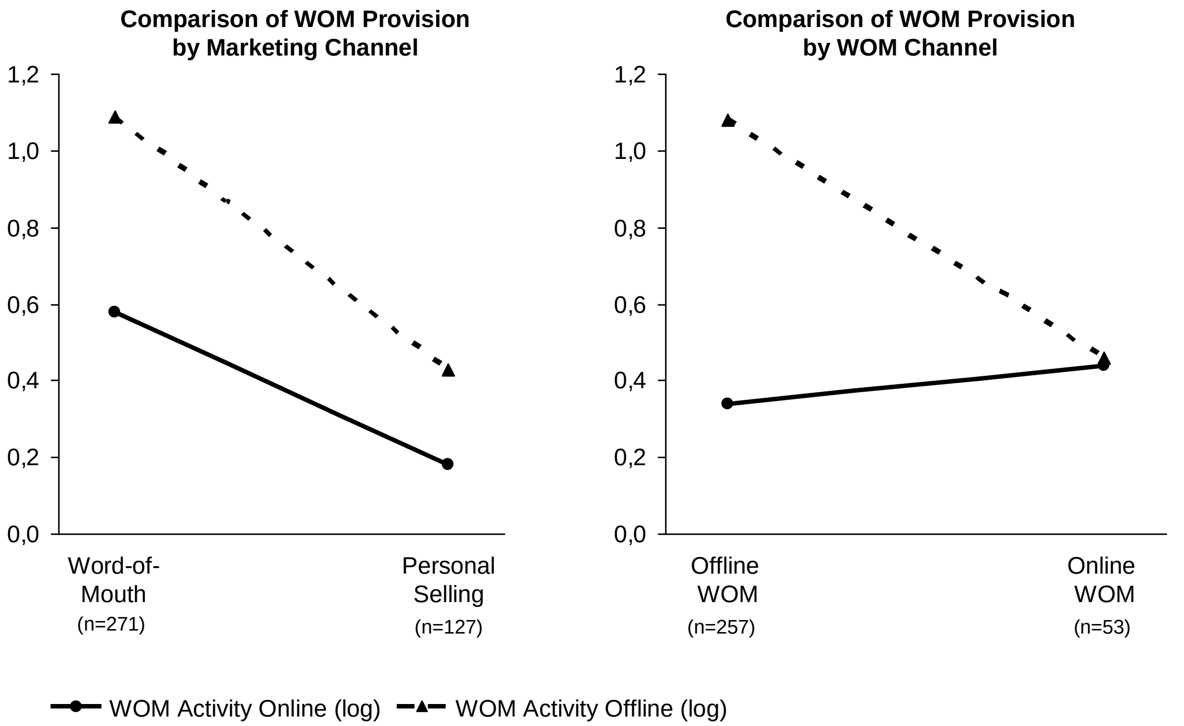


Figure 12: Study 1 – Comparison of WOM Provision by Acquisition Channel

5.4.1 Marketing Channel: Word-of-Mouth vs. Personal Selling

For the comparison of Marketing Channels, two groups of users are specified: users coming from recommendations of friends to the online social community and users coming via the personal selling channel to the platform without receiving any additional referral. Separate ANCOVAs on each dependent variable show overall significant effects including the Marketing Channel type and the covariates on all dependent variables ($p<.10$).

Parameter estimates are used for subgroup comparison. The personal selling group is used as the reference group (having a value of zero). The comparison of WOM-referred users (WOM) and users coming solely from the personal selling channel (PS) shows a significant difference for all dependent variables other than satisfaction. Compared to the group of users acquired by personal selling, the analyses suggest that users referred by a friend or acquaintance identify more with the community, participate more on the platform in terms of active participation, passive participation, and frequency of use, have higher WOM intentions, and provided more WOM to other people online and offline. Further WOM-referred users value information consumption lower but social interaction higher than users acquired by personal selling. The exact results are illustrated in Table 9. Altogether, these results support most of the above outlined hypotheses. Only hypothesis H1a is not supported.

	Dependent Variables									
	Satisfaction		Identification		Active Participation		Passive Participation		Frequency of Use	
	B	p	B	p	B	p	B	p	B	p
Constant	4.439	0.000 ***	3.141	0.000 ***	1.056	0.002 ***	1.700	0.000 ***	2.051	0.000 ***
Covariate										
Age	0.005	0.515	-0.009	0.359	-0.009	0.384	-0.015	0.085 *	-0.035	0.009 ***
Tenure	0.004	0.350	0.006	0.334	0.024	0.001 ***	0.019	0.001 ***	0.050	0.000 ***
Gender	-0.263	0.036 **	-0.363	0.035 **	-0.459	0.015 **	-0.041	0.788	-0.709	0.003 ***
Marketing Channel	0.112	0.406	0.656	0.000 ***	0.954	0.000 ***	0.954	0.000 ***	1.103	0.000 ***
Hypothesis Test	H1a	n.s.	H2a	supp.	H3a	supp.	H4a	supp.	H5a	supp.

	Dependent Variables									
	WOM Intention		WOM Provision Online		WOM Provision Offline		Social Interaction Value		Information Cons Value	
	B	p	B	p	B	p	B	p	B	p
Constant	3.725	0.000 ***	0.091	0.527	0.611	0.001 ***	2.659	0.000 ***	5.579	0.000 ***
Covariate										
Age	0.012	0.311	0.004	0.331	-0.009	0.096 *	0.004	0.674	-0.011	0.245
Tenure	0.000	0.999	0.002	0.452	0.018	0.000 ***	0.013	0.058 *	-0.002	0.715
Gender	-0.519	0.013 **	-0.063	0.427	-0.275	0.004 ***	-0.496	0.008 ***	-0.031	0.860
Marketing Channel	0.395	0.078 *	0.375	0.000 ***	0.511	0.000 ***	1.328	0.000 ***	-0.925	0.000 ***
Hypothesis Test	H6a	supp.	H7a	supp.	H8a	supp.	H9a	supp.	H10a	supp.

Marketing Channel = Word-of-mouth (=1) vs. Personal Selling (without received WOM; =0);
n=398 (WOM=271; Personal Selling=127)
** p<.10; ** p<.05; *** p<.01*
Hypotheses Test: supp. = supported; n.s. = not supported

Table 9: Study 1 – Separate ANCOVA Results for Marketing Channel

5.4.2 WOM Channel: Offline vs. Online Word-of-Mouth

Further ANCOVAs are conducted in order to test the difference between the WOM Channels. In the ANCOVAs all users who received only offline referrals (offline WOM) and all users who received only online referrals (online WOM) are included. The overall ANCOVAs, including WOM Channel type and the covariates, are significant for all dependent variables observed (p<.10), except for online WOM provision and information consumption value.

More important for testing the hypotheses is the comparison of offline WOM and online WOM channels. Therefore, parameter estimates of offline WOM, with online WOM as the reference group (having a value of zero), are used. Users who received WOM via offline communication are more satisfied, willing to give WOM in the future, gave more offline

WOM to other people, and participate more with respect to active participation, passive participation, and frequency of use. For identification and online WOM provision no significant effects are observed. On the motivational dimensions offline and online WOM receivers do not differ significantly in their information consumption values, but they show significant differences on their social interaction value. These results support hypotheses H1b, H3b, H4b, H5b, H6b, H8b, H9b and H10b, while H2b, H7b are not supported (see Table 10 for the results).

	Dependent Variables														
	Satisfaction			Identification			Active Participation			Passive Participation			Frequency of Use		
	Beta	p		Beta	p		Beta	p		Beta	p		Beta	p	
Constant	4.538	0.000	***	3.644	0.000	***	0.298	0.544		1.035	0.008	***	1.416	0.023	**
Covariate															
Age	0.001	0.876		-0.006	0.618		0.029	0.036	**	0.012	0.263		-0.008	0.647	
Tenure	-0.003	0.541		-0.004	0.614		0.031	0.000	***	0.032	0.000	***	0.056	0.000	***
Gender	-0.427	0.002	***	-0.641	0.001	***	-0.595	0.007	***	-0.242	0.166		-0.671	0.017	**
WOM Channel	0.351	0.053	*	0.237	0.338		0.574	0.046	**	0.729	0.001	***	0.812	0.026	**
Hypotheses Test	H1b	supp.		H2b	n.s.		H3b	supp.		H4b	supp.		H5b	supp.	

	Dependent Variables														
	WOM Intention			WOM Provision Online			WOM Provision Offline			Social Interaction Value			Information Cons Value		
	Beta	p		Beta	p		Beta	p		Beta	p		Beta	p	
Constant	3.708	0.000	***	0.444	0.020	**	0.670	0.007	***	3.085	0.000	***	5.690	0.000	***
Covariate															
Age	0.008	0.583		0.002	0.699		-0.014	0.049	**	0.009	0.517		-0.033	0.011	***
Tenure	-0.010	0.261		-0.001	0.716		0.021	0.000	***	0.016	0.050	*	-0.005	0.480	
Gender	-0.607	0.012	**	-0.067	0.431		-0.270	0.017	**	-0.727	0.001	***	-0.080	0.701	
WOM Channel	0.640	0.040	**	-0.092	0.408		0.532	0.000	***	0.512	0.069	*	-0.219	0.421	
Hypotheses Test	H6b	supp.		H7b	n.s.		H8b	supp.		H9b	supp.		H10b	supp.	

WOM Channel = Offline (=1) vs. Online WOM (=0)
n=310 (offline WOM=257; online WOM=53)
** p<.10; ** p<.05; *** p<.01*
Hypotheses Test: supp. = supported; n.s. = not supported

Table 10: Study 1 – Separate ANCOVA Results for WOM Channel

5.4.3 Step-Down Analysis

It has to be noted that most of the dependent variables are to some degree interrelated with motivation (social interaction needs and/or information consumption needs). Variation in a particular variable may be due to the dependence of that variable on the motivational variables. “Step-down analyses provide useful information because they test whether variation in a certain dependent variable is due to a direct association with the manipulation or to its dependence on other dependent variables" (Bagozzi and Yi 1989, p. 274). Because of the theoretical a priori ordering of the dependent variables in the form that attitudes and behavior follow motivational factors (see for example Dholakia, Bagozzi, and Pearo 2004; Lampe et al. 2010; Matthwick, Wiertz, and de Ruyter 2008) and the given correlations between motivation and most of the remaining dependent variables, step-down analyses are conducted. Thereby, the objective is to examine a possible mediation effect of motivation (social interaction value and information consumption value) between communication channel and the outcome variables in order to test whether the communication channel still has a direct impact on the other dependent variables. By examining dependent variables in a predetermined order, the step-down analysis gauges the distinct contribution of each variable to the between-group variance as the variable is added to the dependent variable set (Bagozzi and Yi 1989; Yi 1993). Therefore, information consumption value and social interaction value are used as additional covariates in the analyses.

	Dependent Variables							
	Satisfaction		Identification		Active Participation		Passive Participation	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	2.800	0.000 ***	1.201	0.002 ***	-0.255	0.557	0.777	0.032 **
Covariate								
Age	0.006	0.333	-0.008	0.330	-0.010	0.287	-0.016	0.051 *
Tenure	0.003	0.543	0.002	0.755	0.019	0.003 ***	0.016	0.004 ***
Gender	-0.172	0.139	-0.177	0.254	-0.263	0.134	0.104	0.476
Social Interaction Value	0.171	0.000 ***	0.363	0.000 ***	0.393	0.000 ***	0.292	0.000 ***
Information Cons. Value	0.212	0.000 ***	0.175	0.000 ***	0.048	0.339	0.026	0.525
Marketing Channel	0.082	0.544	0.335	0.062 *	0.477	0.019 **	0.591	0.001 ***

	Dependent Variables							
	Frequency of Use		WOM Intention		WOM Provision Online		WOM Provision Offline	
	Beta	p	Beta	p	Beta	p	Beta	p
Constant	0.100	0.855	0.346	0.441	-0.360	0.065 *	-0.290	0.201
Covariate								
Age	-0.037	0.003 ***	0.015	0.136	0.005	0.294	-0.009	0.091 *
Tenure	0.043	0.000 ***	-0.004	0.554	0.001	0.613	0.016	0.000 ***
Gender	-0.447	0.044 **	-0.318	0.080 *	-0.029	0.713	-0.198	0.030 **
Social Interaction Value	0.522	0.000 ***	0.378	0.000 ***	0.065	0.002 ***	0.149	0.000 ***
Information Cons. Value	0.101	0.111	0.425	0.000 ***	0.050	0.027 **	0.091	0.001 ***
Marketing Channel	0.504	0.050 *	0.287	0.172	0.334	0.000 ***	0.397	0.000 ***

Marketing Channel = Word-of-mouth (=1) vs. Personal Selling (without received WOM; =0);
n=398 (WOM=271; Personal Selling=127)
** p<.10; ** p<.05; *** p<.01*
Hypotheses Test: supp. = supported; n.s. = not supported

Table 11: Study 1 – Step-Down Analysis Results for Marketing Channel

In Table 11, the main effects and group comparisons are included for the Marketing Channels of interest. Step-down analysis reveals that WOM-referred users still differ significantly in comparison to personal selling acquired users regarding their identification with the community, active participation, passive participation, frequency of use, and online and offline WOM provision when motivational factors are treated as covariates. However, the effect of satisfaction is still insignificant. Further, the effect of WOM intentions becomes insignificant. This suggests that the variation in this variable is due to its dependence on motivational factors rather than due to the direct influence of the Marketing Channel. Therefore, the effect of Marketing Channel on WOM intention does not hold when information consumption value and social interaction value together are considered as mediating variables.

	Dependent Variables											
	Satisfaction			Identification			Active Participation			Passive Participation		
	Beta	p		Beta	p		Beta	p		Beta	p	
Constant	2.792	0.000	***	1.597	0.001	***	-1.044	0.054	*	0.127	0.769	
Covariate												
Age	0.005	0.546		-0.005	0.603		0.023	0.056	*	0.007	0.466	
Tenure	-0.006	0.173		-0.010	0.118		0.022	0.002	***	0.026	0.000	***
Gender	-0.227	0.070	*	-0.321	0.055	*	-0.224	0.250		0.038	0.806	
Social Interaction Value	0.256	0.000	***	0.427	0.000	***	0.515	0.000	***	0.392	0.000	***
Information Cons. Value	0.168	0.000	***	0.128	0.005	***	-0.044	0.406		-0.053	0.209	
WOM Channel	0.256	0.112		0.046	0.831		0.301	0.230		0.517	0.010	**

	Dependent Variables											
	Frequency of Use			WOM Intention			WOM Provision Online			WOM Provision Offline		
	Beta	p		Beta	p		Beta	p		Beta	p	
Constant	-0.561	0.415		-0.316	0.568		0.059	0.805		-0.235	0.440	
Covariate												
Age	-0.014	0.375		0.019	0.120		0.002	0.649		-0.012	0.079	*
Tenure	0.046	0.000	***	-0.015	0.043	**	-0.002	0.503		0.019	0.000	***
Gender	-0.204	0.410		-0.232	0.243		-0.014	0.872		-0.170	0.120	
Social Interaction Value	0.643	0.000	***	0.466	0.000	***	0.070	0.002	***	0.127	0.000	***
Information Cons. Value	-0.001	0.990		0.455	0.000	***	0.029	0.203		0.090	0.002	***
WOM Channel	0.482	0.130		0.501	0.051	*	-0.122	0.270		0.487	0.001	***

WOM Channel = Offline (=1) vs. Online WOM (=0)
n=310 (offline WOM=257; online WOM=53)
** p<.10; ** p<.05; *** p<.01*
Hypotheses Test: supp. = supported; n.s. = not supported

Table 12: Study 1 – Step-Down Analysis Results for WOM Channel

Table 12 shows the results of the step-down analysis for the WOM Channels. Again, information consumption value and social interaction value were treated as additional covariates. The overall impact of WOM Channel on passive participation, WOM intention and offline WOM provision is still significant when motivational factors are treated as covariates. The influence of WOM Channels on satisfaction, active participation, and frequency of use become insignificant. Therefore, those variables are dependent on the motivational factors, predominantly on social interaction value. The direct effect of WOM Channels on those dependent variables does not hold true. Thus, online and offline WOM-referred users do not significantly differ in their effect on those variables when controlling for motivations. Here, the mediating role of motivation with regards to some of the dependent variables can also be observed.

In order to check for the robustness of these results, alternative methods are used. All main analyses were repeated using separate regression analyses with heteroskedasticity-consistent standard errors, including the acquisition channel and the respective covariates as the independent variable on each of the outcome factors as the dependent variable. The results do not show any substantial differences to the ones described above. This confirms the findings. Additionally, the non-parametric Kruskal-Wallis test and bootstrapped regressions are used, which address potential issues of non-normality. In the Kruskal-Wallis test, the direct effect of the acquisition channel on each variable is estimated. The bootstrapped regressions used the same data as in the models described above in the main analyses. Again, all results from the main analyses are supported. On this way, the robustness of the results was checked for potential deviations from regression assumptions, resulting in supportive results for the analyses described in this section (see Appendix 4). Therefore, in a next step the potential mediation effects of user motivations can be analyzed.

5.4.4 Mediation Analysis

5.4.4.1 Mediation Effects of Motivations – Methodology

Additional analyses are used to assess whether the effects of the acquisition channels are mediated through user motivation because the results of the step-down analyses indicated potential mediation effects. Thereby, the objective is to investigate the mediator role of motivation and test whether the direct effects of the acquisition channels on user attitudes and behavior still hold.[50] Because motivation is a central trigger for behavior, which has been demonstrated in recent research by its influence on attitudes and behavior in online communities (e.g., Dholakia, Bagozzi, and Pearo 2004; Lampe et al. 2010; Matthwick, Wiertz, and de Ruyter 2008), the users' value perceptions are included as potential mediators in the research model. Different mediation models are tested which consist of the type of acquisition channel as the independent variable, the two motivation variables as mediators (social interaction value and information consumption value), and the dependent variables satisfaction, identification, active participation, passive participation, frequency of use, online and offline WOM provision, and WOM intention. Figure 13 illustrates the different mediation models.

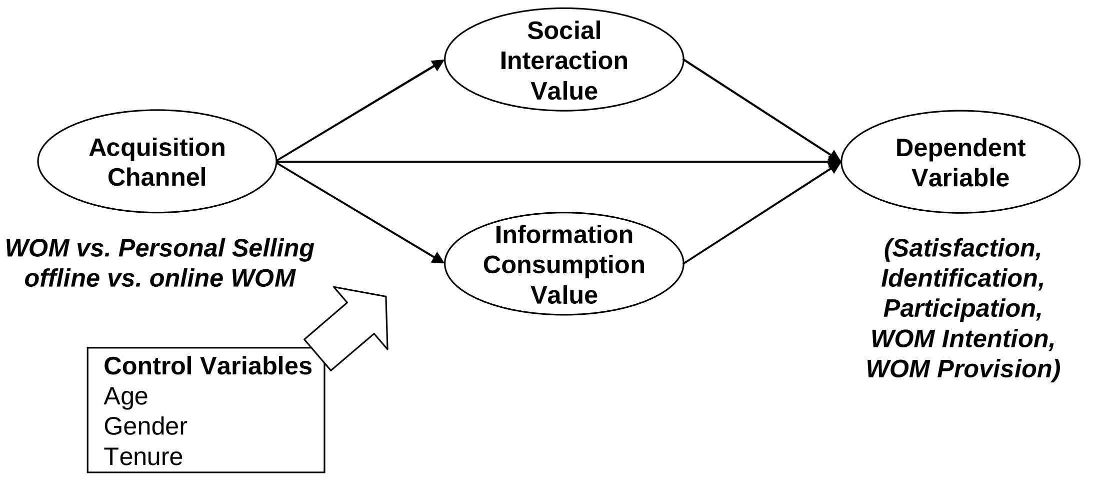


Figure 13: Study 1 – Mediation Model of Acquisition Channels

The objective of this study is to analyze the direct effects of the acquisition channels and how they change when mediation is considered. Preacher and Hayes' (2008) procedure to perform regression-based multiple mediation analyses is employed in order to test whether motivational factors mediate the relationship between the mode of user acquisition and attitudes and behavior.[51] This method has the advantage of being independent from distributional assumptions regarding the parameter estimates for the indirect path. The mediation is tested by an evaluation of statistical significance levels of the indirect effect from the independent

[50] This study focuses on the mediation between the acquisition channels and important outcomes for online social communities. It is acknowledged that there are also interdependencies between the outcome variables in focus. For example, the relationship between satisfaction, identification and active participation will be investigated in chapter 6 of this thesis. In this study however, the focus is set on understanding how the channels differ in their behavioral and attitudinal outcomes of the users and how motivation mediates.
[51] The most commonly utilized method for testing for mediation is one suggested by Baron and Kenny (1986). However, researchers have recently suggested that a direct test of the mediating effect is superior to the Baron and Kenny method (Preacher and Hayes 2004).

variable (X) via the mediator (M) on the dependent variable (Y). Thereby, mediation is indicated when the effects of X on M and of M on Y are both significant (Preacher and Hayes 2008). The direct effect of X on Y without consideration of mediators must not necessarily be significant for mediation to occur (Preacher and Hayes 2008). The Preacher and Hayes method makes use of the bootstrapping procedure. Thereby, the bootstrapping procedure repeatedly takes samples from the original sample, which are used to derive estimates from this re-sampled data. The distributions of the estimates from all re-sampled data serve as empirical, nonparametric approximation of the sampling distribution of the indirect effect and are used to obtain confidence intervals for the indirect effect (Preacher and Hayes 2008). If the bootstrapped confidence interval (CI)[52] of the indirect effect through the mediator does not include zero, the mediating effect is significant (Preacher and Hayes 2008).

For each path from X through M on Y, three regressions are estimated here (X→M, M→Y, and X→Y with inclusion of mediators). In multiple mediation regression the mediation via one mediator is controlled for the other mediator, which is an advantage over several single mediator regressions (Preacher and Hayes 2008). The following bootstrap analyses use 5,000 bootstrap samples.

5.4.4.2 Direct Effects of Motivations on Attitudes and Behavior

All effects of social interaction value on satisfaction, identification, WOM intention, WOM provision and participation are significant and positive. Therefore, a higher perceived social value is associated with stronger attitudes and behavior towards the online social community. In the same manner, higher information consumption value leads to higher satisfaction, identification, WOM provision and WOM intention. However, it does not significantly influence the users' participation on the platform. These results are mainly consistent in both subsamples: the one comparing the Marketing Channels and the one comparing the WOM Channels. The only exception is that for the analysis sample of WOM Channels, the effect on online WOM provision is insignificant. See the section "Direct Effects – M → Y" in Table 13 for a summary of these results.

[52] In the analysis and results discussion of this section it is referred to the bias-corrected and accelerated confidence intervals when indicating the CI values. Upper and lower values of the confidence interval are obtained from a 90% CI.

			Acquisition Channel			
			Marketing Channel[1)]		WOM Channel[2)]	
			Coeff.	p	Coeff.	p
Direct Effects						
X --> M						
Acquisition Channel	-->	Social Interaction	1,3279	,0000 ***	,5123	,0690 *
Acquisition Channel	-->	Information Consumption	-,9253	,0000 ***	-,2188	,4212
M --> Y						
Social Interaction	-->	Satisfaction	,1710	,0000 ***	,2559	,0000 ***
Social Interaction	-->	Identification	,3633	,0000 ***	,4270	,0000 ***
Social Interaction	-->	Active Participation	,3927	,0000 ***	,5153	,0000 ***
Social Interaction	-->	Passive Participation	,2915	,0000 ***	,3920	,0000 ***
Social Interaction	-->	Frequency of Use	,5218	,0000 ***	,6426	,0000 ***
Social Interaction	-->	WOM Intention	,3780	,0000 ***	,4656	,0000 ***
Social Interaction	-->	WOM Provision Online	,0654	,0021 ***	,0704	,0018 ***
Social Interaction	-->	WOM Provision Offline	,1487	,0000 ***	,1268	,0000 ***
Information Consumption	-->	Satisfaction	,2124	,0000 ***	,1682	,0000 ***
Information Consumption	-->	Identification	,1746	,0001 ***	,1284	,0046 ***
Information Consumption	-->	Active Participation	,0478	,3387	-,0436	,4060
Information Consumption	-->	Passive Participation	,0265	,5249	-,0529	,2090
Information Consumption	-->	Frequency of Use	,1010	,1111	-,0009	,9895
Information Consumption	-->	WOM Intention	,4255	,0000 ***	,4547	,0000 ***
Information Consumption	-->	WOM Provision Online	,0497	,0269 **	,0295	,2028
Information Consumption	-->	WOM Provision Offline	,0906	,0006 ***	,0902	,0024 **
X --> Y (mediation model)						
Acquisition Channel	-->	Satisfaction	,0815	,5444	,2562	,1121
Acquisition Channel	-->	Identification	,3355	,0625 *	,0459	,8308
Acquisition Channel	-->	Active Participation	,4766	,0192 **	,3006	,2302
Acquisition Channel	-->	Passive Participation	,5912	,0005 ***	,5166	,0104 **
Acquisition Channel	-->	Frequency of Use	,5038	,0503 *	,4824	,1304
Acquisition Channel	-->	WOM Intention	,2868	,1723	,5013	,0508 *
Acquisition Channel	-->	WOM Provision Online	,3342	,0003 ***	-,1219	,2700
Acquisition Channel	-->	WOM Provision Offline	,3971	,0002 ***	,4872	,0006 ***

Indirect Effects			Bca 90% Confidence Intervals					
X --> M --> Y			Coeff.	Lower CI	Upper CI	Coeff.	Lower CI	Upper CI
Via Social Interaction								
Acquisition Channel	-->	Satisfaction	,2270	,1500	,3288	,1311	,0234	,2531
Acquisition Channel	-->	Identification	,4824	,3499	,6436	,2188	,0309	,4153
Acquisition Channel	-->	Active Participation	,5214	,3650	,7079	,2640	,0414	,5127
Acquisition Channel	-->	Passive Participation	,3871	,2670	,5345	,2008	,0328	,3821
Acquisition Channel	-->	Frequency of Use	,6929	,4888	,9240	,3293	,0438	,6263
Acquisition Channel	-->	WOM Intention	,5020	,3584	,6854	,2385	,0383	,4595
Acquisition Channel	-->	WOM Provision Online	,0868	,0349	,1469	,0361	,0069	,0843
Acquisition Channel	-->	WOM Provision Offline	,1975	,1339	,2783	,0650	,0120	,1377
Via Informtion Consumption								
Acquisition Channel	-->	Satisfaction	-,1965	-,2968	-,1248	-,0368	-,1144	,0369
Acquisition Channel	-->	Identification	-,1616	-,2703	-,0790	-,0281	-,1039	,0190
Acquisition Channel	-->	Active Participation	-,0443	-,1281	,0262	,0095	-,0070	,0663
Acquisition Channel	-->	Passive Participation	-,0245	-,0968	,0399	,0116	-,0052	,0619
Acquisition Channel	-->	Frequency of Use	-,0934	-,2048	-,0023	,0002	-,0348	,0379
Acquisition Channel	-->	WOM Intention	-,3937	-,5651	-,2565	-,0995	-,3046	,0983
Acquisition Channel	-->	WOM Provision Online	-,0459	-,0915	-,0134	-,0064	-,0372	,0025
Acquisition Channel	-->	WOM Provision Offline	-,0839	-,1396	-,0410	-,0197	-,0693	,0146

1) WOM vs. PS (WOM=1; PS=0); n=398
2) Offline vs. Online WOM (Offline WOM=1; Online WOM=0); n=310
**** p<.01; ** p<.05; * p<.10; 5,000 Bootstrap Samples*
X = Independent Variable; M = Mediator; Y = Dependent Variable
CI = Confidence Interval; Bca = Bias Corrected and Accelerated
Covariates: Age, Gender, Tenure

Table 13: Study 1 – Results of Direct and Indirect Effects in all Mediation Models

5.4.4.3 Mediation between Marketing Channel and User Attitudes and Behavior

Satisfaction. As already presented, the direct effect of the Marketing Channel on satisfaction is not significant. Nevertheless, significant direct effects in a model without consideration of mediators are not necessary for mediation to occur (Preacher and Hayes 2008). The analysis demonstrates that the channel effect is mediated through motivation because the paths' from the Marketing Channel on both social interaction value and information consumption value are significant and the effects of the two motivations on satisfaction are also significant. The analysis yields a mean indirect effect for mediation through social interaction value of .23 (90% CI: .15; .33) and for mediation via information consumption value of -.20 (90% CI: -.30; -.12), demonstrating the significance of both mediations. Because WOM-referred users show higher social interaction value, the indirect effect through this variable is positive. On the other hand, the indirect effect through information consumption value is negative due to the lower value WOM-referred users perceive from information compared to personal selling acquired users.

Identification. The direct effect of the Marketing Channel on identification is still significant after including the mediation paths though the effect decreased. This is still consistent with hypothesis H2a. When testing for mediation of the user's motivation, the analysis reveals significant effects of the Marketing Channel on both motivations; and from both motivations on identification, which also results in significant indirect effects. The estimates are .48 (90% CI: .35; .64) for social interaction value and -.16 (90% CI: -.27; -.08) for information consumption value. Both bias-corrected and accelerated (BCa) CIs do not include 0, thus showing significant indirect effects, and thus a partial mediation on identification.

Word-of-Mouth. The Marketing Channel does not significantly influence the users' WOM intention directly but has a significant indirect effect through both motivations (90% CI: .36; .69 and -.57; -.26). Therefore, higher WOM intention is affected by higher perceptions of social and informational value of the platform. The indirect effect is higher through social interaction, indicating stronger influence of the Marketing Channel via social interaction value on WOM intention. Further, both online and offline WOM provision are significantly affected by the Marketing Channel, directly and indirectly via user motivations. Again the effect size is higher (larger point estimate) for the mediation through social interaction value than through information consumption value.

Participation. There is a significant difference between WOM-referred users and users coming from personal selling in their participation behavior, in terms of active participation, passive participation, and frequency of use. The effects of the Marketing Channel are still significant after accounting for motivational mediators. However, the effects decrease through the inclusion of mediators. Social interaction value is found to partially mediate these effects. In contrast, information consumption value does not impact participation significantly. It does not take a mediator role on active and passive participation. However, the indirect effect of the Marketing Channel on frequency of use is slightly significant. Here, social interaction value is more important and the analysis reveals strong direct and indirect effects of the Marketing Channel on participation.

Overall, social interaction need partially or fully mediates between Marketing Channel and all dependent variables, thus supporting hypothesis H11a. Information consumption needs mediate all effects, except those on active and passive participation. Therefore, hypothesis H11b is only partly confirmed.

5.4.4.4 Mediation between WOM Channel and User Attitudes and Behavior

Because of the insignificant relationship between the WOM Channel and information consumption value, motivation driven by information consumption value does not mediate any effect of the WOM Channel on the dependent variables. Therefore, potential mediating effects only occur via social interaction value. Because it was hypothesized that there is no mediation effect of information consumption value, hypothesis H12b is supported.

Satisfaction. The analysis shows a significant indirect effect of the WOM Channel on satisfaction, mediated by social interaction value (indirect effect of .13). As the direct effect is not significant in any model, the significant effects of WOM Channel on social interaction value and from the latter on satisfaction lead to a full mediation.

Identification. In the same way as for satisfaction, the analysis supports the mediation of the WOM Channel through social interaction value on identification. Thereby, users’ referred through offline WOM are associated with higher levels of social interaction value, which in turn lead to higher identification with the community. The indirect effect of the WOM Channel is .22.

Word-of-Mouth. A significant indirect effect of the WOM Channel of .24 (90% CI: .04; .46) is also found for the effect on WOM intention. However, there is also a direct effect of the WOM Channel on WOM intention. This effect is lowered in the mediation model compared to the direct effects model, but still significant, which indicates a partial mediation in this case. The comparison of online- and offline-referred users shows that the WOM Channel does not have a significant direct impact on online WOM provision, but on offline WOM provision. However, social interaction value mediates between the WOM Channel and both WOM activities (online and offline), which is indicated by the significant indirect effects of the WOM Channel on WOM provision of .04 (online) and .07 (offline).

Participation. The direct effects of the WOM Channel on active participation and frequency of use are not significant in the mediation model but there exist significant indirect effects of .26 and .33 respectively, which indicate mediation by social interaction value. The direct effect of passive participation remains significant in the mediation model. Nevertheless, its effect is also mediated by social interaction value (indirect effect: .20; 90% CI: .04; .38). Users who were referred offline show a higher participation on the platform, which is fully mediated for active participation and frequency of use and partly mediated for passive participation by their need to socially interact with other users of the community.

Overall, social interaction value mediates the effect of WOM Channel on all dependent variables. Therefore, hypothesis H12a is confirmed.

5.4.4.5 Direct Effects of the Interpersonal Acquisition Channel with Mediation

Despite the significant mediating role of user motivations between acquisition channels and the dependent variables, some main effects of the acquisition channels remain significant: the effects of the marketing channel on identification, active participation, passive participation, frequency of use, and online and offline WOM provision. This indicates that users coming to the platform through referrals are more active on the platform and show more positive attitudes towards the community. This underlines the value of WOM as an acquisition channel. Further, with regards to online and offline WOM, only passive participation, WOM intention and offline WOM provision are still significant when including the mediators in the model. This suggests that offline WOM referred users give more WOM than online-referred users and that they are more involved in the platform by receiving more activity from other members. The effects on active participation, frequency of use and satisfaction are mediated through their motivation for social interaction.

5.4.4.6 Verification of Mediation Model

In order to test for the robustness of the results from the mediation study, all mediation analyses are jointly estimated in one model using structural equation modeling (SEM).

AMOS is used as the statistical software to execute the analysis. Two models are specified: one for the Marketing Channel as the independent variable and one for the WOM Channel. Besides the acquisition channel variable, the models include social interaction value and information consumption value as the mediating variables, and satisfaction, identification, active participation, WOM intention, and online and offline WOM provision as the dependent variables.[53] The model includes all paths for mediation as well as the direct paths of the acquisition channel on the dependent variables, as suggested by Preacher and Hayes (2008). The results of the specified SEM models are similar to the ones from the multiple-mediation models presented above. The overall goodness-of-fit statistics for the model, using maximum likelihood estimation with 5,000 bootstrap samples, is satisfactory for both the Marketing Channel Model with $\chi^2(104)=337.83$, $p<.05$, CMIN/DF=3.248, NFI=.908, CFI=.934, TLI=.913, RMSEA=.075, SRMR=.070 and the WOM Channel Model with $\chi^2(104)=230.39$, $p<.05$, CMIN/DF=2.215, NFI=.915, CFI=.951, TLI=.936, RMSEA=.063, SRMR=.056.[54]

The significance levels of all effects in the model with the Marketing Channel as the independent variable are similar to the significance levels from the multiple mediation regressions using the procedure of Preacher and Hayes (2008). The only difference is that the direct effect of the Marketing Channel on identification is not significant in the SEM. When using the WOM Channel as the independent variable, all significance levels are similar. Appendix 5 provides an overview of the results from the SEM analyses.

The mediated SEM was also tested against a direct effects SEM as the alternative model, which included the acquisition channel as the independent variable and all other variables as dependent variables without mediator effects. A significantly better fit in all goodness-of-fit statistics was found for the mediation model. The fit statistics of the direct effects models are $\chi^2(116)=811.07$, $p<.05$, CMIN/DF=6.992, NFI=.779, CFI=.803, TLI=.769, RMSEA=.123, SRMR=.207 for the Marketing Channel and $\chi^2(116)=707.84$, $p<.05$, CMIN/DF=6.102, NFI=.740, CFI=.771, TLI=.732, RMSEA=.128, SRMR=.226 for the WOM Channel. This underlines the findings that the acquisition channels effects are mediated on the dependent variables.

---

[53] Only active participation is used to represent the participation variables in order to keep the model and the number of paths at a reasonable size. As the multiple-mediator regression analyses already showed, the effects for the three participation variables are similar; therefore the most relevant variable “active participation” is kept in the model. Please note that the SEM model does not control for age, gender and tenure.

[54] Although not all goodness of fit criteria meet the strict cutoff values of Hu and Bentler (1999), the models are still providing a good fit. Particularly, the important SRMR indicator is below the suggested cutoff of .08 (Hu and Bentler 1999).

5.5 Discussion

Interpersonal communications are often regarded as effective channels to acquire new customers. In this study, the difference between personal selling and WOM and between online and offline WOM as channels to acquire new users are explored. In line with past research on WOM recommendations (e.g., Engel, Kegerreis, and Blackwell 1969; Herr, Kardes, and Kim 1991; Katz and Lazarsfeld 1955; Schmitt, Skiera, and van den Bulte 2011; Villanueva, Yoo, and Hanssens 2008), the results confirm that WOM is more beneficial for the online community than other marketing channels. However, this study adds insights on how specific interpersonal channels differ in the form of consumer-to-consumer and employee-to-consumer communications. Although personal selling has been attributed to be a valuable marketing channel (Crosby, Evans, and Cowles 1990; Reynolds and Beatty 1999), it is demonstrated that WOM-referred users outperform users acquired by personal selling on attitudinal and behavioral variables. Beyond existing WOM literature, this is the first study to examine the difference between online and offline WOM with regards to post-adoption attitudes and behavior. The results reveal that offline WOM recommendation shows stronger direct effects on passive participation, WOM intention and offline WOM provision than online WOM referrals. In addition, WOM-referred users show significant and positive indirect effects on all dependent variables through the mediation of social interaction value. Thus, offline WOM is regarded as more advantageous than online WOM for attracting engaged and active users.

Motivations to Use the Online Community. The results suggest that different acquisition channels attract different types of users. Specifically, social needs to use the platform and connect to other members are emphasized in WOM communications because of the social context of the different communication channels. Offline WOM accentuates the connection and interaction value of the community even more than online WOM. On the other hand, information value is found to be of higher relevance for users who were brought to the community through personal selling communications with employees. The weak tie between sales person and prospective user, who hardly know each other, leads to focus less on social aspects of the platform, but more on the access to relevant information. Though there is no difference in information value between online and offline WOM referred users. The differences in motivation provide additional arguments for the superiority of WOM communications over personal selling and of offline WOM over online WOM.

WOM vs. Personal Selling. There are significant direct effects of the Marketing Channel on identification, active participation, passive participation, frequency of use, WOM intention, and online and offline WOM provision. All direct effects, except on WOM intention, remain significant when including the users' motivations as mediators in the model. This indicates that the social context of WOM referrals has an important influence on the users' post-adoption attitudes and behavior. The relationship to the WOM sender, and therefore the bond to the community, is stronger through WOM. It is proposed that the WOM sender helps to integrate the new user better in the online social community compared to the sales representative, which causes higher social orientation in the use and perception of the platform.

In addition, it is evident that the WOM channel emphasizes the social interaction value, while the personal selling channel focuses on the information consumption value of the online community. Thereby, users attracted through WOM are motivated more to socially interact than personal selling acquired users. The motivation for social interaction (partially or fully) mediates the effect of the Marketing Channel on all attitudinal and behavioral variables because it is strongly related to these outcomes. Particularly, WOM intention is fully mediated. Therefore, WOM referred users are more willing to give WOM in the future, not primarily because they attribute high relevance to the WOM channel, but rather because of the social value they regard to the online community. Nevertheless, WOM intention is also mediated by the information consumption value, suggesting that higher evaluations of the content on the platform lead to higher WOM intentions. Further, the Marketing Channel does not have a significant total direct effect on satisfaction. This can be explained by the different needs and expectations of users coming from the different channels. Both social and informational value perceptions impact the evaluation of the online community which affects the users' satisfaction. The positive effect of the Marketing Channel on social value and its negative effect on information value work conversely, resulting in similar satisfaction levels. Satisfaction is thereby most likely based on the confirmation and disconfirmation of different kinds of needs for the two user groups. Focusing on the expectation-disconfirmation model (Oliver 1997) related to the fulfillment of different needs helps to explain satisfaction levels of the users better than the characteristics of the channels themselves. Altogether, different motives need to be considered in the users' satisfaction and WOM intentions. The value that the users attribute to social and functional (information) aspects leads to higher levels in the outcomes. But more importantly, the channel influences the level of motivation or expected benefits through the sender of the marketing message. The importance of motivations in different channels is supported by the high number of partial mediations that occur.

Offline vs. Online WOM. The type of WOM Channel shows significant direct effects on satisfaction, active participation, passive participation, frequency of use, WOM intention, and offline WOM provision. However, only some direct effects remain significant when considering motivations as mediators. On passive participation, WOM intention and offline WOM provision, there is still a significant direct effect of the WOM Channel. When coming from offline WOM referral, users receive more activity directly from other members on the platform. They are probably integrated well in the community but their own activity is not directly affected strong enough to differ compared to online-referred users. Further, offline-referred users provided more offline WOM themselves and are more willing to give WOM in the future. These users put more relevance to the WOM channel than online WOM receivers and they tend to specifically use the offline channel more frequently. They might value the offline channel because of its direct contact to other people. On the other side, online-referred users might be less convinced by the recommendation they received via the online WOM channel so that they are less willing to give WOM by themselves.

Social motivation plays an important role within the WOM Channel. It mediates the effects of the channel type on all outcomes. It is proposed that the offline contact to the WOM sender provides a richer social context for the transmission of the message and the integration in the online community that it increases the receiver's social needs. Offline-referred users show more favorable attitudes and behavior in the community through social motivation because they have higher levels of social interaction value. Especially active participation, frequency of use and satisfaction are fully mediated by the offline-referred users' social motives. Both user groups do not differ in the overall participation but in their motivation to interact, which in turn drives their participation behavior. Further, there is no direct significant difference in identification with the group of people in the community. As social identity theory posits, the identification of the user is related to the group (Tajfel 1974). Consequently, if offline- and online-referred users can similarly identify with their respective groups of contacts, the acquisition channel would not have a direct effect on group building within the community. However, there is an indirect effect, suggesting that WOM referred users identify more with the community because of their higher social interaction need. If driven by the need to connect and interact, the users know their contacts better through this interaction or are more positively attuned to the members of the community by their willingness to get to know people. Moreover, active WOM provision via the online channel does not differ for the two WOM acquisition channels. One reason might be that both user groups prefer offline WOM provision, as offline is the superior channel to persuade close friends. The key role of social interaction value has been underlined in this section. This is because information consumption value has no relevance in differentiating the WOM Channels, as users from the online and offline WOM channel do not differ in their information consumption motive. Obviously, both groups are equally interested in the content, which may stem from the homophily between sender and receiver with respect to the topic of the online social community.

Key Findings. The acquisition channel plays an important role in evaluating the quality of the users' relationship to the community. Users attracted by WOM and personal selling differ significantly in their post-adoption attitudes and behavior. When focusing on the WOM Channel, the results show that offline WOM referred users are more satisfied and reveal higher participation and WOM activity. These effects can be attributed to two sources. First, the characteristics of the channel are important as different channels provide distinct social contexts in which the communication occurs. The relationship to the sender of the marketing message has an impact on how the new user is integrated in the online community and how his needs are fulfilled. Second, the characteristics of the users, in terms of their activated motives through the acquisition channel, are important factors that influence attitudes and behavior. Distinct motives mediate the effects of the channel type on attitudes and behavior. Therefore, the channel directly and indirectly affects attitudes and behavior, which are developed after registration in the online community. As already proposed by Bolton, Lemon, and Verhoef (2004), specific benefits, i.e. economic and social benefits, inherent in different acquisition channels can lead to a higher value of the customer to the firm. Thus, it is important to understand the channels better in what perceptions they stimulate in prospective customers and how customers differ by channel type. Not only the effects of the channels need to be understood but also the reasons why these effects occur. It is important to differentiate customer groups by certain characteristics, such as their motivation, because different channels attract different types of users with distinct attitudinal and behavioral attributes.

5.6 Managerial Implications

The findings of this study lead to several managerial implications for the use of interpersonal communication channels by marketing managers of interactive online services. The analyses show that WOM-referred users are more beneficial for the platform than users coming from personal selling. Further, it is demonstrated that offline WOM recommendations are superior to online WOM recommendations because offline referred users show higher levels of post-adoption behavior. Consequently, providers of interactive services, be it online social communities or other Web 2.0 services, should promote positive WOM provision. This can be achieved by targeting the influencing factors of WOM giving, thereby increasing customer satisfaction, service quality, commitment, trust, and perceived value (de Matos and Rossi 2008 for the effect of antecedents on WOM activity). Therefore, community operators must focus on the value proposition of the platform towards the users and continuously improve the service focusing on the users' needs. To stimulate WOM provision, referral programs can be an effective means to increase acquisition of highly valuable customers (Schmitt, Skiera, and van den Bulte 2011).

Nevertheless, it would be false to conclude that only WOM users or, more specifically, only offline-referred users are important to the platform. Every user contributes to the overall value of interactive service firms (be it either posters or lurkers), thus the service provider should target different types of users through different channels. This study shows that the different channels attract differently motivated users. Therefore, the use of various acquisition channels can facilitate the building of a more balanced customer portfolio. Addressing distinct user segments helps to reduce risk and volatility (e.g., Tarasi et al. 2011), for example in site traffic generation. The service firm can more accurately plan site visits when providing a solid and high quality content offering because most users acquired via personal selling focus more on informational content. The satisfaction of information needs is less volatile to friends of users switching to other platforms and will bring more stable results through users who are less loyal, but also of less risk to the firm. Thus, personal selling is still an adequate acquisition channel to address information-driven users who are outside the existing social networks of members and who would otherwise hardly be reached. In order to effectively attract those users to the online service, sales representatives should be developed and a stronger relationship marketing orientation in the sense of a partnering role with the users is demanded. Gremler, Gwinner, and Brown (2001) even found that interpersonal bonds or relationships between employees and customers can significantly influence positive WOM provision. A closer connection of sales people and prospective customers could also increase the social needs of personal selling acquired users, though the costs of building and maintaining more intense relationship must be taken into account.

Both WOM channels are important to reach different users. Therefore, it is proposed to stimulate online and offline WOM referrals in parallel and distinctively. As described above, offline WOM constitutes stronger and closer ties and thus facilitates faster integration into existing online social networks of the WOM sender. In consequence, offline WOM can be promoted through offline events which stimulate offline communication between members and non-members. On the other hand, online WOM should be used to reach weaker and more distant ties. As Granovetter (1973) notes, weak ties are important and valuable to spread information to people outside the current network, reaching further sub-networks and new customers. Here, incentives for online recommendations, like offering virtual currency incentives, premium services or small gifts, could help to promote this. Many online services already introduced vouchers or monetary rewards for recommendations (e.g., Xing, Groupon, etc.). Ryu and Feick (2007) find that rewards are particularly effective in increasing referrals to weak ties. They claim that for weak ties, giving a reward to the provider of the recommendation is important; for strong ties, providing at least some of the reward to the receiver of the referral seems to be more effective. Moreover, easy to use referral tools, like tell-a-friend functionality, should be used to increase recommendations to friends and acquaintances with weak tie connections.

5.7 Limitations and Future Research

This study reveals important insights for researchers and practitioners. Nevertheless, it comes with some limitations. As it is the first study to take a closer look at the effects of online versus offline WOM channels, specifically the recommendation to join an online social community, this study should be repeated in other settings to validate the results of the difference between online and offline referrals. Although online communities share many similarities with other online interactive services, like communication tools, news websites with comment functionalities, or product review websites, they provide a specific context, as social networks might play a stronger role than in other settings. It is proposed to study the effects in other industries, which are not as dependent on online social network elements.

The results suggest that users coming from certain channels, that is WOM and more specifically offline WOM, are more beneficial for the online community as they spread the word and participate in the community. However, this study does not focus on the financial benefits users from different channels generate for the operator. The next step would be to translate the positive attitudes and behavior of users acquired by different channels into monetary value. It is therefore suggested to focus on these financial consequences when researching different interpersonal acquisition channels.

Further, the study mainly focused on one-to-one or one-to-few communications between senders (sales representatives or WOM givers) and prospective customers. As especially online WOM provides easy-to-use tools on the Internet to spread the word in a one-to-many fashion, a comparison between WOM channels with a broader audience could be promising.

In addition, WOM can occur in online and offline channels simultaneously. The effect of overlaps in channels and an integral use of different channels would need to be further investigated in order to gain a better understanding on how these two channels interrelate.

Finally, this study contributes to a better understanding on how user acquisition channels differ. However, the focus here was to investigate the direct effects and the mediator role of user motivations. An investigation of further factors that impact the effect of the WOM channel on attitudes and behavior should be considered in future studies.

6 User Activation: The Effects of Individual Network Structure, Attitudes and Motivation on User Participation

The effects of different communication channels to acquire new users were presented in the last chapter. After becoming members of the online social community, the users develop certain perceptions and attitudes towards the community and show different levels of active participation behavior. In order to make the community more attractive for the total user base, community operators need to stimulate active user participation. This chapter is concerned with the influencing factors of active participation and how certain user groups differ in the effects of these factors.

6.1 The Need for Understanding the Drivers of Active User Participation in Online Social Communities

Facebook, the largest social networking site on the web, states that it connects more than 800 million people and that the average user creates 90 pieces of content per month (Facebook 2011). Besides Facebook, there are hundreds of other online communities in the market. Regardless of their business model and orientation – be it social networking, knowledge-sharing, problem-solving, leisure and travel, or any other theme – user participation is the central aspect for their success. From a marketing perspective, participation and involvement can significantly impact an individual’s brand engagement, brand loyalty and recommendation behavior through brand communities (e.g., Algesheimer, Dholakia, and Herrmann 2005; McAlexander, Schouten, and Koenig 2002; Shang, Chen, and Liao 2006; Woisetschlaeger, Hartleb, and Blut 2008) and create value for other members through user conversations and product reviews (e.g., Godes and Mayzlin 2004; Hennig-Thurau et al. 2004; Nambisan and Baron 2007). For many online communities, member participation is not only a means to strengthen a brand or create content in a cost-efficient way, it is also an important outcome itself, as it affects revenues, for instance through advertising (Clemons 2009; Trusov, Bodapati, and Bucklin 2010).

As users connect and interact with each other in online communities, their behavior is affected by the social structural context and their relationships to other users in the community (e.g., Tsai and Ghoshal 1998). For community operators, social network measures are easy to retrieve. Having objective information on how users are influenced by their position in the network can help them shape the network in a way to increase user participation. Despite the rise of the network paradigm in marketing (Achrol and Kotler 1999; Algesheimer and Wangenheim 2006) and the managerial importance of user participation and its relationship to the social context, a deep understanding of network structure as a driver of online community participation is still missing. In particular, this study tests the effects of the user's position in the network in terms of his degree and betweenness centrality, ego-network density, and the online-offline configuration of the friend network on user participation.

Nevertheless, the impact of attitudinal factors should not be underestimated. Scholars from various disciplines have investigated such predictors of online community participation, including satisfaction, social identity or motivations, only to name a few (e.g., Dholakia, Bagozzi, and Pearo 2004; Wasko and Faraj 2005; Woisetschlaeger, Hartleb, and Blut 2008). Besides structural components, certain self-reported attitudinal and relational measures are valuable to explain user behavior when objective network measures are available. This study is the first to combine attitudinal factors and the individual user's network configuration within a closed and complete online social community as predictors of active user participation.

Besides the need to understand the antecedents of user participation, their effect can depend strongly on the types of users. Users may have different motivations to participate in an online community. Thereby, users with different motivations show distinct attitudes and behavior (e.g., Dholakia, Bagozzi, and Pearo 2004; Hennig-Thurau et al. 2004; Matthwick, Wiertz, and de Ruyter 2008). One central motivation of online community users is the need to interconnect with others (e.g., Dholakia et al. 2009; Wiertz and de Ruyter 2007). Although individual motivations and needs have been recognized as important drivers of attitudes and behavior in online communities, its moderating role for the effects of the social context on user behavior has not been studied so far. In this study, it is demonstrated that a differentiation of users' participation behavior and its predictors by the users' social motives provides further insights on how specific user groups need to be treated according to their needs.

In summary, this study contributes to the existing literature on online community participation in two ways. First, the effects of structural measures in combination with the user's attitudes on online community participation are investigated. Social capital theory provides an appropriate frame for understanding the effect of such a combination of structural and attitudinal data on active user participation in online communities. Structural measures on the individual's position in the network are included, which have not been used before to get insights on user participation, i.e. centrality, ego-network density and the online-offline configuration of the user's network. Although recent research utilized actor networks to identify the influence of users on other members' behavior in the community (Trusov, Bodapati, and Bucklin 2010) or to examine diffusion in and of social networks (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011), there is a need to better understand their impact on an individuals' online community participation. In order to investigate the relevance of additional attitudinal measures, satisfaction, identification and reciprocity are also included as predictors of user participation.

Second, the effects on user participation are compared for two user groups, based on their motivations to participate. For the community provider, it is crucial to know how specific user groups are motivated to participate and how they can be stimulated to contribute more in the community. Past research has based group comparisons only on user or community characteristics, like length of membership (Matthwick, Wiertz, and de Ruyter 2008), brand knowledge (Algesheimer, Dholakia, and Herrmann 2005), size of community (Dholakia, Bagozzi, and Pearo 2004), or level of participation (Ridings, Gefen, and Arinze 2006). In this study, users are differentiated by their level of motivation for social interaction. This distinction has interesting implications for customer management.

6.2 Research Overview on Structural Elements as Predictors of Online Community Participation

The reasons for the interest in online community participation are manifold. The users' participation in online communities can generate value through (1) revenues from advertising, membership fees or selling transactions, (2) increased customer loyalty and retention, (3) innovations derived through ethnographic observation of customers, or (4) increasing organizational efficiency and effectiveness by sharing knowledge, solving problems of other customers or providing self-service to employees, experts, or other interest groups (e.g., Algesheimer, Dholakia, and Herrmann 2005; Clemons 2009; Dholakia et al. 2009; Hagel and Armstrong 2006; Nambisan and Baron 2007; see also chapter 2.3 for a discussion of an online community's value for marketing). Consequently, a large number of studies emerged to explain the influencing factors of user participation in online communities. In chapter 2.5, an extensive overview of literature investigating online community participation is given. Basically, from a marketing perspective the most relevant factors investigated in past research can be subsumed as individual attributes, attitudes and perceptions towards the community, and network structure. First, individual attributes, such as motivations and personal characteristics have been widely adapted as predictors of participation in online communities (e.g., Dholakia, Bagozzi, and Pearo 2004; Hennig-Thurau et al. 2004). Second, an even larger set of different attitudinal factors, like identification with the community, commitment, satisfaction, involvement, or trust have been demonstrated to influence user intentions and behavior (e.g., Algesheimer, Dholakia, and Herrmann 2005; Casalo, Flavian, and Guinaliu 2008b; Woisetschlaeger, Hartleb, and Blut 2008). Finally, as the interconnection between users is more visible in online social networks of today, there is a growing interest in the impact of the online network structure on behavior. Though, only a few studies investigated the effect of structural components on behavior in the area of online communities (see Table 14 for a summary of selected studies). In recent research, scholars used network measures to explain adoption behavior (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011), the influence of a user's log-in behavior on other user's activity (Trusov, Bodapati, and Bucklin 2010), and success of open source software communities on a community level (Toral et al. 2009). With regards to individual participation in the community, most studies focused on the users' number of contacts or the overall strength of social ties to other users in the community as the only structural measures (Chen 2007; Chiu, Hsu, and Wang 2006; Nov and Ye 2008; Wasko and Faraj 2005). However, the explicit constitution of the users' networks of contacts, in conjunction with their attitude towards the community, has not been studied as predictors of active user participation. Although community participation can take on different forms, it is crucial that users provide enough content and interact with other users to keep the community alive. Therefore, the focus of this study is to understand how structure and attitudes influence active user behavior.

Authors	Structural variables	Attitudinal variables	Moderators	Outcome	Data Source
Wasko and Faraj (2005)	Centrality (degree)	Reputation, enjoy helping, self-rated expertise, tenure in field, commitment, reciprocity		Volume and helpfulness of knowledge contribution	Objective data and survey
Chiu, Hsu, and Wang (2006)	Social interaction ties	Trust, norm of reciprocity, identification, shared language, shared vision, personal outcome expectations, community-related outcome expectations		Quantity and quality of knowledge sharing	Survey
Chen (2007)	Post-usage social interaction ties	Website use satisfaction, knowledge quality, system quality		Continuance intention	Survey
Nov and Ye (2008)	Structural embeddedness (number of contacts of a user)	Enjoyment, commitment to the community, self development, tenure in community		Photo sharing (number of user‘s public photos)	Objective data and survey
Toral et al. (2009)	Betweenness centrality, number of core users and brokers, ratio of core developer and active developers, Outdegree of core and active developers			Success of open source software communities (number of threads, size, number of active developers)	Objective data
Goldenberg et al. (2009)	Hubs (large number of social ties)			Adoption of virtual items	Objective data
Trusov, Bodapati, and Bucklin (2010)	Log-in of contacts			Log-in decision	Objective data
Katona, Zubscek, and Sarvary (2011)	Degree, betweenness, network density, population density, network size		Network density	Community diffusion - probability of adoption	Objective data
This Study	Degree, betweenness, ego-network density, share of real-world contacts	Satisfaction, identification, reciprocity	Social motivation	Active community participation (interaction with other users and the community)	Objective data and survey

Table 14: Study 2 – Overview of Online Community Studies including Structural Components

6.3 Theoretical Background and Hypotheses

6.3.1 Social Capital in Online Social Communities

Naturally, in online community research, the question arises why users contribute to such communities. Social research suggests that individuals choose to participate because of the existence of social capital (Coleman 1990; Putnam 1995). In chapter 3.3, the concept of social capital was introduced, which provides a relevant theoretical basis for understanding online community participation. Thereby, social capital emphasizes the value of social structures, where relationships are an important resource for social action (e.g., Bourdieu 1986; Burt 1992; Coleman 1988). Capital is generated not only by the independent users of the online community, but specifically from the relationships and interactions with other users and the platform. In online social communities, the value is inherent in the connection of users, the communication between them and the exchange and combination of knowledge, information and digital goods (such as photos or videos). Users connect and interact with each other in various forms, e.g. establishing friendships, exchanging messages, providing support to others, submitting comments or ratings. Relationships are established through these interactions and are a source for social capital. Thereby, social capital in online communities is relevant for the individual users as well as the collective community, because social capital can be obtained and used by individuals for their personal benefit, and be a public good that is collectively owned and serves the community as a whole (Bourdieu 1986; Burt 1992; Burt 1997; Coleman 1988). Social capital in online communities can consequently lead to functional and social benefits, like the exchange of information, knowledge and social support (e.g., Adler and Kwon 2002; Matthwick, Wiertz, and de Ruyter 2008).

As outlined in chapter 3.3, Nahapiet and Ghoshal (1998) propose three dimensions of social capital: structural, relational and cognitive capital. The ***structural dimension*** constitutes the properties of the focal user's position in the network, which influences the user's behavior (Tsai and Ghoshal 1998; Wasko and Faraj 2005). In this study, it is measured by a user's centrality, ego-network density, and his network configuration in terms of online and offline contacts. The ***relational dimension*** focuses on the particular relations users have and that influences behavior. Among its main elements are identification, obligations and expectation (Nahapiet and Ghoshal 1998). For example, in online social communities two users may occupy equivalent positions in the network having the same number of friends, but if they differ in their personal and emotional attachment to other users (e.g., in terms of identifying oneself with the contacts), their actions are also likely to differ. The relational dimension is described by identification, reciprocity and satisfaction in this study. The ***cognitive dimension*** refers to those resources providing shared representations, interpretations, and systems of meaning within groups (Nahapiet and Ghoshal 1998). It also consists of individual expertise, which facilitates the understanding of a common language, and experience with applying the expertise (Wasko and Faraj 2005). As many communities are built around specific topics, it is expected that at least part of the cognitive dimension is related to the tenure in the community, because established users get to know the shared language used in the community and have knowledge about the community topics. Hence, this study uses tenure as an indicator for the cognitive dimension.

Online social communities are based on user connections and interactions. The users' voluntary communication, interaction and establishment of ties with other users foster social capital (e.g., Matthwick, Wiertz, and de Ruyter 2008). Social capital is therefore inherent in such online social networks and in many cases visible in form of explicit friend lists and public discussions. Social capital theorists argue that social capital is the precursor of combination and exchange of resources (Nahapiet and Ghoshal 1998), where active user participation describes this combination and exchange, including users' interaction and information exchange. Therefore, it is an explicit consequence of the relationships between users and with the community as a whole. In this study, the focus is on the individual user level in order to better understand the impact of social capital on user participation. In the following section, hypotheses are developed relating the single dimensions to user participation in the online community as an outcome of social capital.

6.3.2 Impact of the Social Capital Dimensions on User Participation

Centrality. Social capital theory posits that structural capital is driven by the user's constitution of and the position in the network (Tsai and Ghoshal 1998). Thereby, network ties provide access to resources and thus increase the potential of valuable information flows (Nahapiet and Ghoshal 1998). An adequate concept to measure the structural dimension is centrality. Although, several approaches to measure centrality exist, degree and betweenness centrality are simple to calculate and the most commonly used centrality measures (e.g., Freeman 1978/79; Wasserman and Faust 1994).[55]

The concept of ***degree centrality*** (or degree) accounts for the number of ties, which are adjacent and thus directly related to an actor (Freeman 1978/79). Therefore, social capital in form of network ties provides potentially better access to information, which is an important basis for action (Coleman 1988; Nahapiet and Ghoshal 1998). Actors with a large number of sources are central to the network, because they can more easily obtain information, have more sources of information at their disposal, and can reach information quicker (de Nooy, Mrvar, and Batagelj 2005; Borgatti 2005). In an online community, a higher number of contacts may lead to more interaction, as many different sources of information are available for the focal user. On the other side, actors with low degree are likely to be seen as peripheral by others. They are in a more isolated position with less direct interaction with most other actors of the network and few contact points to active participation and communication (Freeman 1978/79).

Ahuja, Galletta, and Carley (2003) and Sparrowe et al. (2001) found that centrality was a strong predictor of individual performance in workgroups. Wasko and Faraj (2005) showed that degree centrality is an important measure that positively impacts knowledge contribution in online communities. Further, Stephen and Toubia (2010) demonstrated that in online seller networks where shops are linked with each other, shops with higher indegree had higher commission revenues. This leads to the following hypothesis:

H1: The user's degree centrality exhibits a positive effect on active user participation.

In addition, ***betweenness centrality*** (or betweenness) takes into account the position of actors in the entire network, and not only to their direct neighbors. It describes the central position of the user in terms of the actor's location on the shortest paths (geodesics) between all other pairs of actors in the network (de Nooy, Mrvar, and Batagelj 2005; Wasserman and Faust 1994). The structural hole argument posits that actors who take on brokerage roles between other actors, thus bridging between sub-groups, have an advantageous position as they can facilitate communication and have access to non-redundant information (e.g., Burt 2000). Granovetter's (1973) "strength of weak ties" approach also underlines the importance of

[55] A more detailed description of network measures and an overview of the mathematical definitions of centrality can be found in chapter 3.2. This study focuses on degree and betweenness centrality, because closeness centrality cannot be calculated in weakly connected networks, which is the case for the empirical research object.

bridging between different circles of friends. This access to new information can stimulate the communication activity of the broker. Central actors might have control over the communication flows between other actors, but more notably they are important intermediaries in the communication network, i.e. they are involved in the interaction of the other actors (de Nooy, Mrvar, and Batagelj 2005; Freeman 1978/79). As information more likely passes the central user, forwarding messages, commenting this information, or getting involved in discussions should occur more often.

Empirical research has found positive relationships between an individual's betweenness centrality and individual performance (Brass 1984; Cross and Cummings 2004; Mehra, Kilduff, and Brass 2001). Further, Tsai and Ghoshal (1998) used betweenness centrality to assess the impact of structural capital in business unit relationships on resource exchange and combination and found a significant positive effect. This leads to the following hypotheses:

H2: The user's betweenness centrality exhibits a positive effect on active user participation.

Ego-network density. Not only the focal actor's ties to other actors constitute an important facet of social capital, but also the configuration of the overall network, where network density is one of the properties of network structure (Nahapiet and Ghoshal 1998). The density of the ego's network describes the number of relationships between direct neighbors of the focal actor in relation to the maximum number of relations between these neighbors (de Nooy, Mrvar, and Batagelj 2005). From a social capital perspective, densely connected individuals lead to network closure, that on the one hand increases trust between the related actors, but at the same time creates redundant paths for information flow (Burt 1992; Burt 2000; Nahapiet and Ghoshal 1998). Network closure can have a negative effect on performance: a dense ego-network increases the probability of its members to know the same information and decreases the opportunity to broker through direct connections among the members (Burt 2000). If the friends' network is strongly connected, the users' information does not have to travel through the focal user, but could also bypass this user. Thus, individuals have less information and control advantages in closed ego-networks. Granovetter (1973) supports the view that a dense network may result into redundant information, because of the absence of relationships to more distant social circles. In contrast, if a user's network is weakly connected the focal user can take on the connector role between friends and therefore centralize activity within the friend network (with the extreme form of a star network).

In a comparison of different studies on social capital and performance, Burt (2000) concludes that network density has a significantly negative association with performance in all study populations. This leads to the following hypothesis:

H3: Ego-network density exhibits a negative effect on active user participation.

Share of real-world friends. Especially when locally organized around geographic regions and cities, online social communities emphasize the integration of online and offline friend networks. Wellman and Hampton (1999) found that online and offline relationships are intertwining. Because structural capital also lies in offline ties, those connections make the resources for information exchange available in the offline world, meaning that there might be less need to use the online channel for interactions. In this case, the Internet may compete for time with offline interaction. As online interactions are inherently inferior to face-to-face interactions, the relationship of offline and online is negative (Wellman et al 2001; Cummings, Butler, and Kraut 2002). Having many online ties also accessible in the offline world leads to the use of offline ties for better access of the information. In addition, online communities are an appropriate mean to maintain intermediate-strength and weaker ties with people one does not meet regularly offline (Wellman et al. 1996). These online contacts potentially provide diverse and non-redundant information (Granovetter 1973), which leads to a higher social capital in the online community derived from those online contacts. Thus, it is hypothesized:

H4: A higher share of real-world friends in the online community has a negative effect on active user participation.

Identification. Recent research suggests that the relational dimension of social capital influences the access to parties for exchange, as well as the motivation to engage in content creation through exchange and combination (Nahapiet and Ghoshal 1998). As one element of the relational dimension, identification with a group can be considered as the sum of the relationships a user has with other users of the online community. In this context, the users' identification with the group means that they come to view themselves as a member of the community, as 'belonging' to it (Dholakia, Bagozzi, and Pearo 2004). Social identity[56] is closely

[56] For a theoretical introduction to the concept of social identity see also chapter 3.4.1.

related to consciousness of kind, which is an important element in communities. It is defined as a feeling of intrinsic connection between members and a collective sense of difference from others not in the community (Muniz and O'Guinn, 2001). The user's identification expresses a kind of tie strength to all contacts of the user. The stronger the identification with other users in the ego's network, the higher the strength of the relationship.

From a social capital perspective, stronger ties among members enhance the concern for collective outcomes, and thus increase the opportunities to exchange information and to communicate (Nahapiet and Ghoshal 1998). Therefore, identification with the community can act as a resource influencing the motivation to interact with other users and participating in the community. Different studies demonstrated the significant positive effect of identification and social identity on the individual's participation in the community (e.g., Algesheimer, Dholakia, and Hermann 2005; Dholakia, Bagozzi, and Pearo 2004; Woisetschlaeger, Hartleb, and Blut 2008). This leads to the following hypothesis:

H5: Identification exhibits a positive effect on active user participation.

Reciprocity. Another important element of the relational dimension of social capital are obligations and expectations. Obligations are expectations developed within particular personal relationships (Coleman 1988). They represent a duty to undertake some activity in the future. This sense of moral responsibility is what produces collective action and contributes to group cohesion (Muniz and O'Guinn 2001). In this context, a norm of reciprocity can develop in online communities between its users. Generally, this norm of reciprocity demands that people should help those who have helped them or give back when they received (Gouldner 1960). It "usually refers to a set of socially accepted rules regarding a transaction in which a party extending a resource to another party obligates the latter to return the favour" (Wu et al. 2006). Individuals who are treated favorably by others feel a sense of obligation to repay benefits received through reaction or return of the favor (Gouldner 1960). On the basis of reciprocity, there are two reasons for the users to actively participate: first, users expect the other community members to react on their posts or ratings; second, the users might feel obliged to answer or react when other users contact them. Therefore, high degrees of reciprocity describe an attitude of give and take within a community. Usually, feelings of obligation emerge on the basis of their past behavior (Gouldner 1960). This means, if users know that others reciprocate they would establish a stronger personal norm of reciprocity. Ridings, Gefen, and Arinze (2002) emphasize the importance of trust and reciprocity for the survival of an online community, as they impact the intention to give information. In online communities, a norm of reciprocity motivates users to share information and knowledge with other members, because of their moral obligation (Chiu, Hsu, and Wang 2006; Wasko and Faraj 2005). This leads to the following hypothesis:

H6: Reciprocity exhibits a positive effect on active user participation.

Overall satisfaction with the community. In addition to obligations, expectations shape the users' overall satisfaction with the community. Satisfaction is regularly based on the expectancy disconfirmation theory and is most often measured as the difference between customer expectations and the actual service or product performance (Oliver 1997). When the product or service exceeds the customers' expectations they will be satisfied. Within the relational dimension, such expectations are likely to influence the users' motivation to contribute to the community and interact with others (Nahapiet and Ghoshal 1998). It is not only the expectation toward single users but also the expectation towards the community as a whole which determines the users' behavior. Thereby, satisfaction is an overall evaluation of performance (i.e. if expectations are fulfilled) and it is based on prior experiences (Anderson and Fornell 1994). When users do not expect that their needs will be satisfied in the future, based on past experiences and disconfirmed expectations, they are less likely to participate in community activities. Satisfaction has been demonstrated to influence participation and continuance intentions in online communities (e.g., Chen 2007; de Valck et al. 2007; Woisetschlaeger, Hartleb, and Blut 2008). Therefore, it is hypothesized:

H7: Overall user satisfaction exhibits a positive effect on active user participation.

Tenure. The cognitive dimension of social capital shall provide a basis for common understanding and interpretation of the communication and interaction that occurs in the online social community (Nahapiet and Ghoshal 1998). Shared language and narratives can help to make sense of the information and content on the platform. To gain a common language, Wasko and Faraj (2005) propose that the individual expertise and experience with the matter of subject is important. Because online communities are often centered around specific topics, users can gain experience and expertise through the usage of the online community itself. Cognitive capital develops through the interaction with other users over time, so that the user learns the skills, knowledge and norms of practice (Wasko and Faraj 2005). Thus, tenure is viewed an indicator of the expertise users gain about an online community topic. However, it does not directly measure the cognitive capital, which inheres in the community members. Consequently, this study uses tenure as a variable to control for the effects of longer duration in the online community, as has been done in past studies. Therefore, no explicit hypothesis is derived in this respect, however, the results for tenure are also reported and discussed in the empirical study section.

6.3.3 Moderating Effects of Social Motivation

From a social capital perspective, individual attributes, such as motivations and abilities, are important to explain why some individuals build up more social capital and engage more willingly in collective action than others (Adler and Kwon 2002; Nahapiet and Ghoshal 1998). Past studies investigated the motives of individuals to use online social communities, with the value of the single motives depending on purpose or type of the community (e.g., Dholakia, Bagozzi, and Pearo 2004; Park, Kee, and Valenzuela 2009; Sangwan 2005). From a managerial perspective, it is important to consider that different user types or user motivations emphasize specific influencing factors of behavior more than others. In online communities, the need to interconnect with other users is a central motivational factor. Recent research has already demonstrated this importance (e.g., Dholakia et al. 2009; Wiertz and de Ruyter 2007; Mathwick, Wiertz, and de Ruyter 2008). Especially in online communities which offer editorial content in addition to user generated content, users can be differentiated by high and low levels of motivation to connect with other users, because value is not only inherent in user interaction, but also in firm-offered content. Basically, it is expected that users with higher motivations to interconnect with other community members (called 'Networkers') have a higher interest in the social networking functions of the website and therefore engage more in the community than those with low social motivations (called 'Non-Networkers'). When being motivated by social needs, those users are more likely to be aware of their position within the network and their role vis-à-vis other members. Wiertz and de Ruyter (2007) found that online interaction propensity, a personal trait to interact in the online community, has a significant effect on knowledge contribution. Hennig-Thurau et al. (2004) showed that based on user motivations different user segments emerge which also differ in their participation behavior on online opinion platforms, suggesting that more socially driven users participate more. Therefore, Networkers are expected to be more active in the community and should be more influenced in their number of contributions by the structural and relational factors. On the other hand, a low level of social need implies that the influence of the predictors on the decision to contribute at all should be more important for Non-Networkers.

H8: Users motivated by social needs show higher levels of centrality, satisfaction, identification, and reciprocity and a lower level of ego-network density than users not motivated by social needs.

H9: The impact of structural and relational capital on user participation is
a) more important for Networkers with respect to the number of contributions, and
b) more important for Non-Networkers with respect to the likelihood of making any contributions.

6.4 Empirical Study – Methodology

Figure 14 summarizes the hypotheses developed in the previous section. To test the hypotheses, data from the online social community is used, which is introduced in chapter 4. The analyses are conducted in three steps: Analysis 1 addresses the effects of network measures on user participation for the entire sample of current users. Analysis 2 complements the structural predictors by including attitudinal measures from the survey. Analysis 3 tests for the moderating effect of social motivation to use the community by splitting the sample in two groups and comparing those groups. For these analyses, two data sources are used: self-reported survey data and objective data.

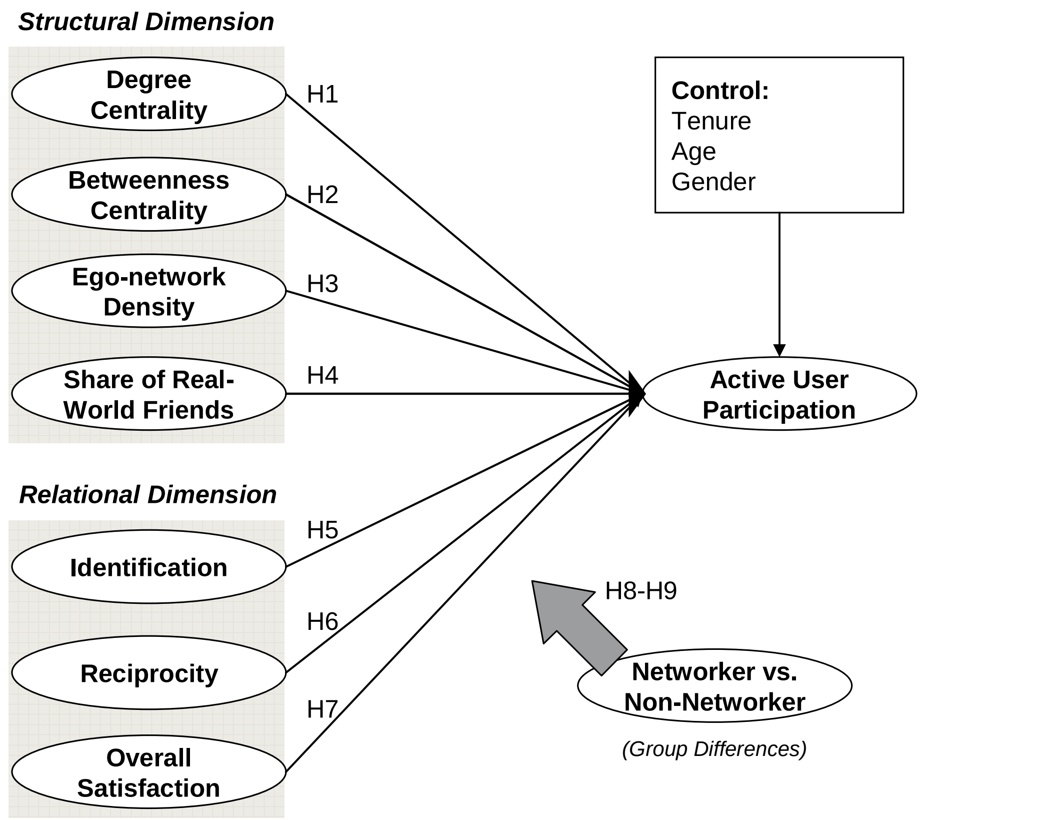


Figure 14: Study 2 – Influence Model on Active User Participation

6.4.1 Data and Sample Characteristics

A two-step procedure is used to gather the necessary data. Two successive time intervals, T1 and T2, are defined, which represent a period of two months each. In T1, the network measures are collected for all users who logged-in at least once in T2 and registered before T2, thus ensuring that all subjects were members in the community in T1 or before. On this way, the structural dimension of each user is extracted, describing the individual position at the end of T1. An overall of 30,666 nodes with at least one contact in T2 were present in this network of current users, representing 1,174,254 connections between those users.[57] In T2, participation data of all users, who were active in T2 and already registered in T1, are collected. Participation was tracked for the entire two months of T2. The temporal separation of

[57] The focus is on users with at least one active contact, because users without contacts a) do not use the platform for social networking, where the objective is to investigate the impact of social structures on user behavior, b) do not have the opportunity to use all of the interactive functionality (e.g. sending messages to friends) to participate, and c) consistently have zero or missing values in their network measures, and thus do not show any variation in those measures. The network of active users (logged-in in T2) is used, because users who left the community for good do not provide any value to the focal user anymore.

network structure and behavior permits to test for the hypothesized effects in the intended direction. In addition, self-reported data was obtained from the online survey described in chapter 4. The data set includes 486 usable survey responses. Again, the survey sample comprises users who registered before T2, logged-in in T2, and have at least one active contact in T2, to have a common basis for both the survey and the network sample.[58]

In the same way as in Study 1 (chapter 5), non-response bias is tested by two approaches. First, the responses of early and late respondents are compared by using time-trend extrapolation (Armstrong and Overton 1977). All variables were tested by the use of t-tests, comparing the mean values of early and late respondents. No significant differences appeared at $p<.05$ level (see Appendix 6). These results suggest that non-response bias is not likely to be a major concern for the study. Second, it is tested for differences between respondents and non-respondents regarding their general characteristics concerning the platform (i.e. age, gender, tenure, degree and betweenness centrality, ego-network density, active participation). A comparison is conducted between the 486 users of interest, who submitted their completed questionnaires, and the remaining 30,180 active users in the overall network. T-tests of degree and betweenness centrality, ego-network density, and gender did not show significant differences. However, t-tests of age and tenure showed significant differences. Nevertheless, the mean age of the respondents is only 1.2 years higher and the mean tenure 3.2 months lower than that of the non-respondents. In order to account for possible effects of age, gender and tenure, they are included as control variables in the analyses. Last, active participation was significantly higher for the users in the survey sample. However, Table 15 shows that there is still a high share of users with low participation levels in the online social community, so that not only heavy users are in the survey sample. In order to check for the validity of the results regarding the impact of the structural dimension on active participation, the analysis for the entire network sample is pursued in Analysis 1 of this study. An overview of the sample characteristics can be gained from Table 16.

[58] Because of the consideration of two timeframes and the explicit network of active contacts, users with missing values in the objective measures are not included in the relevant survey sample for this study.

	Network Sample		Survey Sample			Network Sample		Survey Sample	
# Activities	# Users	% Users	# Users	% Users	# Activities	# Users	% Users	# Users	% Users
0	17,973	58.61	175	36.01	25-29	288	0.94	11	2.26
1	2,771	9.04	31	6.38	30-39	465	1.52	17	3.5
2	1,430	4.66	30	6.17	40-59	598	1.95	22	4.53
3	857	2.79	12	2.47	60-79	382	1.25	17	3.5
4	653	2.13	12	2.47	80-99	266	0.87	8	1.65
5	484	1.58	11	2.26	100-199	621	2.03	14	2.88
6	436	1.42	13	2.67	200-299	265	0.86	13	2.67
7	335	1.09	10	2.06	300-499	222	0.72	6	1.23
8	294	0.96	11	2.26	500-999	176	0.57	6	1.23
9	235	0.77	5	1.03	1000-1999	63	0.21	4	0.82
10-14	871	2.84	25	5.14	2000-3999	19	0.06	3	0.62
15-19	539	1.76	9	1.85	4000-9999	6	0.02	2	0.41
20-24	417	1.36	19	3.91	ALL	30,666	100%	486	100%

Table 15: Study 2 – Distribution of Active User Participation Levels

		Survey Sample[1)]		Network Sample[2)]	
Network and Behavioral Data	# of Measures	Mean	Standard Deviation	Mean	Standard Deviation
Active Participation[3)]	1	6.50	7.09	3.21	5.60
Degree[4)]	1	3.07	1.73	3.18	1.68
Betweenness[4)]	1	11.69	5.91	11.58	5.57
Ego-Network Density	1	0.22	0.23	0.21	0.21
Share of Real-World Friends	1	0.41	0.33	-	-
Tenure of Membership (years)	1	1.89	1.28	2.16	1.26
Age (years)	1	22.48	7.97	21.26	5.70
Gender (0= male, 1=female)	1	0.54	0.50	0.52	0.50

Latent Constructs from Survey Sample[1,5)]	# of Measures	Mean	Standard Deviation	Cronbach's Alpha	Lowest Item-to-total correlation	Composite Reliability	Average Variance Extracted
Satisfaction	3	4.65	1.18	0.81	0.64	0.84	0.65
Identification	3	3.36	1.59	0.90	0.78	0.90	0.75
Reciprocity	2	3.08	1.62	0.89	0.81	0.90	0.82
Social Interaction Value	3	3.62	1.87	0.90	0.74	0.91	0.76

1) n=486;
2) n=30,666; except for Age, where n=30,431 due to missing values
3) categorized measure
4) log-transformed measure
5) CFA: executed only for multidimensional measures using ADF (asymptotically distribution-free) estimation

Table 16: Study 2 – Descriptive Statistics and Measurement Model Evaluation

6.4.2 Measurement of Constructs

The empirical studies use either objective data or a mixed approach, combining self-reported measures with objective data from the community platform, which is the preferred method to address common-method bias (e.g., Podsakoff et al. 2003). The factors from the relational dimension, i.e. identification, reciprocity, and satisfaction, are collected through the online survey. They are measured using multi-items from existing scales, adapted to the context of this study. The survey development and the operationalization of the scales are explained in more detail in chapter 4. The share of offline friends is also retrieved from the survey. It is measured by asking the users how many friends they have in the community in total, followed by a question on how many of those friends they regularly meet in the real world. The ratio of offline to total friends comprises the share of real-world friends.

All network analytical measures are calculated using the social networks software Pajek (de Nooy, Mrvar, and Batagelj 2005). Therefore, the network of users in T1, who also logged-in in T2, is used. Based on this active friend network, which is an undirected network of explicit friendships between users, degree, betweenness centrality and ego-network density are computed. For the formula and description of the network measures, see chapter 3.2.2.

Because of the nature of the centrality measures in large online social communities, the variance of these measures is rather high. Therefore, the variance and the effect of extreme outliers should be reduced. To gain more stable results, the variables are transformed in the following way: degree was directly log-transformed (ln(x+1)); betweenness centrality was first multiplied with the inverse ratio of the lowest non-zero value to determine the lowest betweenness value at value 1 and then also log-transformed.

For the dependent variable, an index is constructed, which describes the active participation on the platform. Active participation includes out-going activity, i.e. communication with other users, comments and ratings.[59] The measure uses aggregated activity data from an overall of two month in T2. Active participation is a classical count data variable, but with a very “long tail”. In order to control for outliers and to stabilize the estimates, the observations are categorized in different intervals. This approach is common for discrete count variables and it helps to reduce the variability, without producing large bias in the parameter estimates (Greene 2002). All values less than 10 stay the same, because the majority of users show

[59] Active participation comprises messages sent to other users, guestbook entries written, virtual gifts sent to other users, comments on articles, and submitted ratings on photos, groups, and other users’ profiles.

fewer contributions (>60 % in the survey sample and >80 % in the network sample)[60]. All other values are grouped in intervals. In the participation index, single activities are equally weighted. However, higher weighting of direct user interaction or public content (guestbook entries, comments), leads to similar results in the analysis (see Appendix 7). In Table 15 the observed values for the count variable "number of active contributions" are presented. Table 17 includes the correlations between the single measures.

Measures	1	2	3	4	5	6	7	8	9	10	11	12
1 Active Participation[1)]		.44 **	.30 **	-.05 **	-	-	-	-	-	-.06 **	-.06 **	-.04 **
2 Degree[2)]	.56 **		.83 **	-.04 **	-	-	-	-	-	.41 **	-.14 **	-.03 **
3 Betweenness[2)]	.43 **	.79 **		-.11 **	-	-	-	-	-	.36 **	-.08 **	.01
4 Ego-Network Density	-.13 **	-.05	-.09 *		-	-	-	-	-	-.09 **	-.03 **	-.05 **
5 Share Real Friends	-.31 **	-.22 **	-.13 **	.26 **		-	-	-	-	-	-	-
6 Identification	.28 **	.13 **	.06	.09	.26 **		-	-	-	-	-	-
7 Reciprocity	.31 **	.12 **	.10 *	-.03	-.04	.47 **		-	-	-	-	-
8 Overall Satisfaction	.26 **	.02	-.01	.00	-.03	.36 **	.41 **		-	-	-	-
9 Social Interaction Value	.51 **	.41 **	.32 **	-.08	-.12 **	.50 **	.47 **	.32 **		-	-	-
10 Tenure	.05	.47 **	.37 **	-.15 **	-.11 *	-.06	-.17 **	-.18 **	.04		.15 **	-.06 **
11 Age	.21 **	-.02	.01	-.09 *	-.19 **	.05	.13 **	.06	.14 **	.03		-.20 **
12 Gender	-.15 **	-.03	-.01	-.05	.05	-.15 **	-.17 **	-.11 *	-.13 **	.23	-.26 **	

** p<.05; ** p<.01*
correlations of Survey Sample on white background; n=486
correlations of Network Sample on grey background; self-reported measures not available for Network Sample; n=30,431
1) categorized measure; 2) log-transformed measures

Table 17: Study 2 – Correlations of Variables

6.4.3 Measurement Model Evaluation of Self-Reported Latent Measures

Standard validity and reliability tests for the survey measures are conducted in the same manner as in Study 1 (chapter 5.3.3). All relevant statistics are provided in Table 16. The reliability of the scales was assessed using Cronbach's alpha, with a minimum value of 0.7 being satisfactory (Nunnally 1978). To evaluate the unidimensionality of the proposed scales, exploratory factor analyses for each construct are conducted.[61] To test for internal consistency and discriminant validity, confirmatory analyses are accomplished. Asymptotically distribution-free (ADF) is chosen as an estimation method, since it shows more security in sam-

[60] Intervals for values higher than 9 are used, because most of these values comprise less than 1% of cases in the two samples. Thus, most information resides in the lower values. Additional analyses of the empirical models are pursued, which use different transformations of the dependent variable to verify the results (see below in this study).
[61] Only one factor was extracted from each scale (criteria used: only factors extracted with eigenvalues higher than 1, factor loadings higher than 0.5, and a significant total explained variance).

ples that might not present multivariate normality (Byrne 2010). The tests are executed for the multi-item constructs used in this study.

Two measures to evaluate internal consistency of constructs are used – composite reliability (CR) and average variance extracted (AVE). The composite reliability is a measure analogous to the Cronbach α coefficient (Fornell and Larcker 1981) with a suggested cut-off value of 0.6 (Bagozzi and Yi 1988). The average variance extracted estimates the amount of variance captured by a construct's measure relative to random measurement error (Fornell and Larcker 1981). Estimates of AVE above 0.5 are considered supportive of internal consistency (Bagozzi and Yi 1988). The results, which are shown in Table 16, are satisfactory, and therefore indicative of good internal consistency.

Discriminant validity verifies that a construct is significantly distinct from other constructs that are not theoretically related to it. A CFA-model is used, which includes four latent constructs (satisfaction, identification, reciprocity and social interaction value). Results show that the model fit the data well. The goodness-of-fit statistics for the model are satisfactory (Hu and Bentler 1999): $\chi^2(38)=64.56$, $p<.05$, CMIN/DF=1.699, NFI=.938, CFI=.973, TLI=.961, RMSEA=.038, SRMR=.047. Further, the correlations among the latent constructs are not higher than 0.8 points (Bagozzi 1994). In addition, a test of discriminant validity was performed, as suggested by Fornell and Larcker (1981). Discriminant validity is achieved if the AVE by the underlying construct is larger than the shared variance (i.e. squared correlations) with other latent constructs. This condition is satisfied for all of the cases. In sum, internal consistency and discriminant validity are satisfactory and permitted to include these constructs in the hypotheses tests.

6.4.4 Specification of Count Data Models

The activity of the online social community users can be treated as count data, i.e. each time the user actively participates it is counted as one more activity carried out. In fact, in user databases these activities are literally counted each time the users interact with the platform. For count data, standard linear regression models are often inefficient and would lead to inconsistent standard errors, producing biased predictions for the dependent variable (Chou and Steenhard 2009; King 1989). Count data models are used when the dependent variable takes on non-negative integer values. The initial model for count data is the Poisson regression model (e.g., Cameron and Trivedi 1998; Winkelmann 2003). The Poisson distribution assumes equi-dispersion, i.e. the equality of the conditional variance and the conditional mean (Cameron and Trivedi 1998; Chou and Steenhard 2009). However, the assumption of equi-dispersion implied by the Poisson distribution is very restrictive. In many cases the conditional variance exceeds the conditional mean, which is also called over-dispersion. Thereby, "over-dispersion in a Poisson regression will lead to deflated standard errors of parameter estimates and therefore inflated t-statistics" (Liu and Cela 2008, p. 3). To account for over-dispersion, the standard parametric model is the negative binomial (NB) (Cameron and Trivedi 1998; Hilbe 2007). In NB models, the Poisson model is extended by an individual, unobserved effect, which reflects either specification error or heterogeneity of the data (Greene 2002).[62]

Another problem of estimating count data models adequately is the presence of many zeros in the event variable. In online communities, not all users actively participate and contribute content. Lurkers just browse the platform and consume content instead of contributing. This "excess of zero activity" is problematic for estimating count data models adequately. Zero-inflated models help to model excess zeros and to better understand the impact on frequency data (e.g., Lambert 1992) – such as online community participation – thereby addressing two questions: a) whether or not contributions are observed (binary yes/no decision), and b) how many contributions are made conditional on the fact that at least one contribution is made. Zero-inflated models are two-stage models, in which the binary and the count model are estimated jointly, therefore it can also be considered a mixture of two statistical processes, one always generating zero counts and the other generating both zero and nonzero counts (Wangenheim and Bayón 2007). Zero-inflated models assume that the zero counts come from two sources. In this study, this means that there are users who would never choose to actively participate and others that would, but did not actively participate during the sample period. A logit model is used to determine the probability count outcome to be zero; the standard Poisson or negative binomial count data model then predicts the number of activities made by the user (the not-always-zero group) (Liu and Cela 2008). One advantage is that the possibility is taken into account that specific variables impact either one of the two stages but not the other (for example, see Lambert 1992; Wangenheim and Bayón 2007). An example in the marketing literature is provided by von Wangenheim and Bayón (2007), who show that the predic-

[62] The error term is mainly gamma-distributed and loosens the equi-dispersion assumption of mean and variance. The most common implementation of the negative binomial is the NB2 model. For more information see for example Cameron and Trivedi (1998).

tors of WOM referrals affect both, the likelihood of giving WOM referrals, and the conditional number of WOM referrals, but could also vary in their impact on each stage.

Model Selection. As already noticed, there are different models available to investigate the effect on count data variables. In order to select the model, that best fits the data of this study, four common count data models are compared: a Poisson regression model (PRM), a negative binomial regression model (NBRM), a zero-inflated Poisson model (ZIP) and a zero-inflated negative binomial model (ZINB). Because Poison and negative binomial models are popular models to analyze count data (e.g., Cameron and Trivedi 1998; Greene 2002; Hilbe 2007; Winkelmann 2003), these two models and their extensions to approach excess zeros are considered as adequate models in a first step. The model selection is then based on a) a likelihood-ratio test to compare the fit of two (nested) models, b) the test statistics of alpha (which is a built-in test in NB models) indicating over-dispersion, and c) the Vuong (1989) test to compare non-nested models.

a) A likelihood-ratio test (LR-test) (e.g., Cameron and Trivedi 1998; Liu and Cela 2008) is used to compare the fit of two (nested) models, one of which (the null model, e.g. PRM) is a special case of the other (the alternative model, e.g. NBRM). The LR-tests show that NB models are preferred over Poisson models. The LR-test comparing the regular Poisson and negative binomial models is $-2(LL_{PRM} – LL_{NBRM})=-2[-1864 – (-1282)]=1164$, which is highly significant ($p<.05$) and indicates that the NBRM is preferred over the PRM. For the zero-inflated models, the LR-test is $-2(LL_{ZIP} – LL_{ZINB})=-2[-1306 – (-1177)]=258$, which favors the ZINB over the ZIP model ($p<.05$).

b) The test statistics of alpha indicates over-dispersion, and thus whether the true specification of the model is Poisson or negative binomial (Erdman, Jackson and Sinko 2008). Alpha is significant in all cases ($p<.05$), suggesting over-dispersion and preference of NB models. Therefore, the PRM and the ZIP model can be rejected, favoring the NBR and ZINB model respectively.

c) The Vuong test (see Cameron and Trivedi 1998; Vuong 1989) is an extension of the likelihood-ratio test specified to situations of non-nested models. The PRM and ZIP model as well as the NBRM and ZINB model are not nested, and thus the log-likelihood ratio test cannot be applied. The Vuong tests confirmed the superiority of zero-inflated models over standard count regression models (ZIP vs. PRM and ZINB vs. NBRM), because the Vuong statistics exceeded a value of 1.96 (Liu and Cela 2008). For example, the Vuong statistic for model 2c (see below) is 6.69, indicating that the ZINB model is preferred over the NBRM.

All four models are presented in Table 18 for the survey sample (applied to the final Model 2c) and in Table 19 for the network sample (applied to Model 1). Overall, the results confirm that the ZINB model fits the data best. Thus, ZINB models with logit splitting in the binary model are applied to test the hypotheses. The formal ZINB model description is provided in Appendix 8. The ZINB models are parameterized using the hypothesized independent variables to estimate their effect on active participation. In addition, age, gender, and tenure are included as control variables. As the ZINB model is a two-stage model and the derived hypotheses on the effect on participation are of general nature, all independent variables are specified for both the binary (whether or not the user contributes) and the negative binomial (conditional number of contributions) model. It is expected, that the predictors should have an effect in the hypothesized direction in both models.

			Zero-inflated Poisson		Zero-inflated Negative Binomial	
	Poisson	**Negative Binomial**	**Logit[1)]**	**Poisson**	**Logit**	**NB**
Intercept	0.024	-0.525 **	-3.030 ***	1.155 ***	-2.970 ***	0.991 ***
Degree	0.244 ***	0.329 ***	0.797 ***	0.166 ***	0.800 ***	0.185 ***
Betweenness	0.032 ***	0.034 *	0.034	0.000	0.035	0.000
Ego-Network Density	-0.317 ***	-0.869 ***	-1.167 **	-0.261 **	-1.163 **	-0.355
Share Real Friends	-0.719 ***	-1.017 ***	-1.498 ***	-0.361 ***	-1.490 ***	-0.414 ***
Identification	0.097 ***	0.135 ***	0.092	0.099 ***	0.080	0.117 ***
Reciprocity	0.019	0.020	0.124	-0.004	0.124	-0.007
Overall Satisfaction	0.111 ***	0.140 **	0.271 **	0.065 ***	0.275 **	0.067 *
Tenure	-0.144 **	-0.196 ***	-0.478 ***	-0.078 ***	-0.482 ***	-0.083 **
Age	0.011 ***	0.023 ***	0.048 **	0.008 ***	0.049 **	0.011 **
Gender	-0.085 **	-0.122	-0.194	-0.053	-0.198	-0.061
Alpha		1.3600 ***			0.2820 ***	
Log-Likelihood	-1864	-1282	-1306		-1177	
AIC	3749	2588	2656		2400	

Dependent Variable: Active Participation (# of activities; categorized); n=486
**** p<.01 ** p<.05 * p<.10; NB=negative binomial;*
1) signs in the logit part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros). Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"

Table 18: Study 2 – Model Comparison for Survey Sample

	Poisson	Negative Binomial	Zero-inflated Poisson		Zero-inflated Negative Binomial	
			Logit[1)]	Poisson	Logit	NB
Intercept	-0.376 ***	-0.784 ***	-2.196 ***	1.329 ***	-2.137 ***	1.133 ***
Degree	0.645 ***	0.707 ***	0.817 ***	0.338 ***	0.829 ***	0.381 ***
Betweenness	-0.023 ***	-0.027 ***	-0.029 ***	-0.026 ***	-0.026 ***	-0.030 ***
Ego-Network Density	-0.501 ***	-0.932 ***	-0.854 ***	-0.323 ***	-0.848 ***	-0.474 ***
Tenure	-0.408 ***	-0.483 ***	-0.550 ***	-0.214 ***	-0.568 ***	-0.237 ***
Age	0.018 ***	0.040 ***	0.044 ***	0.009 ***	0.053 ***	0.011 ***
Gender	-0.080 ***	-0.126 ***	-0.093 ***	-0.048 ***	-0.078 **	-0.076 ***
Alpha		3.481 ***			0.866 ***	
Log-Likelihood	-102257	-55665	-65455		-53515	
AIC	204528	111346	130938		107061	

Dependent Variable: Active Participation (# of activities; categorized); n=30,431
**** p<.01 ** p<.05 * p<.10; NB=negative binomial;*
1) signs in the logit part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"

Table 19: Study 2 – Model Comparison for Network Sample

6.5 Results of Main Analysis

A series of three main analyses is conducted. Analysis 1 addresses the impact of network measures on user participation for the entire sample of current users. Analysis 2 complements the structural predictors by including attitudinal measures from the survey. This study should verify the findings from Analysis 1 in cases where attitudinal measures are available and also test for the impact of those attitudinal measures when network variables are taken into account. Analysis 3 tests for the moderating effect of social motivation to use the community by splitting the sample in two groups and comparing those groups.

6.5.1 Analysis 1 – Network Effects on User Participation

To evaluate the effect of social network characteristics on user participation behavior, this analysis uses the entire sample of users with at least one contact, who registered before T2 and logged-in in T2. A ZINB model is specified (Model 1), consisting of the structural (degree, betweenness centrality, and ego-network density) and control variables (tenure, age, gender), since attitudinal variables are not available for the full sample. The results in Table 20 show that degree and ego-network density have significant effects in the hypothesized direction for the binary logit-model and the NB-model. A higher degree results in a higher level of active participation. A strongly connected network of neighbors has a negative impact on participation behavior. Further, this study shows a significant, but negative effect for betweenness, which is in contrast to the hypothesized positive effect.

6.5.2 Analysis 2 – Network and Attitudinal Effects on User Participation

In Analysis 2, ZINB models are parameterized using both network and self-reported measures to estimate their effect on user activity. Again, tenure, age, and gender are included as control variables. Because attitudinal measures are only available for survey respondents, this analysis uses the reduced survey sample. Table 20 provides an overview of three ZINB models with the respective parameter estimates and significance level. Model 2a includes only the structural dimension, while Model 2b includes only the relational dimension. Model 2c is the final model and presents the combination of both dimensions. A comparison of the models based on the log-likelihood and AIC statistics show that Model 2a is superior to Model 2b, but a combined model with both structural and attitudinal predictors has a better fit than the model with structural drivers alone (lower AIC value), which underlines the value of additional self-reported measures.

In the final Model 2c, degree has a significant positive effect in both parts of the ZINB model. This suggests that the decision to contribute at all and the number of contributions are affected by the number of contacts. Betweenness does not show any significant effect. The impact of ego-network density is in the expected direction and statistically significant only in the logit-part of the model. The share of offline friends has a negative effect on participation, implying that a higher share of offline friends leads to less interaction with the community.

The attitudinal variables provide diverging results. While overall satisfaction has a significantly positive effect in both parts of the model, identification significantly impacts only the conditional number of activities. Last, reciprocity shows no significant effect on user participation. An interesting finding is that some effects of Model 2b change in Model 2c. When controlling for the network dimension, satisfaction becomes significant in the NB-part of the model, while reciprocity becomes insignificant in the logit-part of the model. In addition, the effect of tenure changes signs when network measures are included in the model. This leads to the presumption that network and attitudinal variables affect each other to some degree. Although this study does not explicitly test for these relationships, these results are discussed in the discussion and limitations section.

In summary, the findings from Analyses 1 and 2 confirm the hypotheses for degree (H1), share of real-world friends (H4) and satisfaction (H7). Ego-network density (H3) and identification (H5) can be partly confirmed, as only one part of the two-stage models show consistently significant effects. Betweenness (H2) is only significant in Model 1, showing a negative effect. Reciprocity does not show a significant impact, thus rejecting hypothesis (H6). When controlling for network structure, tenure shows a significant negative effect, meaning that users with lower tenure participate more actively. At the same time, older users are also more active. Although gender is only significant in Model 1, the results suggest that male users contribute more.

	Model 1[1)]		Model 2a[2)]		Model 2b[2)]		Model 2c[2)]		Hypo-theses
	Logit[3)]	Neg.Bin.	Logit	Neg.Bin.	Logit	Neg.Bin.	Logit	Neg.Bin.	
Intercept	-2.137 ***	1.133 ***	-1.795 ***	1.468 ***	-1.939 ***	1.062 **	-2.970 ***	0.991 ***	-
Degree	0.829 ***	0.381 ***	0.901 ***	0.246 ***	-	-	0.800 ***	0.185 ***	supp.
Betweenness	-0.026 ***	-0.030 ***	0.012	-0.008	-	-	0.035	0.000	n.s.
Ego-Network Density	-0.848 ***	-0.474 ***	-1.676 ***	-0.394	-	-	-1.163 **	-0.355	partly supp.
Share Real Friends	-	-	-	-	-	-	-1.490 ***	-0.414 ***	supp.
Identification	-	-	-	-	-0.058	0.134 ***	0.080	0.117 ***	partly supp.
Reciprocity	-	-	-	-	0.282 ***	0.028	0.124	-0.007	n.s.
Overall Satisfaction	-	-	-	-	0.235 **	0.058	0.275 **	0.067 *	supp.
Tenure	-0.568 ***	-0.237 ***	-0.573 ***	-0.120 ***	0.163 *	0.073 **	-0.482 ***	-0.083 **	-
Age	0.053 ***	0.011 ***	0.064 ***	0.015 ***	0.030 *	0.010 **	0.049 **	0.011 **	-
Gender	-0.078 **	-0.076 ***	-0.290	-0.126	-0.157	-0.019	-0.198	-0.061	-
Alpha	0.8663 ***		0.3365 ***		0.4184 ***		0.2820 ***		
Log-Likelihood	-53515		-1207		-1280		-1177		
AIC	107061		2445		2590		2400		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized)
**** p<.01 ** p<.05 * p<.10; supp.=supported; n.s.=not supported; neg.eff.=negative effect*
1) n=30,431; 2) n=486
3) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros). Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"

Table 20: Study 2 – Results of ZINB Models for Analyses 1 and 2

6.5.3 Analysis 3 – Social Motivation and the Effects on User Participation

In order to better understand the influencing factors on user behavior, two relevant groups of users are compared: Networkers and Non-Networkers. Specifically in online social communities with both editorial and user-generated content, not all users are highly driven by the need for interconnection with other users, as the consumption of information or entertainment can also provide benefits to use the platform. The group of Networkers is defined as those users, who state at least an average of four points on the 7-point social interaction value scale. Non-Networkers are defined with less than four points. By doing this, 260 users are assigned to the Non-Networker group, 226 users to the Networker group. To estimate the differences in the structural and relational dimensions (H8), separate regression models with degree, betweenness, ego-network density, identification, reciprocity, and satisfaction as the dependent variables are used. The user groups are represented by a binary variable. Age, gender, and tenure are included as control variables. For the analysis ordinary least squares is used to estimate the coefficients and compute heteroskedasticity-consistent standard errors (Breusch-Pagan test, p<.05). For all hypothesized variables except ego-network density, the regression analyses reveal a significant difference (p<.05), supporting hypothesis H8. Table 21 presents the results of the regressions for user group comparison.

	Mean Scores		OLS-Regression Results for Binary Group-Variable[1)]		
Dependent Variables	Networker (n=226)	Non-Networker (n=260)	Beta	T-Value	P-Value
Degree[2)]	3.68	2.54	1.10	8.20	.000
Betweenness[2)]	13.43	10.18	3.12	6.48	.000
Ego-Network Density	0.20	0.24	-0.03	-1.60	.110
Identification	4.13	2.69	1.43	11.28	.000
Reciprocity	3.84	2.43	1.38	10.25	.000
Overall Satisfaction	5.00	4.35	0.66	6.33	.000

1) Statistics provided for the effect of the binary group variable (Networker=1; Non-Networker=0) on the dependent variables;
Control variables: age, gender, tenure; binary Variable: estimated with robust standard errors
2) log-transformed measure

Table 21: Study 2 – Comparison of Networkers and Non-Networkers

Further, two independent ZINB models are specified, one for each user group. They include the same predictor variables as in the final Model 2c and estimate the effects on active participation. The two models are presented in Table 22. The absolute log-likelihood (1177) and the AIC (2400) of Model 2c are larger than the sum of the log-likelihoods (1135) and AICs (2362) of the two separate models. Because the difference in the log-likelihoods is significant (LR-test: $chi^2(23)=83.78$; p<.01) and models with lower AIC are preferred (Cameron and Trivedi 1998), a separation into two models is supported. To test hypothesis H9, the significance levels of the predictors' effects in the two models are assessed. Overall, the results show that for Networkers all independent variables, with the exception of betweenness, have a significant effect on the conditional number of contributions. Interestingly, reciprocity has a negative effect, which will be addressed in the discussion. The only main variables significant in the binary model are degree and share of offline friends. For Non-Networkers, degree, ego-network density, satisfaction, reciprocity, and share of real-world friends are significant in the binary model, while only identification is significant in the NB-part of the model. This provides evidence, that for Networkers the conditional number of contributions is more influenced by the predictors than their decision to contribute at all. For the Non-Networkers, basically the effects are more important for their decision to contribute, but less for the number of contributions they make. Although both user groups do not differ in all aspects, in general hypotheses H9a and b are supported for most of the predictors.

	Model 2c[2)]		**Non-Networkers**[2)]		**Networkers**[3)]		**Difference**[4)]	
	Logit[5)]	**NB**	**Logit**	**NB**	**Logit**	**NB**	**Logit**	**NB**
Intercept	-2.970 ***	0.991 ***	-2.508 **	-0.276	-2.879 *	2.097 ***	-	-
Degree	0.800 ***	0.185 ***	1.073 ***	0.108	0.716 ***	0.183 ***	n.s.	supp.
Betweenness	0.035	0.000	0.012	0.042	0.018	-0.016	n.s.	n.s.
Ego-Network Density	-1.163 **	-0.355	-2.178 ***	-0.347	0.985	-0.540 **	supp.	supp
Share Real Friends	-1.490 ***	-0.414 ***	-1.044 *	-0.378	-2.686 ***	-0.394 ***	n.s.	supp.
Identification	0.080	0.117 ***	-0.072	0.139 *	0.182	0.054 *	n.s.	n.s.
Reciprocity	0.124	-0.007	0.325 **	0.035	-0.103	-0.048 **	supp.	supp.
Overall Satisfaction	0.275 **	0.067 *	0.373 **	0.082	0.262	0.056 *	supp.	supp.
Tenure	-0.482 ***	-0.083 **	-0.620 ***	0.002	-0.457 ***	-0.138 ***	-	-
Age	0.049 **	0.011 **	0.026	0.022 *	0.074 **	0.006 *	-	-
Gender	-0.198	-0.061	-0.916 **	0.205	0.407	-0.169 **	-	-
Alpha	0.2820 ***		0.5793 ***		0.1021 ***			
Log-Likelihood	-1177		-503		-632			
AIC	2400		1052		1310			

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized)
***** p<.01 ** p<.05 * p<.10; NB=negative binomial; supp.=supported; n.s.=not supported*
1) n=486; 2) n=260; 3) n=226
4) Differences between networkers and non-networkers are defined as one group showing a significant effect of an independent variable on participation, while the other groups does not reveal a significant effect of the same variable on participation for the same model (either logit or negative binomial model)
5) signs in the logit-part of the model are reversed to indicate the effect on “contributions” instead of “not contributing” (certain zeros). Positive effects are thus interpreted as “the greater the independent variable, the more likely the users were to contribute”

Table 22: Study 2 – Results of ZINB Models for Networkers and Non-Networkers

6.5.4 Verification of Results

To verify the results from Analysis 1, the analyses are conducted for a second observation period, which describes the network of active users in T0 (a timeframe of two month before T1) and the participation from T1. The results obtained from this analysis are similar to the ones in Analysis 1, thereby confirming the findings from Analysis 1 (see Appendix 9).

In a second test, the models are estimated using different transformations of the dependent variable. As the choice of the number of categories used to account for extreme values and high variance is rather arbitrary, different categorizations of the original data are applied to represent different aggregation levels for the count data variable (see Appendix 10 for the different intervals used). Overall, the results for the different transformations of the dependent variable are almost identical. Thus, the above findings from Models 1 and 2c are confirmed. The analyses for Non-Networkers and Networkers with different transformations of user participation also yielded similar results.[63] (see Appendix 11)

Although degree and betweenness centrality of an individual's network are conceptually distinct, they are often highly correlated (e.g., Mehra, Kilduff, and Brass 2001). Therefore, it is tested for the collinearity of degree and betweenness by examining the variance inflation factors (VIF) associated with all predictors, where VIF scores greater than 10 are indicative for multicollinearity issues (Field 2005). None of the variables violated this criterion (highest VIF=3.41 in survey sample, highest VIF=3.52 in network sample), thus multicollinearity should pose no serious threat to the validity of the analyses. Nevertheless, to verify the results, Models 1 and 2 are estimated with the untransformed betweenness centrality measure. On this way the correlation with degree is much lower in both models ($r<.22$), but this leads also to a higher variance and exposure to outlier effects, which was aimed to be avoided in the main analysis. The results are similar to the findings above, as most effects are confirmed. Though, the effect of betweenness becomes significantly positive in the logit-part of all three models, providing partly support for hypothesis H2. However, this effect might be attributed to extreme values and is therefore not considered robust. Additionally, in Model 2c ego-network density becomes insignificant in the logit-part, but in Model 1 and 2a stays significant, thus still supporting hypothesis H3. The effects of the remaining covariates in all models stay the same. (see Appendix 12)

In a last test to verify the results of Model 2c, the activity of a user's contacts is included as a further control variable. Thereby, active participation of friends shows a significant effect in both parts of the ZINB model. This suggests that higher activity of a user's contacts stimulates the participation of the focal user, because more information is available and more reaction to others' activities is expected to occur. Most effects of the independent variables remain significant despite controlling for friends' activity. Nevertheless, the inclusion of the friends' participation leads to an insignificant effect of degree in the logit-part of the model.

[63] Only two effects are not predominantly the same: for Non-Networkers, the share of real-world friends did not show a significant effect in most alternative models in the logit-part; for Networkers the significant effect of satisfaction in the NB-part of the model could not be supported by three models with alternative transformations of activity. Nevertheless, in both cases the results from Study 3 were supported by at least one other alternative model. The overall tendency still holds, that for Networkers the effect on the conditional number of activities and that for Non-Networkers the effect on the decision to participate is more important.

This is not surprising, as a higher degree is also most likely associated with a higher total friends' activity, thus suppressing the effect of degree in this case. (see Appendix 13)

6.6 Discussion

Although recent community research used objective network measures to explain adoption behavior and overall community success, this study is the first to provide a comprehensive investigation of structural and relational effects on active user participation. In three analytical steps, the impact of structure and attitudes is demonstrated.

Structural Dimension. From a social capital perspective, human behavior can be affected by the social network of an individual. The findings support the importance of social structures for behavior. In line with social theories, structural capital is greater when many direct contacts provide access to potentially more information, thus leading to a positive effect of a user's degree on participation. At the same time, a high ego-network density, meaning that the ego's direct contacts are well connected, leads to lower user activity. This is consistent with Granovetter's (1973) idea of many strong ties and Burt's (2000) concept of network closure leading to redundant information in dense networks.

Surprisingly, betweenness does not consistently show significant effects across all studies, and reveals even negative effects in the full network sample when controlling for degree (see Analysis 1). One reason might be the assumption of using only "shortest paths" inherent in the definition of betweenness centrality. For example, if second or third level neighborhoods are more densely connected, degree and ego-network density ceteris paribus, betweenness centrality is lower as the number of geodesics is reduced. Nevertheless, more information could flow through the focal actor, as more paths are available via which new information from more distant circles of users can reach the ego. Local betweenness might be a more adequate measure in this context, i.e. being in a central position between one's direct neighbors, as too distant actors might be irrelevant (e.g., Katona, Zubcsek, and Sarvary 2011).

Further, the share of real-world friends among one's online contacts is negatively related to user participation. This supports the argument that if the offline channel is available to access social capital, it will be preferred over the online channel. This means that there is less need to use the online channel for interactions, if information and social support can be reached offline. Because online interactions are inferior to face-to-face interactions the relationship of offline and online is negative (Wellman et al 2001; Cummings, Butler, and Kraut 2002). Thus, online interaction is supplemented and not complemented by offline interaction. On the other hand, this suggests that online social community bear greater social capital when the user is connected to people beyond his offline network. Thereby, contact to more distant ties creates value for the user in the online network.

Relational Dimension. Nahapiet and Ghoshal (1998) propose a distinction of the relational and structural dimension, because they describe different facets of social capital. This study provides support for this view. Attitudinal factors can significantly improve the understanding of what makes users contribute in online communities. The results suggest, that identification with the own group of contacts has a positive influence on the conditional number of activities, but not on the decision to participate at all. Users who strongly identify with their group contribute more to the community, which is consistent with other community studies (e.g., Algesheimer, Dholakia, Herrmann 2005; Dholakia, Bagozzi, and Pearo 2004). Further, the results show that more satisfied users are more likely to participate and show higher levels of activity when controlling for the structural dimension. However, without controlling for a user's network effects, satisfaction does not show significant effects on the conditional number of contributions, which leads to the presumption of interaction effects with the structural dimension. Contrary to former research, reciprocity does not significantly affect user behavior in Model 2c. As reciprocity depends on the imputed value of the benefit received (Gouldner 1960), the effect of reciprocity might be undermined if the value is not perceived as sufficient. Further, in Model 2b (Table 20) reciprocity is significant in the binary part of the model, but not in the final model. This means that its effect is suppressed in Model 2c by the structural variables. Nevertheless, when accounting for the level of social motivation, Analysis 3 reveals that reciprocity has an impact for different user groups on participation behavior, which emphasizes the need for user segmentation.

Tenure. Higher tenure does not lead to higher participation. This is in contrast to former research (e.g., Wasko and Faraj 2005; Nambisan and Baron 2007), and the view that more experienced members contribute more. Though, when not controlling for network structure tenure is positively significant. A longer membership correlates with the number of friends, meaning that experienced members have more friends and thus contribute more. But, given the same number of friends, new users are more active than experienced ones.

Social Motivation. The study shows that different user groups – Networkers and Non-Networkers – significantly differ in their network position and attitudes. Users motivated by social needs are more satisfied, can identify more with the community, show a stronger norm to reciprocate, and are in more advantageous positions in the network. Further, the differentiation of users by their social needs demonstrates that different types of users are affected differently in their behavior. For Networkers, all predictors except betweenness are significant influencers of the number of contributions they make. Interestingly, reciprocity shows a significantly negative effect. This could be explained that if the reciprocity expectations are not met in the past, users have either lower norms of reciprocity today, but nevertheless participate, or they still have higher norms of reciprocity, but participate less because they fear that their participation is not reciprocated. In contrast, the decision of whether to contribute at all is only affected by degree and share of real-world friends. One reason for this observation could be that the social needs already explain the decision to participate at all for most of the users. On the other side, Non-Networkers use the community platform to a lesser extent out of social needs, but might be driven more by functional needs. For them the information and editorial content provided on the website might be sufficient to satisfy those needs. Therefore, they are more affected by factors triggering to contribute at all. Although, not driven by high social needs, Non-Networkers are also affected by structural capital in form of degree, ego-network density and share of real-world friends in their decision of whether to participate at all. This is interesting, as structural capital drives behavior, although those users are less interested in social interaction. For example, those users could be needed as connectors to distribute information in their networks, which would "force" the user to be involved. Further, for Non-Networkers identification is specifically important, as it is the only factor influencing the number of contributions.

Key Findings. Overall, these results provide three important theoretical implications. First, the study provides support for the impact of an individual's network configuration on user behavior. Specifically, degree, ego-network density, and the share of real-world friends are important influencing factors of user activity. This contributes to existing research in showing that not only the number of ties is important, but also how the group of online contacts is composed. Second, the relational dimension provides additional explanation of user participation and holds although objective data is used. Even though the structural and relational dimension of social capital can be distinguished with regards to their facets, the analyses provide basis for the presumption of important interaction effects between structural and relational constructs, which needs further investigation. Third, customer segmentation helps to better understand social capital drivers. Social capital affects different stages of user participation for Networkers and Non-Networkers, and some effects only become visible when looking at specific user groups as they are differently affected.

6.7 Managerial Implications

From a managerial perspective specific attention should be paid to promote the connection with other users, provide a basis for a stronger relationship to the community, and customize content and functionality to the needs of specific user groups.

Community operators should use the objective information on the users' individual network configuration and try to shape their network in a way to increase user participation. It is important to facilitate the establishment of ties to other users in a way to open up new information sources and connecting otherwise not connected users. Especially users with few friends should be promoted to connect with others to establish a basis for interaction, for example through automatic suggestions of whom they may know. Recommending ties to distant users, i.e. users who do not know each other from the offline world, is even more effective. As groups may become more cliquish over time, meaning that one's neighbors connect to each other, streams of new members to these sub-groups are important to refresh the interest and get access to new interactions. Thus, the community system may not focus to recommend close users from the same region, but specifically users from other cities and regions with similar interests, for example users who visited the same offline events or forums on the community site. Another way to access new streams of content is to reconnect with existing distant contacts, who are not in the close circle of friends. First, the community system could simply suggest reconnecting with specific users. Second, as many systems provide some kind of newsfeed of the friends' activities, weak tie contacts could be prioritized in the news listings to facilitate reconnection.

The effects of the attitudinal factors suggest that a better identification with the community facilitates higher user participation. Thematic online communities could offer local events and activities, like brand fests in brand communities, to achieve a sense of belonging and to provide a rich social context which facilitates the process of socialization. Those kinds of events should be carefully offered, as offline friends supplement for online activities, therefore the focus must be to bring different circles of friends together, not already closely connected clusters. Another way to increase identification with the community online would be dedicated thematic groups to meet like-minded people. These groups should be kept at a reasonable size, because large groups could become anonymous. Further, it is important to keep the expectation of the users and satisfy their needs. Especially for communities providing editorial content, quality and quantity of the content are key factors.

The study demonstrates that different user groups are differently affected in their behavior. Therefore, it is crucial for community operators to understand the different needs of their users. One simple approach to gain such information would be a mini-survey after registration and recurring every few months to know what the current needs of the users are. Two or three simple questions should evaluate social, functional and other needs relevant for the respective type of community. With this knowledge, operators can better personalize content on the landing page after logging-in, for example providing articles, photos and videos for users with information needs, and news about last activity of friends for socially driven users. Consequently, Non-Networkers should be provided with relevant content and recommendations to meet other users online and offline. Especially, if those users have not participated yet, introduction to active users and groups of interest could help to make them “a little” active and connect with others, which could potentially turn them into Networkers in the future. For Networkers it is specifically appropriate to connect and reconnect with distant and new users. When Networkers decided to participate, the number of contributions is influenced by almost all observed factors. This suggests that community operators must continuously stimulate the activity of Networkers by providing a steady stream of new information or anchors for discussion and interaction. Therefore on their landing page more focus would be set on displaying friends’ activities and prioritizing the recommendation to connect to new users and users from other regions with similar interests. The need for new users is also underlined by the fact, that users with lower tenure contribute more (all other things being equal). Therefore, new users help to increase directly the activity. Further, more experienced users could be stimulated to be more active by these new information sources.

6.8 Limitations and Future Research

Although the results are insightful and have interesting theoretical and managerial implications, this research comes with limitations.

Overall structural positions are included in the analysis. Though, dyadic factors between the focal actor and other users like homophily or tie strength were not investigated. Identification was used as a proxy to include tie strength to the group as a whole. But tie strength to single members could give additional insights on customer activity related to specific other members in the network and not to the community as a whole. Therefore, a more dyadic view on the influencing factors could help to understand micro level dynamics.

In many online communities a wide range of activities is offered. Some users might use online communities for specific activities like communication with others, while other users are more interested in playing games or watching videos and photos. Depending on the community's theme, a differentiation of user activities could be useful in order to understand what influences which user activity.

Further, tenure is used as an indicator for the cognitive dimension of social capital because of the unavailability of more detailed data in this context. Therefore, considering other factors that reflect the cognitive dimension more explicitly, like shared language and stories in the online community, could help to gain further insights on this dimension.

As demonstrated in Analysis 2, it is suspected that there is a relationship between structural and attitudinal measures. For example, satisfaction and reciprocity could be related to a user's degree. Users with many friends could be more satisfied with the community, as there is more interaction with other users. To test such hypotheses, further investigation of the relationships between structural and relational dimensions and the direction of such effects are recommended.

7 User Retention: The Effects of Social Structure and Engagement on Defection from Online Social Communities[64]

In the last chapter, a better understanding of the influencing factors of active user participation was provided. For community operators, the activation of users is an important goal, which helps to make the online community more attractive and sustainable. But not all members actively contribute and stay forever on the platform. The user's life cycle ends with the decision to leave to community. As pointed out in chapter 2.4.3 a critical step in the membership development process is to build loyalty and keep a sufficient number of users on the platform in order to capture the value of the community. Therefore, community operators need to understand why their users leave the online social community in order to develop measures to build a long-term relationship with their members. The aim of this section is to address the question of what influences user defection in online social communities.

7.1 The Importance of User Retention in Online Social Communities and the Need to Understand User Defection

Though there has been enormous growth of the online community market in the past years, this development is slowing down. While some online social communities are still able to grow their user base, many smaller and more regional online social communities suffer from decreasing numbers of members. For example, in the German online social community market some communities lost up to 25% of their users in 2011 compared to 2010 (Schneller 2011). The trend also shows some social media fatigue among certain user groups of such sites, where users get bored of their social network and use the community less than when they first signed up (Gartner 2011). As the online community market becomes more saturated it is more important to retain existing users. Traditionally, customer retention is an important theme for marketing managers, because retention is often seen as less expensive than acquisition of new customers, and it can also yield an increase in profits for the firm (e.g., Reichheld

[64] This study was conducted jointly with my colleague Christine Geser, Chair of Services and Technology Marketing, TUM School of Management, Technische Universität München.

and Sasser 1990; Rosenberg and Czepiel 1984). Therefore, companies try to anticipate customer defection and retain the relationship to their customers. For online community operators it is even more important to keep a critical mass of members using the platform, in order to remain attractive for existing and prospective users. Once users are registered on the platform, user retention should be a major objective. Knowledge about what influences the members in their decision to stop using the community and which users are more likely to defect should help to develop measures to strengthen the relationship of the users to the community.

In marketing, the impact of different factors on customer churn has been investigated from many different perspectives. Thereby, past research explains customer attrition by personal characteristics (e.g., Bhattacharya 1998), the customers' attitudes and perceptions (e.g., Bolton 1998), and specific customer behavior (e.g., Dover and Murthi 2006; van den Poel and Lariviere 2004). Although the network perspective plays an increasing role in marketing (e.g., Achrol and Kotler 1999; Algesheimer and Wangenheim 2006), the effect of the social context of the customer on his defection behavior has received only limited attention so far (e.g., Nitzan and Libai 2011). In the context of a telecommunications provider, Nitzan and Libai (2011) provided valuable insights on how defecting contacts of customers affect the defection of themselves. They found that a higher number of defected neighbors also affect one's own defection decision. Also in other research areas, it has been found that individuals are affected in their leaving behavior by other people, such as when quitting smoking (Christakis and Fowler 2008) or leaving employment in a firm (Castilla 2005).

As users connect and interact with each other in online communities, their structural capital in terms of the users' location within the network is expected to influence their usage behavior (Tsai and Ghoshal 1998). Because of the nature of online social communities, network effects and social influence have been regarded specific attention (Dholakia, Bagozzi, and Pearo 2004; Wasko and Faraj 2005). In particular, recent research incorporated social structural aspects, for example, to explain adoption behavior (Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011), the influence of a user's log-in behavior on other user's activity (Trusov, Bodapati, and Bucklin 2010), or the success of open source software communities on a community level (Toral et al. 2009). In study 2 (chapter 6), it was already pointed out that there is a lack of knowledge about how the users' position in the network and the overall network structure influence user contribution in the online social community. In addition, the effect of the social network structure and the position of the users in the network on user defection have not yet been studied. Because users are connected to different people in the online community, it is expected that their social environment has a direct effect on the decision to stay with the community or to leave. This is of considerable interest for the online community operator, because the knowledge of how structure influences defection can help to shape the community in a way to reduce user defection. This is the first study that employs a user's position in the online network, the configuration of his close network of contacts, and his engagement in the community to explain the likelihood of defection from an online social community.

This study contributes to marketing and online community literature in several ways. First, the focus is to understand the effect of the users' social connections on their decision to leave the community. Thereby, social network measures in terms of a user's centrality and ego-network density are used to explain defection behavior, which have not been studied in this context. Further, the effects of the composition of the users' social networks are investigated, concentrating on the focal user's similarity with his contacts' and the share of defected users among one's contacts. Although network measures have been used in the past to study user behavior in online communities (e.g., Goldenberg et al. 2009; Katona, Zubcsek, and Sarvary 2011; Trusov, Bodapati, and Bucklin 2010; Wasko and Faraj 2005), there is no knowledge about their impact on the users' propensity to leave the community. Second, changing effects over time are investigated for all variables. The study reveals that for some influencing factors, their impact decreases over time. This is important for the timing of specific operator initiated measures to retain their users. Third, the study is conducted with a unique data set. By using the data from a free online community service with a closed network of users (see chapter 4 for the description of the research object), this study provides a valuable context for observing the effect of the social context on defection. Longitudinal data of two years is used to follow a cohort of users from their registration to their defection or end of the observation period. Thereby, time-varying covariates are employed in a survival analysis, which provides a more accurate estimation of the effects on defection.

7.2 Theoretical Background and Hypotheses

7.2.1 The Social Context of Online Communities and User Retention

For community operators, an important question is why their users continue or stop using the platform. Because online community members interact with other members and the community as a whole, the impact of the social-structural context of online communities on member behavior should be accorded explicit recognition. Social network theory posits that the social network of an individual affects human behavior. With regards to user retention, one important aspect to keep users on the platform is the benefits they gain from the relationship to the community. Thereby, it has been suggested that individuals choose to keep social contacts because of the existence of social capital, which emphasizes the value of social structures for social action (e.g., Bourdieu 1986; Burt 1992; Coleman 1988; for a more detailed introduction to social networks and social capital, see chapter 3.2 and 3.3). An important aspect of social capital is that investment in social relations results in benefits. Benefits can be realized on an individual and collective level and they include functional benefits from exchanging information and knowledge as well as solidarity benefits and social support provided by other individuals or the community as a whole (e.g., Adler and Kwon 2002; Burt 1997; Coleman 1988; Matthwick, Wiertz, and de Ruyter 2008). Therefore, the value of using an online social community lies in the opportunity to access resources and interact with other people, for example through communication, sharing ideas, helping each other, or exchanging digital goods such as photos and videos.

Compared to human capital, which is embodied in the skills and knowledge of an individual, social capital is not owned by an individual, but is jointly owned by the community (Burt 1992). Social capital is inherent in the social structure and offers only advantages through the relationships with other users. If the relationship to the community or other members is absent, the social capital is lost. Therefore, leaving the community would result in a loss of capital for the defecting individual. When a member leaves, all connections are cut to the community. Because social capital facilitates the exchange and combination of resources and information (Nahapiet and Ghoshal 1998), it is important to keep connected to ensure further access to the social capital in the community. If the capital that an individual member can gain from social interaction with the online community is large enough, the member should continue using the site. Hence, when his social benefits are higher than the costs, the user would stay. Alternatively, if there are high opportunity costs when leaving, meaning that one would lose a lot of social capital and benefits, the relationship would be maintained. Consequently, the user would evaluate the benefits that can be achieved in the online social community depending on the connections to other users. The composition of the user's network and the relationships to one's contacts play an important role, which is expected to influence the user's decision to leave.

In the following, hypotheses are developed on the basis of social theories about the effect of a user's social structure and engagement in the community on the individual's likelihood to defect.

7.2.2 Ego-Network Structure

Centrality. In social theories the structure and configuration of the individuals' networks play an important role for their behavior. Valuable information flows can arise from network ties, which provide access to resources and information (Nahapiet and Ghoshal 1998). Here, centrality is an adequate concept to measure a user's position in the network. Degree and betweenness centrality are simple to calculate and the most commonly used centrality measures (e.g., Freeman 1978/79; Wasserman and Faust 1994).[65]

As pointed out in chapter 3.2.2, degree centrality (or the degree) of a user describes the number of contacts or 'friends' to which the focal user is connected (Freeman 1978/79). A person who is highly connected to others acts as an intersection in the sense that a large amount of information can pass through this person. This means that users with a high degree centrality are more involved in the network than persons with a low degree (Freeman 1978/79). Not only direct contacts, but also indirect, second-level contacts, i.e. the contacts of the focal user's direct contacts, can help to increase the flow of valuable information to the focal user. The reason is that through their contacts the focal users potentially gain access to more information sources. The more indirect contacts one has, the more exposed to knowledge beyond one's own friendship circle (Granovetter 1973). Access to information would lead to higher exchange of resources (Nahapiet and Ghoshal 1998), resulting in a higher future exchange and loyalty of the user to his network. If social capital is created in an organization or community, through community specific ties, the likelihood of remaining in the community would be higher (Dess and Shaw 2001). Overall, when more sources of information and support are available to the focal users, they should stay longer in the community as they want to keep these ties that provide functional and emotional benefits.

Betweenness centrality is a second measure of the user's central location in the online social community. Being in a central position, i.e. lying on a high number of shortest paths between all other pairs of actors in the network (de Nooy, Mrvar, and Batagelj 2005; Wasserman and Faust 1994), increases the potential access to relevant information. This is because actors, who take on brokerage roles between other actors, constitute a bridge between sub-groups, so that non-redundant communication and information can flow through this actor (e.g., Burt 2000). Central actors are important intermediaries in the communication network, i.e. they are involved in the interaction of the other actors (de Nooy, Mrvar, and

[65] A more detailed description of network measures and an overview of the mathematical definitions of centrality can be found in chapter 3.2. This study focuses on degree and betweenness centrality, because closeness centrality cannot be calculated in weakly connected networks, which is the case for the empirical research object.

Batagelj 2005; Freeman 1978/79). Being more involved in the activity of other users and accessing relevant resources increases the value a user can gain from the online community. Because social capital lies in the connection to other members (Nahapiet and Ghoshal 1998), users with more ties should gain higher social capital benefits. Higher benefits make the network more attractive for the user. This should lead to a lower willingness to leave the online community, because breaking up with sources of information and social support would result in losing social capital. In a study on adoption behavior, Katona, Zubcsek, and Sarvary (2011) investigated the influence of network effects on the diffusion process in an online community. They find that people adopt more likely if the number of their friends who already adopted is high. Accordingly, it is proposed here that a high number of contacts who use the online community should influence the focal user to continue using the service, and thus reduce the likelihood to defect. This leads to the following hypotheses:

H1: Users with a higher degree centrality are less likely to defect.

H2: Users whose contacts have a higher average degree are less likely to defect.

H3: Users with higher betweenness centrality are less likely to defect.

Ego-Network Density. According to Burt (2000), density is often discussed as network closure and describes how closely connected all contacts are to one another. Here, ego-network density is the extent to which a community member's (online) contacts are connected to each other.[66] From a social capital perspective, network closure facilitates the access to information and sanctions that make it less risky for people in the network to trust one another (Burt 2000; Coleman 1988; Nahapiet and Ghoshal 1998). Although network closure increases the probability of redundant information and decreases the opportunity to broker (Burt 2000), the establishment of close-knit subgroups is associated with stronger ties among friends (Granovetter 1973). When a user's friends are strongly connected, emotional intensity, companionship, interaction, and a sense of mutuality emerge (Granovetter 1973; Walker, Wasserman, and Wellman 1993). A family-like environment with strong attachment arises, where users would not be likely to break up the relationship with their community. Therefore, closure enables the development of norms, trust, and identities (Coleman 1988), which in turn leads to more engagement and loyalty (e.g., Algesheimer, Dholakia and Herrmann 2005;

[66] The density of the ego's network describes the number of relationships between direct neighbors of the focal actor in relation to the maximum number of relations between these neighbors (de Nooy, Mrvar, and Batagelj 2005).

Dholakia, Bagozzi, and Pearo 2004). In an organizational setting, Cappelli (2000) argues that the more employees are interacting and building relationships between colleagues, the more loyal these employees become towards the company they are working for. Thus, it is hypothesized:

H4: Users with a higher ego-network density are less likely to defect.

Similarity. In network theory, similarity or homophily refers to the degree to which individuals are similar in terms of certain attributes, such as demographics, social status, or lifestyle (McPherson, Smith-Lovin, and Cook 2001). In fact, online social communities already build upon the similarity of people in certain aspects, as they often center around a specific topic and bring together like-minded people. However, within the online community, people differ. There is a tendency that socially similar people connect with each other more often and more closely than people who are dissimilar (Granovetter 1973; McPherson, Smith-Lovin, and Cook 2001). When individuals are similar to other members, they are more likely to develop greater levels of interpersonal attraction, companionship, trust, and understanding with these contacts (Ruef, Aldrich, and Carter 2003). Subsequently, similar users experience higher levels of social capital inherent in their relationships with the community, because social capital also emerges from obligations, identification and trust to other users (Nahapiet and Ghoshal 1998). Higher similarity, indicating increased social capital benefits from these relationships, would therefore decrease the potential of users to leave the online community. Further, as similar people more often connect to each other, they would in turn be less likely to disconnect. It has been found that higher similarity between sender and receiver of information has positive effects on the influence of the senders (Evans 1963; Gilly et al. 1998; Wangenheim and Bayón 2004b). In addition, a study based on telecommunications data revealed that similarity with defecting neighbors increases the likelihood of the focal customer's defection (Nitzan and Libai 2011). This supports that higher similarity to one's active contacts should decrease the risk of leaving the community.

Similarity between people can be measured along various dimensions (McPherson, Smith-Lovin, and Cook 2001). Locally organized online communities have the goal to connect with other users and meet new people from the same and different regions. Because users may prefer to get information on leisure time activities (as the topic of the online social community studied here) from people who are like themselves, gender and geographic proximity are chosen as relevant factors to observe similarity. Consequently, it is hypothesized:

H5a: Users with a higher share of contacts from the same geographical region are less likely to defect.

H5b: Users whose contacts are from regions that are on average closer to the own region are less likely to defect.

H5c: Users with a higher share of contacts of the same gender are less likely to defect.

Share of Defected Contacts. Not only the user's current contacts, but also the group of defected contacts, are expected to have an impact on the focal user's decision to leave the community or not. If many members of one's circle of friends leave, more information sources are lost. Less access to information, through direct and indirect contacts, would lead to lower exchange of resources (Nahapiet and Ghoshal 1998). If social capital, which is inherent in the connections to other members, leads to a lower likelihood of defection, the opposite should happen when the number of defected friends is higher, i.e. the risk of leaving increases.

In addition, collective behavior of leaving the platform influences the focal user's own leaving behavior. According to Granovetter (1978), collective behavior occurs in situations where people have to decide between two options and their decision is dependent on the number of other individuals that have already chosen a certain option (see also chapter 3.4.4 for a brief introduction to collective behavior models). In this case, the user has the choice of leaving or staying in the community, where the decision is based on the costs and/or benefits to do so and depends in part on how many others make which choice. At some point, when the proportion of people who take one choice is large enough, the perceived benefits to the individual of doing the thing in question exceed their perceived cost. Therefore, with many members leaving, the benefits of staying are lower than the costs at some point, and the user would decide to leave as well. In online communities, the individual's social structure is of particular relevance, because the number of contacts who leave the community is weighted much higher by the focal user than the defection of any other member in the entire community (Granovetter 1978). Consequently, the users' decision to defect will depend in part on the proportion of the users' contacts, who have already left the community. Because the benefits of remaining in the community decrease with a higher share of defected contacts, more users would decide to leave the community when this share increases.[67] This kind of social influence has been demonstrated to occur in various situations. For example, Christakis and Fowler (2008) found that the quitting behavior of smokers is positively influenced by the smokers' friends, who have already stopped smoking. In a telecommunications context, Nitzan and Libai (2011) demonstrated the positive effect of the number of defecting neighbors on the focal customer's defection. Thus, it is hypothesized:

H6: Users with a higher share of defected contacts are more likely to defect.

7.2.3 Community Engagement

Active User Participation. As online communities subsist on user contributions and exchange of information, active user participation plays a central role in the community context (e.g., Hagel and Armstrong 2006; Rheingold 1993). In an online social network, interaction with the community and other users describes the relationship to the community. A high quantity of interaction is related to stronger relationships, because its strength can be described as the amount of time spent and frequency of contact (Granovetter 1973). Therefore, higher participation in the community increases the overall strength of the relationship to the community, as more information has been exchanged with other users. The social capital in the community increases because participation fosters relationships (Nahapiet and Ghoshal 1998). People who have strong interpersonal relationships are less likely to terminate the relationship to an institution or organization (Dess and Shaw 2001). As a result, higher engagement in the online community leads to increased commitment and loyalty to the community (e.g., Algesheimer, Dholakia and Herrmann 2005; Woisetschlaeger, Hartleb and Blut 2008).

In addition, user participation can serve as an indicator for customer satisfaction (Bolton and Lemon 1999; Nitzan and Libai 2011). Customer satisfaction is a function of expectations and disconfirmation (Oliver 1997). Expectations concerning a product or service work as a reference point with which the customers compare the outcome of the product or service. When the product or service exceeds customers' expectations they will be satisfied. As expectations also play an important role in the relation to ones network (Nahapiet and Ghoshal 1998), they are likely to influence the users' motivation to also participate on the platform in

[67] Granovetter (1978) describes threshold models to explain at which point such a decision would occur. The decision is very complex involving different individual benefits and costs levels and also incorporating the quality of relationships, which is out of the scope of this study to elaborate on. Nevertheless, on an overall basis, a higher share of inactive friends should be positively related with leaving the community.

the future. Following Bolton and Lemon (1999), it is assumed that a decrease in the users' participation level over time indicates dissatisfaction. In a study conducted in the context of an interactive television entertainment service and a cellular communications service, Bolton and Lemon (1999) affirmed that high customer satisfaction is followed by high usage levels. When users do not expect that their needs will be satisfied in the future, based on past experiences and disconfirmed expectations, they are less likely to participate in online community activities. The change of participation level over the past months is therefore an indicator of satisfaction. A positive change suggests increased, a negative change decreased satisfaction. Bolton (1998) shows that customer satisfaction ratings have a positive impact on the duration of the relationship between a cellular phone service provider and its customers. Satisfaction and participation have also been demonstrated to influence continuance intentions in communities (e.g., Chen 2007; Lampe et al. 2010; Lin and Lee 2006; Woisetschlaeger, Hartleb, and Blut 2008). This leads to the following hypotheses:

H7a: Users showing higher active participation are less likely to defect.

H7b: Users showing higher positive changes in their active participation are less likely to defect.

Verified Membership. Trust is an important facet, which supports the willingness to engage in social exchange, and may indicate greater openness to the potential value creation through information exchange (Coleman 1988; Nahapiet and Ghoshal 1998). A trusted relationship leads to commitment and cooperative behaviors that are conducive to relationship marketing success (Morgan and Hunt 1994). Moorman, Zaltman, and Deshpande (1992, p. 315) define trust as "a willingness to rely on an exchange partner in whom one has confidence". In the online world, trust takes an essential part of the relationship between actors who do not necessarily know each other from the beginning (Ba 2001). Members of online communities should ensure that personal information is not misused by the community. Nevertheless, if community members offer personal information, they show confidence to the community operator and its members, though there is uncertainty about what could happen by revealing their personal information (Gross and Acquisti 2005). Relationships that are perceived as very trustful are highly valued; therefore, users will want to commit themselves to such relationships (Morgan and Hunt 1994).

A verified membership in an online community is one explicit way to show such trust and commitment towards the community. In order to verify himself as a 'real' member, the user provides his true personal data in the verification process. In return he is labeled a ‘real’ member on the community site. As the relationship between trust and commitment is evident (Morgan and Hunt 1994), being a verified member also indicates that the user is more committed to the relationship with the community, because the user puts effort in the verification process and the authentication seal only has value if one wants to stay in the community for a longer time. In addition, verified members also create trust for other community members in the sense that they are perceived as honest and reputable, so that other members can have confidence in them (Ba 2001).

Based on a review of studies on trust, Shankar, Urban, and Sultan (2002) conceptualize that online trust is an important predictor of commitment, intentions to act, loyalty and repeat usage behavior. Thereby, trust can be viewed as both a belief in the trustworthiness of others and a behavioral intention to rely on the community in a situation of vulnerability (Shankar, Urban, and Sultan 2002). Empirical evidence of the positive relationship between trust, commitment and retention is provided by several studies (e.g., Morgan and Hunt 1994; Gustafsson, Johnson, and Roos 2005). Also, in the online context, it has been demonstrated that loyalty results from trust and commitment (e.g., Casalo et al. 2009; Zhang et al. 2010). Thus, the following hypothesis is derived:

H8: Verified members are less likely to defect than non-verified members.

7.3 Empirical Study – Methodology

Figure 15 summarizes the hypotheses developed in the previous section. To test the hypotheses, data from the online social community is used, which is introduced in chapter 4.

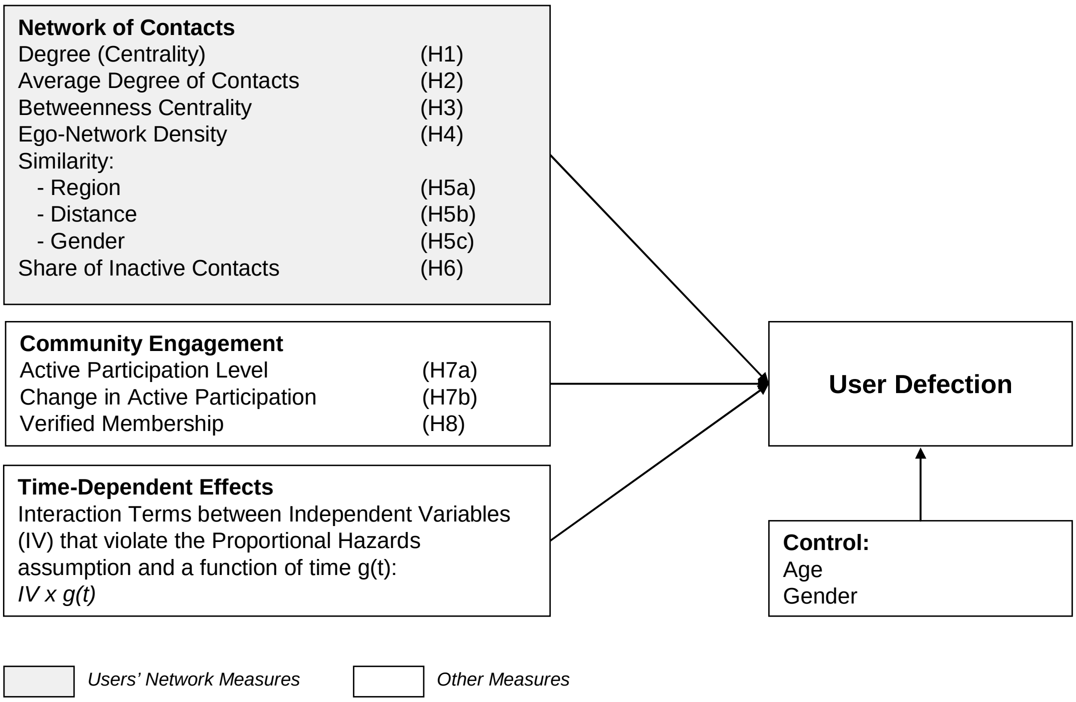


Figure 15: Study 3 – User Defection Model

7.3.1 Data and Sample Characteristics

The operator of the online social community collected the data for this study over a time span of 25 months between 2008 and 2010. In order to study the duration time of users a cohort analysis is used, which is a powerful tool to investigate customer lifetime related issues (e.g., Parasuraman 1997; Reinartz and Kumar 2000). Thereby, all users who registered in the same month for the service are followed until the end of the observation window. This includes both users who are still using the community and users who left the community. The analyses are based on 5752 users, who registered in November 2008, consequently the observations are not left-censored.[68,69] The mean survival time is 9.65 months (median = 7.0) for

[68] Only those users are taken into account, who returned at least once to the platform after registration. Users who only registered, but did not return to the online community show very little interest in using the service and

the entire sample. This cohort contains 83% of failures (users identified as defected), i.e. 17% of the subjects are right censored (survived longer than the end of the time period used in the analysis). The overview of the sample characteristics in Table 23 includes means and standard deviations for the time-varying variables based on the last available observation for each subject. Cohort 1 is used in the main analysis, while Cohort 2 is used in the verification tests in chapter 7.3.6. A description of the relevant variables can be found in the next section.

	Cohort 1		**Cohort 2** *(Validation Sample)*	
	# or % of users		*# or % of users*	
Subjects	5752		5497	
Failures	4775		4553	
Censored Subjects (%)	17%		17%	
Observations	55488		50806	
Gender (female)	55%		55%	
Share of Users Without any Contacts	74 %		75%	
Verified Membership (yes)	8 %		8%	
	mean[1)]	*std.dev.[1)]*	*mean*	*std.dev.*
All Users				
Survival Time (months)	9.65	8.15	9.24	7.79
	(Min = 1; Max = 22)		(Min = 1; Max = 21)	
Age (years)	23.58	5.81	23.60	5.85
Degree (# of active contacts)	11.35	88.85	11.28	64.64
Avg. Active Participation Level (per subject)	2.95	24.85	3.73	27.97
Users with Contacts				
Degree (# of active contacts)	43.88	170.59	45.42	123.64
Average Degree of Active Contacts	213.51	249.04	219.94	257.06
Betweenness[2)] (in active user network)	8.53	6.82	9.05	6.57
Ego-Network Density (of active contacts)	0.20	0.27	0.21	0.26
Average Similarity with Active Contacts				
- same Region (%)	0.61	0.41	0.62	0.40
- avg. Distance to Active Contacts (km)	21.28	31.12	22.42	33.12
- same Gender (%)	0.43	0.34	0.42	0.33
Share of Defected Contacts	0.18	0.25	0.17	0.23
Average monthly Participation Level (per subject)	10.86	47.87	14.16	53.56

1) Means and standard deviations calculated on basis of last available observation for each subject
2) Transformed betweenness measure
Note: All statistics based on the data set with cutoff threshold = 3 months

Table 23: Study 3 – Sample Characteristics

provide only very limited information and data on factors which led to their immediate defection. Therefore, these users are of little value for the current study, which focuses on social effects and user behavior.

[69] Cases with missing values are not included. In fact, missing values only appear in the age variable and make up only a small proportion of the entire sample: 53 subjects, which corresponds to ~1 % of all subjects in Cohort 1 have been withdrawn. Thus, missing values are not a big issue in this study.

7.3.2 *Measurement of Constructs*

Because survival analysis is chosen as an appropriate tool to test the hypotheses on user defection, this study uses longitudinal data. Consequently, it is based on objective data available on the relevant user cohort. The dataset consists of multiple observations for each user. Each observation represents one month for a single user, with the last month including the event of defection or the end of the relevant observation period (censorship). Because most measures are available for each month individually, changes in the structure and relationships are considered in the analysis. The dependent variable is duration time, which reflects the time until the event of defection is observed. Therefore, the entire membership history in the online community is observed for each user of Cohort 1 in a 25-month time window, including the status of defection as well as a set of covariates described below. The correlation table of all observations is provided in Appendix 14, which does not indicate any collinearity issues.

User Defection. Defection is defined as having not logged in on the platform for a certain period of time or when the user unsubscribed from the online social community.[70] Therefore, the last login date of the user determines whether the user is classified as still active (i.e. not defected) or defected. When the last login date is older then a specified threshold value, dated back from the end of the observation period, the user is classified as defected. Hiatus heuristics are used to determine user defection status, because such heuristics are easy to use as a means to define user defection in practice. They also perform well in determining active and inactive customers (Wübben and Wangenheim 2008). For the hiatus heuristic, there needs to be a cutoff threshold *c* below which users are classified as active, and above which users are classified as defected (Wübben and Wangenheim 2008). If a user has not logged in for more than a time span of length *c*, he is considered defected; otherwise, he is considered active. Sensitivity analyses are conducted to assess different threshold values and their effectiveness in correctly classifying active and defected users. For this purpose, each user's defection status is determined by the last login date, taken at the end of the observation period, and the cutoff threshold is set to one, two, three, four, five, and six months[71]. For example, if a cutoff value of three months is taken, a user who has not logged in within the last three months before the end of the observation window is declared as defected in the month of last login (de-

---

[70] Note that deregistrations occur to a much lesser extent than just not returning to the community. Therefore, it is also important to define defection by the time the user has not logged in recently.

[71] Given that online communities operate in a fast-moving market, hiatus lengths is tested of up to six months.

fection event=1). To check if users have been correctly classified, i.e. if they have been online after the observation period, the last login information available from data nine months after the end of the observation period is examined and the correct classification is assessed. If subjects exhibit any login in the period after the observation window, they are designated still active, otherwise they are designated as defected. Therefore, it is possible to determine if active users were still active, and if defected users did not return in the time after the observation period. The correct classifications for six cutoff values (1 to 6 months) are compared. A cutoff threshold of *c* is considered optimal, if it maximizes the percentage of overall correctly classified active and defected users. Table 24 shows the results of correctly classified active and defected users for each cutoff threshold. The table illustrates that the best result is given at a cutoff date of three months, showing the highest total percentage of correctly identified users. This cutoff threshold is also consistent with the community operator's definition of inactive users. Thus, defection is determined for all users who had their last login date three months before the end of the observation window. In the main analyses only the results for models based on the three-months cutoff threshold are presented.[72]

	Cohort 1					
Cutoff Threshold[1)]:	1 month	2 months	3 months	4 months	5 months	6 months
Number of Subjects	5737[2)]	5737	5737	5737	5737	5737
Active, correctly classified (%)	47.13%	58.87%	64.96%	68.14%	69.99%	71.67%
Defected, correctly classified (%)	98.46%	97.00%	95.53%	94.05%	93.20%	91.92%
Overall correctly classified (%)	88.32%	89.47%	89.49%	88.93%	88.62%	87.92%
Active, but classified defected (%)	52.87%	41.13%	35.04%	31.86%	30.01%	28.33%
Defected, but classified active (%)	1.54%	3.00%	4.47%	5.95%	6.80%	8.08%
Overall incorrectly classified (%)	11.68%	10.53%	10.51%	11.07%	11.38%	12.08%

Note: Classification based on comparison of last login date at the end of the observation window compared to the last login data nine months after the observation window
1) Cutoff Threshold: users who have not been logged in for more than x months are designated as defected
2) n=5737 due to missing values in the validation data of last login dates

Table 24: Study 3 – Analysis of Inactivity Cutoff Thresholds

Independent variables. The dataset provides time-constant as well as time-varying covariates. Time-constant covariates include age (last age at the end of the observation period), gender, and verified membership. These variables show the same values over time for each individual user. In addition, the users' network configuration and behavior is followed over time. Time-varying covariates include degree, average degree of the users' contacts, be-

[72] However, robustness tests with other threshold values in the models are pursued, which are presented later in this study.

tweenness centrality, ego-network density, average similarity with active contacts, share of defected contacts, and participation variables. All time-varying covariates are measured on a monthly basis to approximate a continuous-time process (e.g., Nitzan and Libai 2011).

Network variables. Network analytical measures – degree, betweenness centrality, and ego-network density – are calculated using the social networks software Pajek (de Nooy, Mrvar, and Batagelj 2005). The interconnections between users are based on their 'friendships', where each user must explicitly connect with another user in a bidirectional manner to establish a friendship tie. For each month in the observation period, the network of all active users is used to determine the users' network position (including users not part of Cohort 1).[73] Note that users can have more contacts in their friends list than they have active contacts. This is because most members do not explicitly unsubscribe from the platform, rather they do not return and are therefore determined as lost for good when they have not returned for a specified period of time (see above for a discussion about the threshold for defection). Because of the focus on the dynamic social context and user defection in this study, it is critical to consider the user's active social network and defected contacts individually. Based on this active friend network, which is an undirected network between users, degree, betweenness centrality and ego-network density are computed for each month of the observation period. For the formula and description of the network measures see chapter 3.2.2.

As explained in chapter 3.2.2, degree represents the number of active contacts of the focal user in the respective month; ego-network density describes how closely connected all active contacts of the focal user are to one another; and betweenness centrality describes the central position of the focal user in terms of his location on the shortest paths between all other pairs of active users in the network. Because of the nature of the betweenness centrality measure in large networks, the values of this measure become very small for a large number of users, consequently showing a high variance. Therefore, the variance and the effect of extreme values for betweenness centrality are reduced. To gain more stable results, the variable is transformed in the following way: betweenness centrality was first multiplied with the inverse ratio of the lowest non-zero value to determine the lowest betweenness value at value 1. The new values are then log-transformed.

[73] When the event of defection takes place in month t (i.e. the last login of the user, where defection is defined above in this study), then the user is still active until and including month t, but determined as defected, and therefore ignored in the network analysis in month t+1 and the following months. Defected users are not included in the actual monthly network of the focal user, because the tie virtually still exists, but is not maintained anymore, and therefore of no value for the focal user.

Average degree of active contacts. The average degree of the focal user's neighbors is calculated for each user *i* in month *t*. For each neighbor *j* the degree of his active contacts is computed, and the sum of all neighbors' degrees in month *t* is averaged by the number of active neighbors *j* in this month ($N_{j,t}$).

$$(7.1) \quad avgContactDegree_{i,t} = \frac{\sum_{j \in SN_i} ContactDegree_{j,t}}{N_{j,t}}$$

Share of defected contacts. The number of defected contacts includes all neighbors of the focal user, who defected in the respective month (according to the definition of defection). Because defection implies that the user in many cases does not return to the platform (silent attrition), the focal user does not necessarily and immediately realize the defection of contacts. As friends who became inactive in the preceding months can also affect the focal user's loyalty, for example when the overall ratio of defected to active friends becomes high, the cumulative number of defected contacts is taken into consideration. Therefore, this study employs the ratio of all defected contacts of user *i* to all relationships this user established with other users until month *t* in total:

$$(7.2) \quad ShareDefected_{i,t} = \frac{\sum_{t=1}^{x} Defected_{i,t}}{TotalTies_{i,t}}$$

Simultaneity between focal users and their neighbors in becoming inactive is addressed by using the exact dates of defection, and thus ensuring that only users are counted as defected who last logged in before the focal user. Although last login data is available on a daily basis, monthly aggregated data is used for the analyses, because not all covariate data is available on a daily basis.

Similarity variables. In the analysis, covariates representing the similarity between the focal user and his active contacts are included. The average similarity of all direct and active neighbors of the focal user is calculated with respect to the characteristics of interest. Three characteristics are used for this purpose: geographical location, distance between geographical locations, and gender.[74] Binary variables (1=match, 0=no match) are introduced to indicate similarity of geographical location and gender. If both users have the same gender or are registered in the same region or city their similarity is 1. In the case of geographical distance,

[74] Age is not used as a similarity variable, because the online community is targeted at a specific age group already, and there are missing values in age for some users which would potentially bias the similarity variable.

similarity represents the distance between the region/city of the focal user *i* and the neighbor *j*. If they both belong to the same region the distance is 0; if they belong to different regions the similarity is the air-line-distance between the centers of these two regions measured in kilometers[75]. In general, similarity is defined by:

$$(7.3) \quad avgSimilarity_{i,t} = \frac{\sum_{j \in SN_i} Similarity_{i,j} * \delta_{j,t}}{\sum_{j \in SN_i} \delta_{j,t}}, \quad \text{with } avgSimilarity{=}0 \text{ if } \sum_{j \in SN_i} \delta_{j,t} = 0,$$

where user *j* is a neighbor of user *i* and therefore part of *i's* social network ($j \in SN$); $\delta_{j,t}$ is a binary variable (0 or 1) that reflects whether user *j* is an active contact of *i* in month *t*.

Participation variables. The users' active participation in the online social community is measured by an index that includes out-going activity, i.e. communication with other users, comments and ratings.[76] Two measures of participation are considered in the analysis. First, the level of participation consists of aggregated activity data for each single month in the data set. Second, in accordance to Nitzan and Libai (2011) and Bolton and Lemon (1999) a change in usage is used to indicate satisfaction. Therefore, a delta-measure is calculated to describe the difference in usage compared to the preceding months. This measure shows how the engagement in the community changed – positively, negatively, or not at all. This change is calculated as the difference between the current level of participation in month *t* minus the mean participation in the three months preceding *t*:

$$(7.4) \quad \Delta Activity_{i,t} = Activity_{i,t} - average(Activity_{i,t-1}, Activity_{i,t-2}, Activity_{i,t-3})$$

An average of the three preceding months is used to avoid bias caused by fluctuations in customers' usage levels over time. Due to missing values in the first three months (inherent in the cohort approach), the delta measure is calculated as the average of the non-missing lagged variables of activity in the first three months, while in month 1 the delta activity equals the activity in this month.

Verified membership. The online community users can achieve a verified membership by applying for authentication of their personal data. The authentication process requires true

[75] The distance was gained by entering the name of the city or region in the provided calculator of the website http://www.luftlinie.org, which calculates the air-line-distance of the first to the second region's center.
[76] In this study, active participation represents the number of messages sent to other users, written guest book entries, virtual gifts sent to other users, comments in groups, and submitted ratings on photos and groups.

personal data, specifically real name and address, where the authentication code can be sent. When the code is correctly entered on the community website, the user appears as a "verified" member, which is presented as a symbol on the user's profile page and is visible for all other users. The variable is therefore a binary variable with value 1 representing the verified membership, and value 0 otherwise.

Control variables. In addition, age and gender are included to control for heterogeneity and the tendency of specific user groups to behave similarly due to their demographic similarities. The robustness tests also include the region, where the user is registered, as an additional control variable.

7.3.3 Modeling Influence on User Defection

7.3.3.1 Survival Analysis and the Cox Proportional Hazards Model

In survival analysis the impact of time and dedicated covariates on the occurrence of an event are modeled. This study focuses on the effects of different covariates on a user's defection, i.e. the decision of not returning to the online social community. Therefore, the objective is to assess the risk or hazard of ending the relationship with the online community. To model duration, survival analysis is pursued by using a Cox regression model, also known as the Cox proportional hazards model (Cox 1972)[77]. Cox proportional hazards (PH) models are superior to OLS or logit regressions as one can make use of censored data[78] and include time-varying covariates in the analysis (e.g., Allison 2010; Helsen and Schmittlein 1993; Hosmer, Lemeshow and May 2008; Kleinbaum and Klein 2005). For the analysis, the semi-parametric Cox regression is used, because it has the advantage that no a priori determination of the underlying distribution for the time until an event occurs is needed (Allison 2010; Box-Steffensmeier and Zorn 2001).[79] Thereby, the Cox proportional hazards model is robust and will closely approximate a correct parametric model (Kleinbaum and Klein 2005). For this

[77] The term Cox proportional hazards model is used here, although the hazards are not always proportional for all of the covariates. This study also shows later how the issue of non-proportionality can be approached.

[78] In survival analysis, the observation time is often completed before the endpoint is reached for all subjects, i.e. the survival time is censored in that subjects survived to a certain point in time beyond which their status is unknown (SAS Institute 2008). The Cox proportional hazards model (Cox 1972) is able to account for both censored and uncensored data.

[79] Parametric models, such as the Weibull, exponential, or log-logistic model were not used for this purpose, as they rely on a priori assumptions of the underlying distributional form. In the case of unknown distribution of survival time, Hosmer, Lemeshow and May (2008) suggest to apply semi-parametric models like the ordinary Cox model applied in this study.

reason, Cox regression models have been widely used in marketing to investigate effects of predictors on some kind of event, for example, customer churn or purchase timing behavior (e.g., Bolton 1998; Jain and Vilcassim 1991; Li 1995; Nitzan and Libai 2011; Reinartz and Kumar 2003; Schmitt, Skiera, and van den Bulte 2011; van den Poel and Lariviere 2004).

The Cox PH model consists of two parts, which describe the hazard at time *t* for an individual with a given specification of a set of predictors *X* (Kleinbaum and Klein 2005):

$$(7.5) \quad h(t, X) = h_0(t) * e^{\sum_{i=1}^{p} \beta_i X_i}$$

The first term represents the baseline hazard $h_0(t)$, which is a function of time, but does not involve the predictor variables. The second term, the exponential expression, involves the predictors, but not time. In this basic form, the predictors (X_i) are time-independent. [80]

It is possible to include time-varying predictors in the Cox PH model. This is a valuable feature, because many of the covariates of interest change over time. The use of time-varying variables is important, as averaged variables or the last available value of an independent variable can be misleading in their effects on survival time (Fisher and Lin 1999). Here, time-varying values of independent variables allow for a more accurate analysis. The basic Cox PH model can be extended by such variables. For this purpose, the counting process method of input is used to include time-varying repeated measurements in the analysis (SAS Institute 2008; Therneau and Grambsch 2000). In the counting process format there are multiple records for each individual, where time-variation of the predictors X_i is taken into account by determining each time interval with a value for each X_i at this time-interval *t* being constant (Allison 2010; Therneau 1996; Therneau and Grambsch 2000).[81]

7.3.3.2 Proportionality of Hazards

For the interpretation of the outcome of Cox regression models, the hazard ratio is the measure of effect. It is calculated as the ratio of the hazard functions for two individuals, i.e. the hazard for one individual divided by the hazard for a different individual. Thus, it de-

---

[80] Parameter estimates for the Cox regression model are obtained by using partial likelihood functions. Please refer for example to Cox (1972), Hosmer, Lemeshow, and May (2008), or Kleinbaum and Klein (2005), Therneau and Grambsch (2000) for more information on partial likelihood estimations.

[81] Each observation describes a semi-closed time (T1,T2] during which the values of the explanatory variables remain unchanged. Each record also contains the censoring status at T2. The individual remains at risk during that interval, and an event might occur in T2. Andersen and Gill (1982) first elaborated on the counting process within this context and built the theoretical basis for using this method in survival models. For more information on the analysis using counting process style of input see for example Therneau and Grambsch (2000) or the SAS User Guide (SAS Institute 2008).

pends only on the exponential part of the Cox model, as the baseline hazards will cancel out (e.g., Box-Steffensmeier and Zorn 2001; Kleinbaum and Klein 2005):

$$(7.6) \quad \frac{h(t, X_i)}{h(t, X_j)} = e^{\beta(X_i - X_j)}$$

One critical assumption for Cox regression models is the proportionality of the hazards (Hosmer, Lemeshow, and May 2008; Kleinbaum and Klein 2005; Therneau and Grambsch 2000). This means that the hazard ratios, e.g. for individuals *i* and *j*, do not change. Consequently, the effects of the covariates are to shift the hazard by a factor of proportionality, while the size of that factor remains constant over time (Box-Steffensmeier and Zorn 2001). There are circumstances in which the PH assumption does not hold, for example, when the influence of a predictor may be greater or smaller at different points in time. Several authors have warned that estimating proportional hazard models in situations where the hazards are, in fact, not proportional can result in biased coefficient estimates and decreased power of significance tests (Box-Steffensmeier and Zorn 2001; Kalbfleisch and Prentice 1980; Schemper 1992). For example, the effect of a variable that is statistically significant but changing over time may be found to be statistically insignificant when using conventional techniques (Box-Steffensmeier, Reiter, and Zorn 2003). Further, not only the estimates of predictors that violated the proportionality assumption might be biased, but also the estimates of all other parameters could be influenced (Box-Steffensmeier and Zorn 2001). This suggests testing if the proportionality assumption holds for all independent variables.

Several approaches are provided in the literature to test the PH assumption. In this study, two common statistical tests are used, which have been demonstrated to have good power in detecting non-proportionality (for a comparison of different tests see Ng'andu 1997). It is tested by using Schoenfeld residuals, as well as time-dependent covariates, which involve interaction terms of each of the covariates with a function of time (e.g., Kalbfleisch and Prentice 1980; Kleinbaum and Klein 2005; Ng'andu 1997).[82] First, the scaled Schoenfeld residuals are examined (Grambsch and Therneau 1994), which is based on a test proposed by Schoenfeld (1982). Schoenfeld residuals are defined for every subject who has an event, with

[82] Although graphical inspection of log-log-survival curves provides a qualitative indication of proportionality or non-proportionality, it is not recommended to solely rely on graphical approaches (Box-Steffensmeier and Zorn 2001). One shortcoming of these graphical evaluations is that they do not consistently and correctly diagnose instances of non-proportionality, especially when considering a high number of covariates and continuous variables (e.g. Kleinbaum and Klein 2005; Schemper 1992). Therefore, more objective statistical approaches are used to test for proportionality.

one residual for each of the predictors. Statistical tests for non-proportionality are used, which are based on these residuals: Harrell's (1986) correlation test for individual variables[83], as well as calculating Grambsch and Therneau's (1994) global test for non-proportionality. Thereby, a statistically significant test is indicative for non-proportionality.[84]

Second, the Cox model is extended to include time-interaction terms for every independent variable with some function of time in order to assess the PH assumption (Kleinbaum and Klein 2005; Ng'andu 1997). Therefore, the basic Cox PH model can be extended by such variables:

$$(7.7) \quad h(t;X) = h_0(t) * e^{\sum_{i=1}^{p}\beta_i X_i + \sum_{i=1}^{p}\delta_i X_i g_i(t)}$$

While the first part in the exponential term contains the predictors being assessed as main effect terms, the second part contains product terms of the predictors with some function of time *g(t)*. Wald statistics are used for assessing the significance of time-interaction terms, when all variables are simultaneously included in an extended Cox regression model. To further test whether the PH assumption is adequate, likelihood ratio tests are used to compare the model without time-interaction terms with the models that include the interaction between each single covariate with time separately (one-at-a-time), adjusting for the main effects of the other covariates. For this purpose, the analysis uses *g(t)=ln(t),* which is a popular function of time when considering time-dependent effects (e.g., Kalbfleisch and Prentice 1980; Ng'andu 1997).[85] A significant likelihood ratio test is an indication for non-proportional hazards and therefore time-dependent effects.

Consequently, the hazard ratio for the extended Cox regression model, which is a function of time, is given as follows (Kleinbaum and Klein 2005):

$$(7.8) \quad H\hat{R}(t) = \exp\left[\sum_{i=1}^{p1}\hat{\beta}_i(X_i^* - X_i) + \sum_{j=1}^{p2}\hat{\delta}_j(X_j^*(t) - X_j(t))\right]$$

[83] If the PH assumption holds for a particular covariate then the Schoenfeld residuals for that covariate will not be related to survival time, i.e. the correlation with time should be near zero (Kleinbaum and Klein 2005). Different statistical software packages use different variations of this test. For example, STATA allows to investigate the correlation between the scaled Schoenfeld residuals (Grambsch and Therneau 1994) for a particular covariate and the log of the survival time. However, the tests of different programs typically (but not always) yield similar results (Kleinbaum and Klein 2005).

[84] For more information on Schoenfeld residuals please refer for example to Schoenfeld (1982), Grambsch and Therneau (1994), or Hosmer, Lemeshow, and May (2008).

[85] This is consistent with the use of the natural logarithm of time in the main analysis, which is described below.

This hazard ratio at a particular time t requires the specification of two sets of predictors, $X^*(t)$ and $X(t)$ at time t. As the hazard ratio is a function of time, the coefficient $\hat{\delta}_j$ indicates an increase (if positive) or decrease (if negative) of the hazard ratio with increasing time (the hazard ratio is not constant). In the test involving time-interaction terms, the hazard ratio is constant for all t only when the coefficient of the interaction term is not significant.

The interpretation of an estimated (constant) hazard ratio of, for example, 1.10 is that the hazard rate increases by 10% for a unit increase in the independent variable (Hosmer, Lemeshow, and May 2008). This is true when the proportional hazards assumption holds and there are no time-dependent coefficient effects. In cases where non-proportionality is observed, the hazard ratio changes over time and for each time-interval t, there is a different hazard ratio. Therefore, the hazard ratio increases (for positive $\hat{\delta}_j$) or decreases (for negative $\hat{\delta}_j$) with time (SAS Institute 2008). For example, a time-dependent hazard ratio is calculated as the time-independent coefficient of the covariate plus the product of the coefficient of the time-interaction term of that independent variable and the function of time. The combined coefficient $\beta_j(t) = \hat{\beta}_i + \hat{\delta}_j * g(t)$ is the actual coefficient of a covariate showing non-proportionality, which depends not only on $\hat{\beta}_i$ and $\hat{\delta}_j$, but also on time (e.g., Golub and Steunenberg 2007; Kleinbaum and Klein 2005). Note, that the model with proportional hazards, for a given covariate j, corresponds to the restriction $\beta(t) = \beta$, i.e. that a plot of $\beta_j(t)$ versus time will be a horizontal line (Therneau and Grambsch 2000).

7.3.3.3 Approaching Non-Proportionality

Different approaches exist to address non-proportional hazards. One often recommended approach is to use interaction terms of covariates with time for those covariates that violate the PH assumption. In addition to their utilization as indicators for non-proportionality in PH assumption tests, these interaction terms are an adequate procedure to incorporate non-proportional time-dependent effects in situations where this is an issue (Allison 2010; Box-Steffensmeier and Zorn 2001; Kleinbaum and Klein 2005; Schemper 1992; Therneau and Grambsch 2000).[86] In essence, non-proportionality can be considered as similar to instances

[86] An alternative way of addressing non-proportionality is to stratify the data by the covariate of interest. Under stratification, the hazard ratios of the remaining covariates are assumed to be constant across strata, but the baseline hazards are allowed to be different for each stratum (e.g., Box-Steffensmeier and Zorn 2001; Kleinbaum

where covariate effects change depending on some other variable, where in this case the other variable is time. An additional advantage is that this approach allows to explicitly model the nature of the non-proportionality, resulting in a more accurately specified model and greater validity in the overall results (Box-Steffensmeier und Zorn 2001).

Therefore, an extended Cox regression model with time-dependent effects is used for all analyses in this study. For the hypotheses tests, different models are specified, which include the respective time-constant and time-varying covariates and the interaction terms with time for those covariates, which violate the proportional hazards assumption. Consequently, the models are parameterized with the predictors of interest and allow for time-varying covariates and time-dependent effects when necessary: $h(t,x) = h_0(t)\exp(\beta(t)x(t))$ (see Hosmer, Lemeshow, and May 2008).[87]

As already explained, the interaction with time allows the predictor to vary monotonically according to some function *g(t)* of time. For non-proportional effects, the functional form of time in the interaction term is important to determine how the hazards change. As the appropriate function of time could not be determined a priori, the analyses are conducted with different time functions to identify *g(t)* which fits the models best, thereby comparing models incorporating *t*, *ln(t)*, t^2, and $\sqrt{t}$ as functions of time (e.g., Box-Steffensmeier, Reiter, and Zorn 2003; Kalbfleisch and Prentice 1980; Kleinbaum and Klein 2005). The following analyses use *ln(t)*, because across all models the application of *ln(t)* as the function of time results predominantly in better or at least equally good model fit (in terms of lower log-likelihood, AIC and BIC statistics) compared to models with other functions of time (a comparison of the different time function models can be found in Appendix 17).

7.3.4 The Online Community Defection Model

The main analyses are basically separated in three models. First, the entire cohort is used as the sample in Model 1. Second, the entire sample is split in two user groups: a) users who do not have any contact at any time of the observation period (named “Not-Connected Users”, Model 2), and b) users who have at least one contact within the observation period (named “Connected Users”, Model 3). The Connected Users are of particular interest in this

---

and Klein 2005). The major drawback of stratified Cox regression models is that the impact of the variable of interest on the hazard rate is not estimated. As the estimates for all variables are required to test the hypotheses in this study, stratified models are not used.

[87] Note, that it is differentiated here between time-dependent effects (in cases of non-proportionality) and time-varying covariates (changing values over time) (see for example Therneau 1996).

study and build the core of the empirical analysis. The reason is that those users are involved in the online social community and the effect of the social context can only be measured for this user group. They represent the ‘true’ community users. However, in online communities which provide firm-generated content in addition to user-generated content Not-Connected Users exist, who rather consume content as they are not involved in the community. For Not-Connected Users social network characteristics are not present. Due to the fact that the majority of users in the cohort do not have any contact, and thus have zero or missing values in all network variables, results on the effects of the social context would be biased. However, as they represent a large number of users in the sample, results for some of the variables are also reported for this user group to investigate some differences to the community users.

Model 1 investigates the effects of degree and community engagement on user defection. It includes degree, active participation level, change in active participation, verified membership, age and gender as independent variables. The PH assumption is tested as described above. The global test provides strong evidence that non-proportional effects exist (p<.01). Further, the test specifically found degree, active participation level, verified membership and gender to violate the PH assumption. Table 25 summarizes the results of both covariate-specific and global tests for non-proportionality. To address non-proportionality and to better assess the time-dependent effects of these variables, the model is estimated including interaction terms between these covariates and time. The hazard of the defection of a user *i* at time *t* is thus given as follows:

$$
\begin{aligned}
(7.9)\quad & h(t,x) = h_0(t)\exp(\beta_1 Degree_{it} + \beta_2 Participation_{it} + \beta_3 DeltaParticipation_{it} \\
& + \beta_4 VerifiedMembership_{it} + \beta_5 Age_i + \beta_6 Gender_i \\
& + \delta_1 Degree_{it} * \ln(t) + \delta_2 Participation_{it} * \ln(t) \\
& + \delta_3 VerifiedMembership_{it} * \ln(t) + \delta_4 Gender_{it} * \ln(t))
\end{aligned}
$$

Model 2 is based on the Not-Connected User group, including all community engagement and control variables, but none of the social network variables. According to the PH assumption tests, presented in Table 25, active participation level, verified membership and gender need to be interacted with *ln(t)*, leading to the following model:

$$
\begin{aligned}
&h(t,x) = h_0(t)\exp(\beta_1 Participation_{it} + \beta_2 DeltaParticipation_{it} \\
(7.10)\quad &+ \beta_3 VerifiedMembership_{it} + \beta_4 Age_i + \beta_5 Gender_i \\
&+ \delta_1 Participation_{it} * \ln(t) + \delta_2 VerifiedMembership_{it} * \ln(t) + \delta_3 Gender_{it} * \ln(t))
\end{aligned}
$$

Model 3a is solely based on the Connected User group and incorporates the same covariates as Model 1. Further, Model 3b extends Model 3a by including the whole set of social network variables to better understand social effects for this group of users (as it does not make much sense to estimate those covariates for the Not-Connected Users). Model 3b can thereby be seen as the final model, which includes the overall set of all relevant variables, and is meant to answer the important question of how an individual's social context influences defection behavior. Therefore, the following variables are added to the model: average degree of the users' contacts, betweenness, ego-network density, share of defected contacts, and similarity in gender, region and distance with a user's active contacts. Again, time-interaction terms are included for all variables violating the PH assumption, namely degree, average degree of contacts, betweenness, ego-network density, similarity in region and distance, and gender (PH assumption tests are also presented in Table 25). The complete Model 3b is specified as:

$$
\begin{aligned}
&h(t,x) = h_0(t)\exp(\beta_1 Degree_{it} + \beta_2 avgContactsDegree_{it} + \beta_3 Betweenness_{it} \\
&+ \beta_4 EgoNetworkDensity_{it} + \beta_5 avgSimilarityregion_{it} + \beta_6 avgSimilaritydist_{it} \\
&+ \beta_7 avgSimilaritygender_{it} + \beta_8 ShareDefected_{it} + \beta_9 Participation_{it} \\
(7.11)\quad &+ \beta_{10} DeltaParticipation_{it} + \beta_{11} VerifiedMembership_{it} + \beta_{12} Age_i + \beta_{13} Gender_i \\
&+ \delta_1 Degree_{it} * \ln(t) + \delta_2 avgContactsDegree_{it} * \ln(t) + \delta_3 Betweenness_{it} * \ln(t) \\
&+ \delta_4 EgoNetworkDensity_{it} * \ln(t) + \delta_5 avgSimilarityregion_{it} * \ln(t) \\
&+ \delta_6 avgSimilaritydist_{it} * \ln(t) + \delta_7 Gender_{it} * \ln(t))
\end{aligned}
$$

	Schoenfeld Residual Test			Time-Interaction Terms - Time-Dependent Effects[1]			Time-Interaction Terms - LR-Test (one-at-a-time)[2]		Use of Time-Interaction Term in Model[3]
	Log t			ln(t)*			ln(t)*		
MODEL 1 - Covariates	ρ	χ2	p	Coef.	SE	p	χ2	p	
Degree (Active Contacts)	0.033	91.94	0.000	0.043	0.003	0.000	412.03	0.000	Yes
Active Participation Level	0.024	16.10	0.000	0.012	0.004	0.004	108.82	0.000	Yes
Active Participation Delta	0.023	2.29	0.130	-0.002	0.003	0.376	1.40	0.237	No
Verified Membership	0.036	8.15	0.004	0.157	0.058	0.007	15.19	0.000	Yes
Age	-0.001	0.01	0.915	-0.001	0.002	0.537	3.48	0.062	No
Gender	0.052	12.88	0.000	0.093	0.027	0.001	11.15	0.001	Yes
Global Test / LR-Test		185.73	0.000		441.82	0.000			
MODEL 2 - Covariates									
Active Participation Level	0.016	1.75	0.186	0.038	0.018	0.029	4.22	0.040	Yes
Active Participation Delta	-0.003	0.04	0.849	-0.017	0.014	0.225	0.00	0.976	No
Verified Membership	0.032	4.09	0.043	0.148	0.075	0.048	3.05	0.081	Yes
Age	-0.026	2.60	0.107	-0.004	0.003	0.111	3.42	0.065	No
Gender	0.029	3.29	0.070	0.056	0.031	0.068	4.00	0.046	Yes
Global Test / LR-Test		12.16	0.033		16.19	0.006			
MODEL 3a - Covariates									
Degree (Active Contacts)	0.033	9.25	0.002	0.015	0.003	0.000	54.14	0.000	Yes
Active Participation Level	0.023	1.34	0.247	0.007	0.005	0.209	15.29	0.000	No
Active Participation Delta	-0.005	0.03	0.870	-0.004	0.004	0.271	2.05	0.152	No
Verified Membership	0.044	1.91	0.167	0.162	0.130	0.212	2.45	0.118	No
Age	0.035	1.75	0.186	0.011	0.007	0.127	0.62	0.432	No
Gender	0.093	7.69	0.006	0.235	0.082	0.004	6.5	0.011	Yes
Global Test / LR-Test		27.54	0.000		68.4	0.000			
MODEL 3b - Covariates									
Degree (Active Contacts)	0.010	0.73	0.392	0.010	0.003	0.000	33.13	0.000	Yes
Avg. Degree of Active Contacts	-0.091	5.01	0.025	-0.001	0.000	0.000	4.66	0.031	Yes
Betweenness (Active Contacts)	0.065	12.04	0.001	0.015	0.008	0.083	16.04	0.000	Yes
Ego-Network Density (Active Contacts)	0.071	4.99	0.026	0.258	0.146	0.078	1.98	0.159	Yes
Similarity (Active Contacts) - Region	-0.141	17.07	0.000	-0.446	0.117	0.000	3.72	0.054	Yes
Similarity (Active Contacts) - Distance	-0.142	5.98	0.015	-0.002	0.001	0.056	0.27	0.603	Yes
Similarity (Active Contacts) - Gender	-0.001	0.00	0.973	0.025	0.110	0.820	0.33	0.568	No
Share of Inactive Contacts	-0.030	0.71	0.399	-0.181	0.210	0.389	1.01	0.314	No
Active Participation Level	0.016	0.67	0.413	0.005	0.005	0.301	9.51	0.002	No
Active Participation Delta	-0.007	0.05	0.831	-0.004	0.004	0.291	1.69	0.193	No
Verified Membership	0.036	1.22	0.270	0.137	0.131	0.296	1.97	0.160	No
Age	0.030	1.00	0.317	0.007	0.007	0.337	0.08	0.779	No
Gender	0.097	8.09	0.005	0.235	0.084	0.005	6.59	0.010	Yes
Global Test / LR-Test		49.02	0.000		73.94	0.000			

**) Time-interaction terms are generated by Covariate x g(t), where g(t) = ln(t) for the executed tests.*
1) Inclusion of interaction terms with time for all covariates simultaneously. p-values represent significance test for each time-interaction term adjusted by all other covariates.
2) Time-interaction terms for one covariate at a time, adjusted by all other covariates. LR-Test of model with time-interaction terms vs. model without time-interaction term with 1 df.
3) Time-interaction terms are included in the respective model if 2 of the 3 tests indicate significant (p<.10) time-dependent effects.
ρ reports the estimated correlations between the scaled residuals and log t
χ2 and p-values indicate the confidence with which the null hypothesis that the hazard ratios for different values of that covariate are constant over time can be rejected.
χ2 statistics have one degree of freedom, except for global test of Schoenfeld Residuals and Time-Dependent Effects (model1 = 6 df; model2 = 5df; model3a = 6df; model3b = 13df)

Table 25: Study 3 – PH Assumption Tests

7.4 Results of Main Analysis

For all analyses STATA's stcox command is used, with *ln(t)* as the functional form of the time-dependent coefficients and Efron (1977) estimation for handling ties. Because there are a large number of ties in the data, Efron's approximation is more accurate than Breslow (1974) approximation (Allison 2010). A cutoff threshold of three months is used for all main analyses, as described and evaluated above (see description of dependent variable "User Defection"). For this purpose, the observation window used in the analyses is set to 22 months. Figure 16 illustrates the relevant observation period. The reason for cutting off the last three months is that all members who were online after 22 months would be considered active for the whole three months past the cutoff date, although it cannot clearly be determined if they are active or defected by using the three month cutoff threshold. Defining defection, looking back from the end of the observation period, could then lead to false assumption for the status of active users in the time span after the cutoff threshold. Therefore, all observations of users are excluded which can not definitely be classified as active or inactive, considering all users with login dates past the cutoff date as censored at the end of month 22. This makes the analysis more accurate.

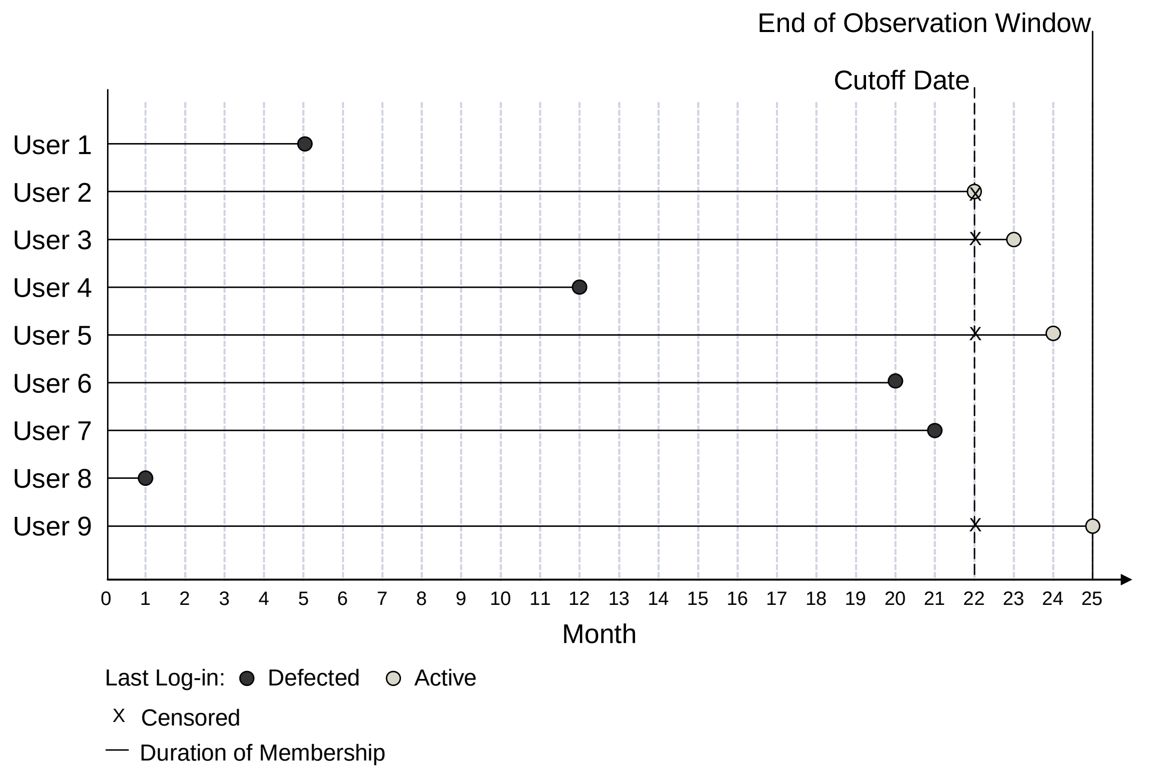


Figure 16: Study 3 – Description of Observation Window and Cutoff Date

Analysis of the Overall Effects for the Total Sample. Model 1 includes degree and the community engagement variables, and reveals that all main effects are significant and in the hypothesized direction (see Table 26). A higher number of contacts reduces the hazard of defection, which supports hypothesis H1. Higher levels of active participation and change in active participation compared to the preceding months are also negatively associated with defection, resulting in lower hazard rates (H7a and H7b supported). Further, users with a verified membership status are less likely to defect, thus confirming hypothesis H8.

Although the main effects are significant, the analysis also reveals that the time-interaction terms are significant, indicating that the effects of degree, active participation level and verified membership decrease over time. Because of the interaction of these covariates with *ln(t)*, the estimates of the direct effects can be interpreted as the effect of that covariate on the hazard of defection in the first month following the registration on the platform (that is when *t=1*). In the following months, the main effect is reduced by the time-dependent effect. For example, it is observed that the effect of the users' degree, strong in the initial month after registration, declines as the membership proceeds, so that in month 22 of membership its influence becomes insignificant.[88] Figure 17 shows the development of the hazard ratios as a function of time for the predictors that violate the PH assumption. Similarly, the time-dependent effects of active participation level and verified membership reduce the main effect, i.e. the negative impact on the likelihood of defection decreases over time for those covariates. While verified membership still has a significantly negative effect over the entire length of the observation period, the effect of the active participation level wanes over time, so that the effect starts to be insignificant in month 14. Thereby, although the hypotheses are in general supported, for some variables they hold only for a limited period of time. All time-dependent effects are presented in Figure 17.

[88] The combined log(HR) of the main effect and the time-dependent effect is $(\beta_i + \delta_i * \ln(t))$. Its standard error is $Std.Err.(\beta_i + \delta_i * \ln(t)) = \sqrt{\text{var}(\hat{\beta}_i) + (\ln(t))^2 * \text{var}(\hat{\delta}_i) + 2\ln(t) * \text{cov}(\hat{\beta}_i, \hat{\delta}_i)}$. A pointwise 95%-confidence interval can thus be obtained by taking the exponent of log(HR)±1.96∗Std.Err.(log(HR)) (Golub and Steunenberg 2007; Putter et al. 2005).

COHORT 1		Model 1				Model 2				Model 3a				Model 3b				Hypotheses	
		All Users				Not-Connected Users				Connected Users - Base				Connected Users - Full				Test	
Covariates		HR	Coef.	SE	p	HR	Coef.	SE	p	HR	Coef.	SE	p	HR	Coef.	SE	p		
Main Effects of Covariates																			
Degree (Active Contacts)		0.875	-0.133	0.010	0.000 ***	-	-	-	-	0.952	-0.050	0.008	0.000 ***	0.966	-0.035	0.008	0.000 ***	H1	supp.
Avg. Degree of Active Contacts		-	-	-	-	-	-	-	-	-	-	-	-	1.002	0.002	0.000	0.000 ***	H2	rev. eff.
Betweenness (Active Contacts)		-	-	-	-	-	-	-	-	-	-	-	-	0.927	-0.076	0.021	0.000 ***	H3	supp.
Ego-Network Density (Active Contacts)		-	-	-	-	-	-	-	-	-	-	-	-	0.455	-0.788	0.346	0.023 **	H4	supp.
Similarity (Active Contacts) - Region		-	-	-	-	-	-	-	-	-	-	-	-	3.742	1.320	0.255	0.000 ***	H5a	rev. eff.
Similarity (Active Contacts) - Distance		-	-	-	-	-	-	-	-	-	-	-	-	1.011	0.011	0.002	0.000 ***	H5b	supp.
Similarity (Active Contacts) - Gender		-	-	-	-	-	-	-	-	-	-	-	-	1.371	0.315	0.101	0.002 ***	H5c	rev. eff.
Share of Inactive Contacts		-	-	-	-	-	-	-	-	-	-	-	-	2.060	0.723	0.153	0.000 ***	H6	supp.
Active Participation Level		0.967	-0.034	0.012	0.004 ***	0.961	-0.040	0.026	0.121	0.996	-0.004	0.002	0.103	0.996	-0.004	0.002	0.097 *	H7a	partly supp.
Active Participation Change (Delta)		0.996	-0.004	0.001	0.002 ***	0.988	-0.012	0.006	0.040 **	0.997	-0.003	0.001	0.021 **	0.997	-0.003	0.001	0.020 **	H7b	supp.
Verified Membership		0.443	-0.814	0.125	0.000 ***	0.648	-0.434	0.136	0.001 ***	0.621	-0.476	0.099	0.000 ***	0.648	-0.434	0.100	0.000 ***	H8	supp.
Age		1.009	0.009	0.003	0.001 ***	1.002	0.002	0.003	0.538	1.013	0.013	0.006	0.026 **	1.008	0.008	0.006	0.199		-
Gender		0.839	-0.175	0.048	0.000 ***	0.855	-0.157	0.049	0.002 ***	0.672	-0.398	0.195	0.042 **	0.629	-0.464	0.198	0.019 **		-
Time-dependent Effects of Covariates [1]																			
Degree (Active Contacts)	x g(t) [2]	1.044	0.043	0.003	0.000 ***	-	-	-	-	1.016	0.016	0.003	0.000 ***	1.011	0.011	0.003	0.000 ***		-
Avg. Degree of Active Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	0.999	-0.001	0.000	0.001 ***		-
Betweenness (Active Contacts)	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	1.016	0.016	0.008	0.055 *		-
Ego-Network Density (Active Contacts)	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	1.299	0.261	0.145	0.072 *		-
Similarity (Active Contacts) - Region	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	0.643	-0.442	0.109	0.000 ***		-
Similarity (Active Contacts) - Distance	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	0.998	-0.002	0.001	0.060 *		-
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-
Share of Inactive Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-
Active Participation Level	x g(t)	1.011	0.011	0.004	0.006 ***	1.025	0.025	0.013	0.058 *	-	-	-	-	-	-	-	-		-
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-
Verified Membership	x g(t)	1.167	0.155	0.058	0.008 ***	1.154	0.143	0.075	0.057 *	-	-	-	-	-	-	-	-		-
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-		-
Gender	x g(t)	1.100	0.096	0.026	0.000 ***	1.067	0.065	0.030	0.031 **	1.224	0.202	0.081	0.013 **	1.238	0.214	0.082	0.009 ***		-
Subjects				5752				4264				1488				1488			
Failures				4775				3901				874				874			
Observations				55488				31121				24367				24367			
Log-Likelihood				-37784				-29583				-5963				-5882			
AIC				75588				59182				11942				11805			
BIC				75677				59248				12007				11967			

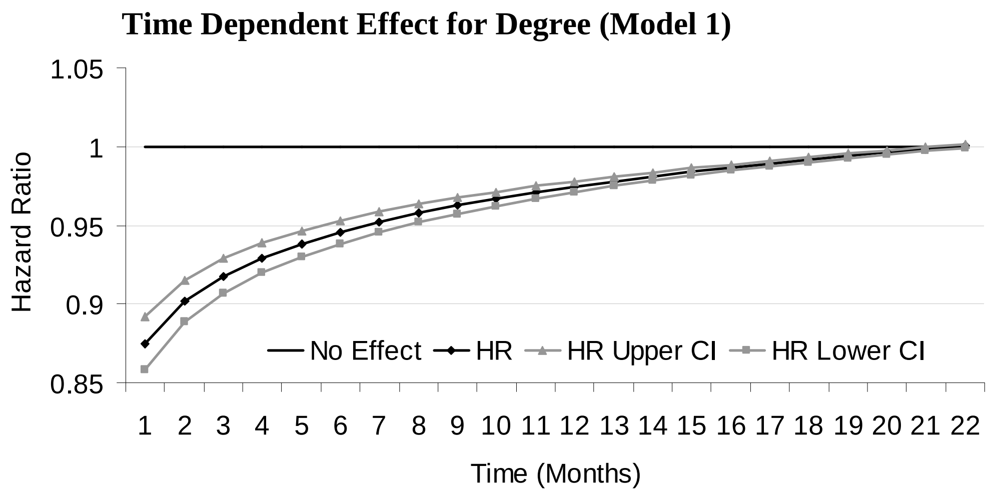
**** p<.01, ** p<.05, * p<.10*

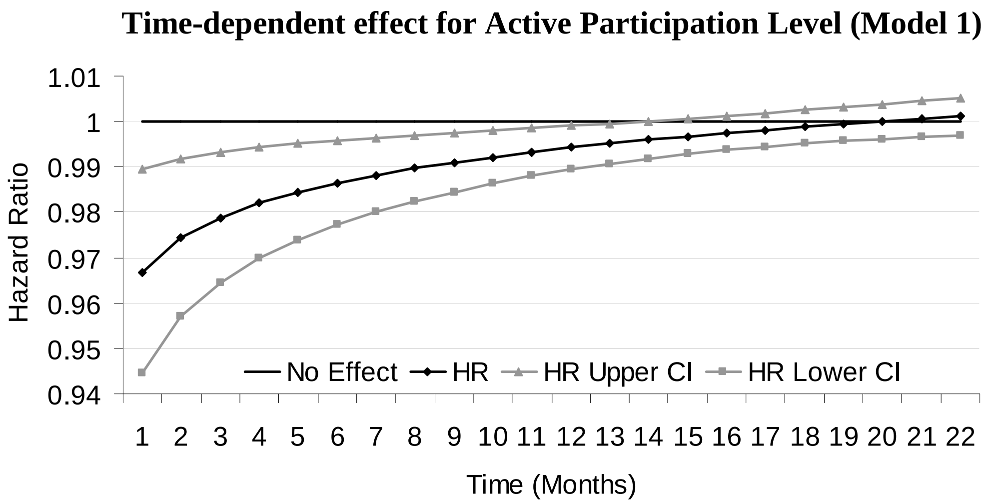
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value; supp.=supported; rev.eff.=reversed effect

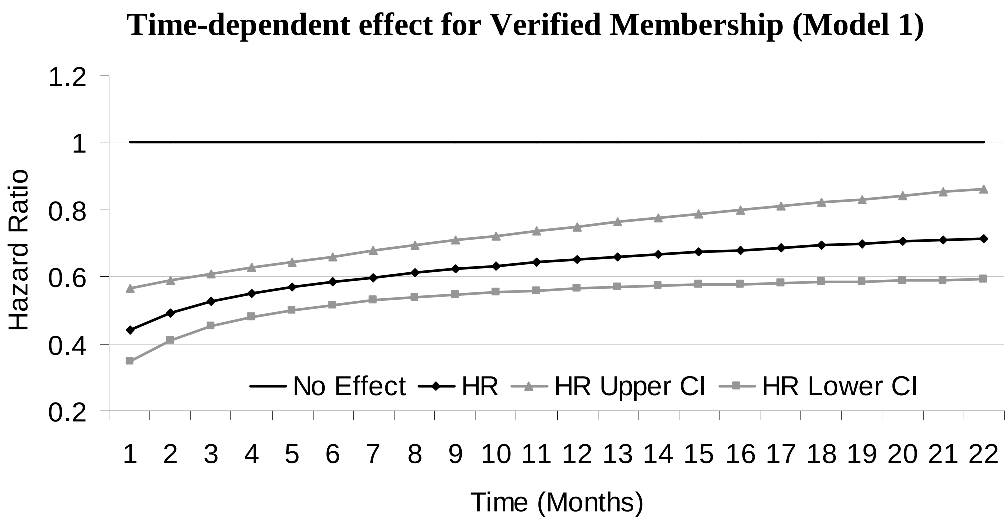
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; PH assumption tests executed for each model seperately; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) g(t)=ln(t)

Table 26: Study 3 – Results of Cox Regression Models for the Effects on User Defection







HR=Hazard Ratio; CI=Confidence Interval (95%)

Figure 17: Study 3 – Time-Dependent Effects of Model 1

Analysis of the Engagement Effects for the Not-Connected Users. Model 2 is based on the group of Not-Connected Users. Because they do not show any social network characteristics, the model includes only the community engagement and control variables. The results, presented in Table 26, show that the change in active participation level and a verified membership have a significantly negative effect on defection. However, the effect of verified membership decreases over time, and results are insignificant for months 7-22. Thus, hypotheses H7b is supported while H8 is only supported for the first 6 months. In contrast to the findings in Model 1, the main effect of the level of active participation is not significant (not supporting H7a). The overall time-dependent effect even indicates that the impact of active participation on defection becomes significantly positive starting in month 10. The time-dependent hazard ratios are presented in Appendix 15.

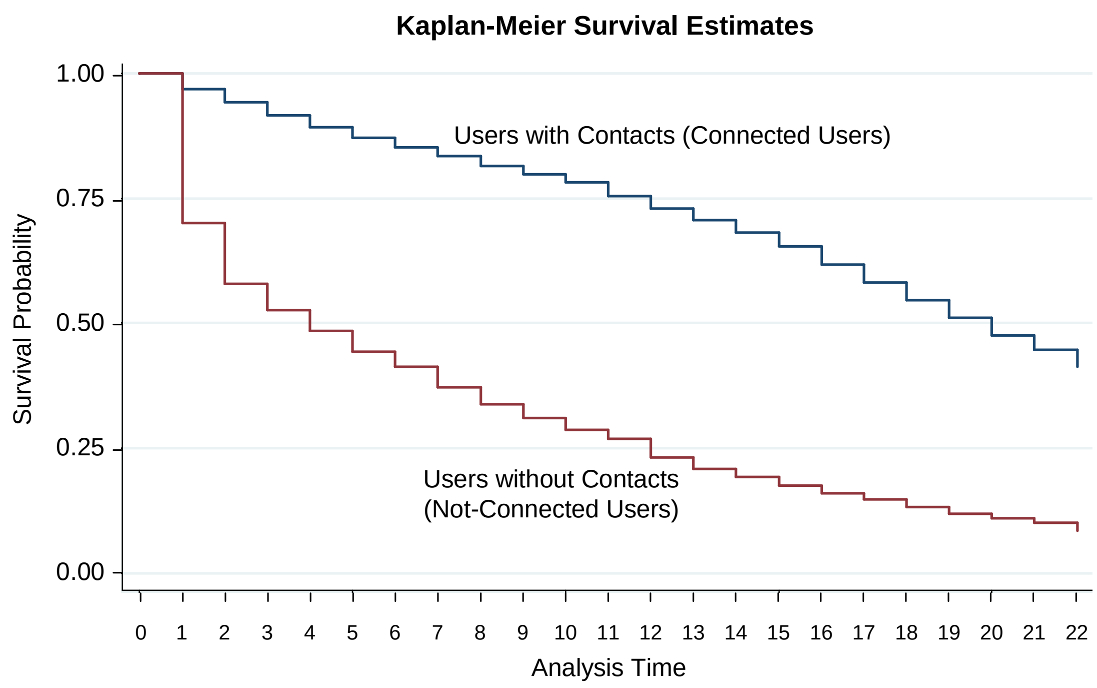


Figure 18: Study 3 – Kaplan-Meier Curves of Different User Groups

Analysis of the Network and Engagement Effects for the Connected Users. As described above, Connected Users are users who have at least one contact, and therefore build the base of the overall online social network. The results of Model 1 suggest that the number of friends is very important to retain the users. In fact, comparing users without any contacts with those who have at least one contact demonstrates that Connected Users have an overall higher survival probability than Not-Connected Users. Kaplan-Meier curves representing both user groups demonstrate this difference (Figure 18), which is also supported by a significant log-rank test.[89] Because the group of Connected Users constitutes the overall social network of the online community and provides data on network positions due to their interconnection with other users, this user group is of specific interest to investigate the impact of the social context on user defection.

First, Model 3a is estimated with the same covariates as in Model 1 for the subsample of Connected Users. The analysis confirms the findings from Model 1, except that the non-proportional effects of active participation level and verified membership are rather proportional (not violating the PH assumption) in this subsample. However, the active participation level is slightly not significant over the entire observation period. In this model H1, H7b and H8 are supported, while H7a is not supported.

Further, the whole set of social network variables is added in the final Model 3b to better understand the impact of social context on defection. In general, Model 3b shows a better model fit with respect to log-likelihood, AIC, and BIC statistics than Model 3a, suggesting that the inclusion of social effects helps to better explain user defection. Table 26 includes the results for Models 3a and 3b.

Despite the inclusion of additional covariates, the effects of degree, change in active participation and verified membership remain significant with only little change in size compared to Model 3a. This confirms hypotheses H1, H7b and H8. The level of active participation also becomes significant when controlling for the network variables, supporting H7a. In addition, all network variables added in Model 3b significantly impact user defection. The average number of the active neighbors' friends has a positive effect on defection. This means that the more friends a user's contacts have, the higher the probability of defection, which describes the opposite direction than hypothesized, leading to a reversed effect for hypothesis H2. Higher levels of betweenness and ego-network density significantly reduce the hazards of defection, which support H3 and H4. Further, a higher share of defected contacts greatly increases the hazards of defection for the Connected Users (supporting H6).

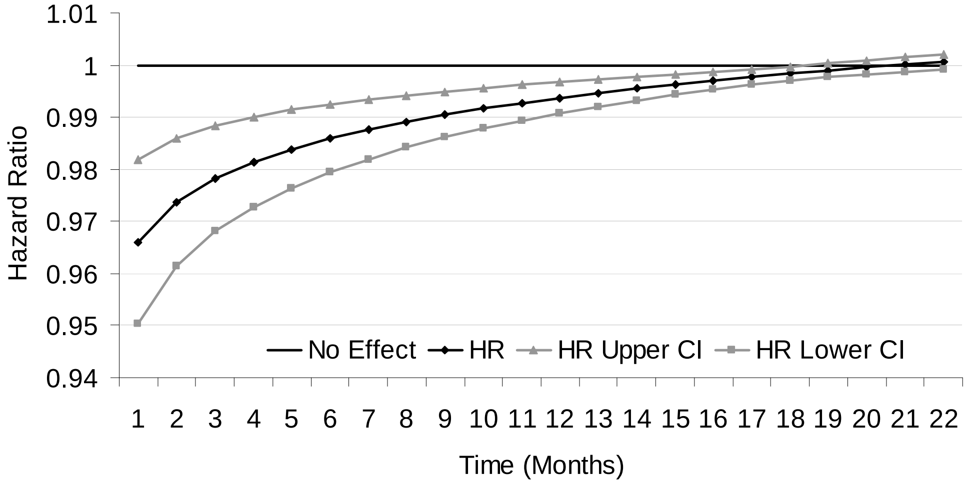
While the average distance between the focal user and his contacts is positively associated with a higher hazard rate (supporting H5b), the impact of similarity in region and in gender on defection provides some surprising results. Contrary to the hypotheses, the hazard rates in Model 3b show that a higher share of contacts from the same region and of the same gender increases the hazards of defection. This would mean that users with more contacts from other

[89] Kaplan-Meier curves represent graphically the estimated survival probability at given points in time. For more information on Kaplan-Meier curves please refer for example to Kleinbaum and Klein (2005).

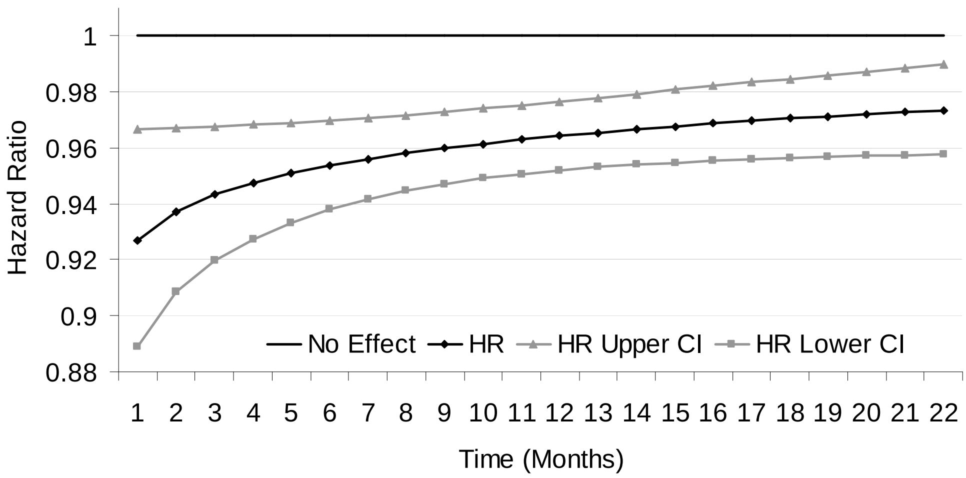
regions and more contacts of the other gender would stay longer. As this is counter-intuitive to some degree at first glance, further investigation with those covariates is pursued later in this study to get additional insights about these effects (see below in chapter 7.3.5).

With respect to non-proportionality, the effects of degree, average degree of active contacts, betweenness, ego-network density, similarity of active contacts in region, and similarity of active contacts in distance are reduced over time, which is indicated by the significant effects of the time-interaction terms. Figure 19 shows the development of effects over time. Degree, betweenness, and ego-network density behave similarly. In the beginning, higher levels of those predictors show a stronger negative effect on defection, i.e. reducing the hazards of becoming inactive. Towards the end of the observation window some effects even turn insignificant. Interesting to see is that betweenness is a more stable influencing factor of defection than degree or ego-network density. While betweenness stays significant through the whole observation period, degree becomes insignificant starting in month 19, and ego-network density turns insignificant already in month 9. For average degree of active contacts, similarity of active contacts in region, and similarity of active contacts in distance the positive effects diminish over time. Here, the positive impact of similarity in distance is significant across all months, while similarity in region is only significant up to month 12, and the degree of the active contacts turns insignificant in month 19.

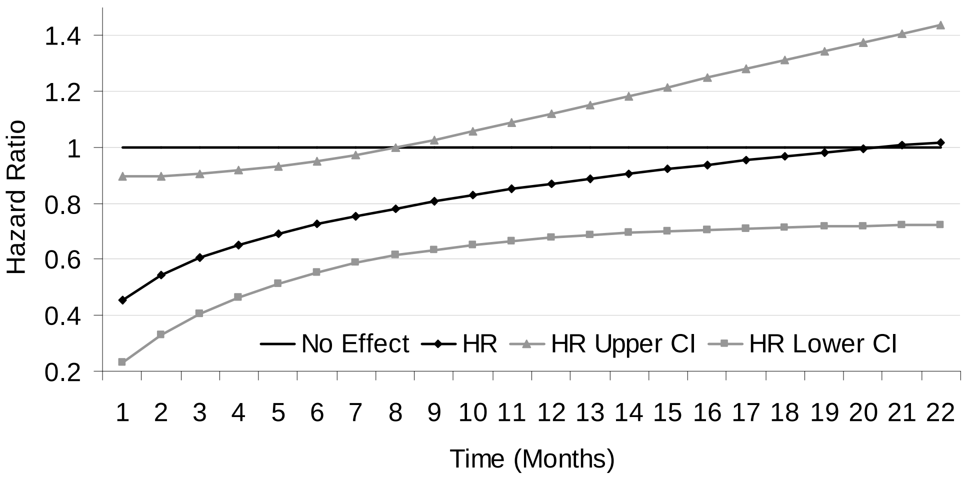
Time Dependent Effect for Degree (Model 3b)



Time-dependent effect for Betweenness (Model 3b)

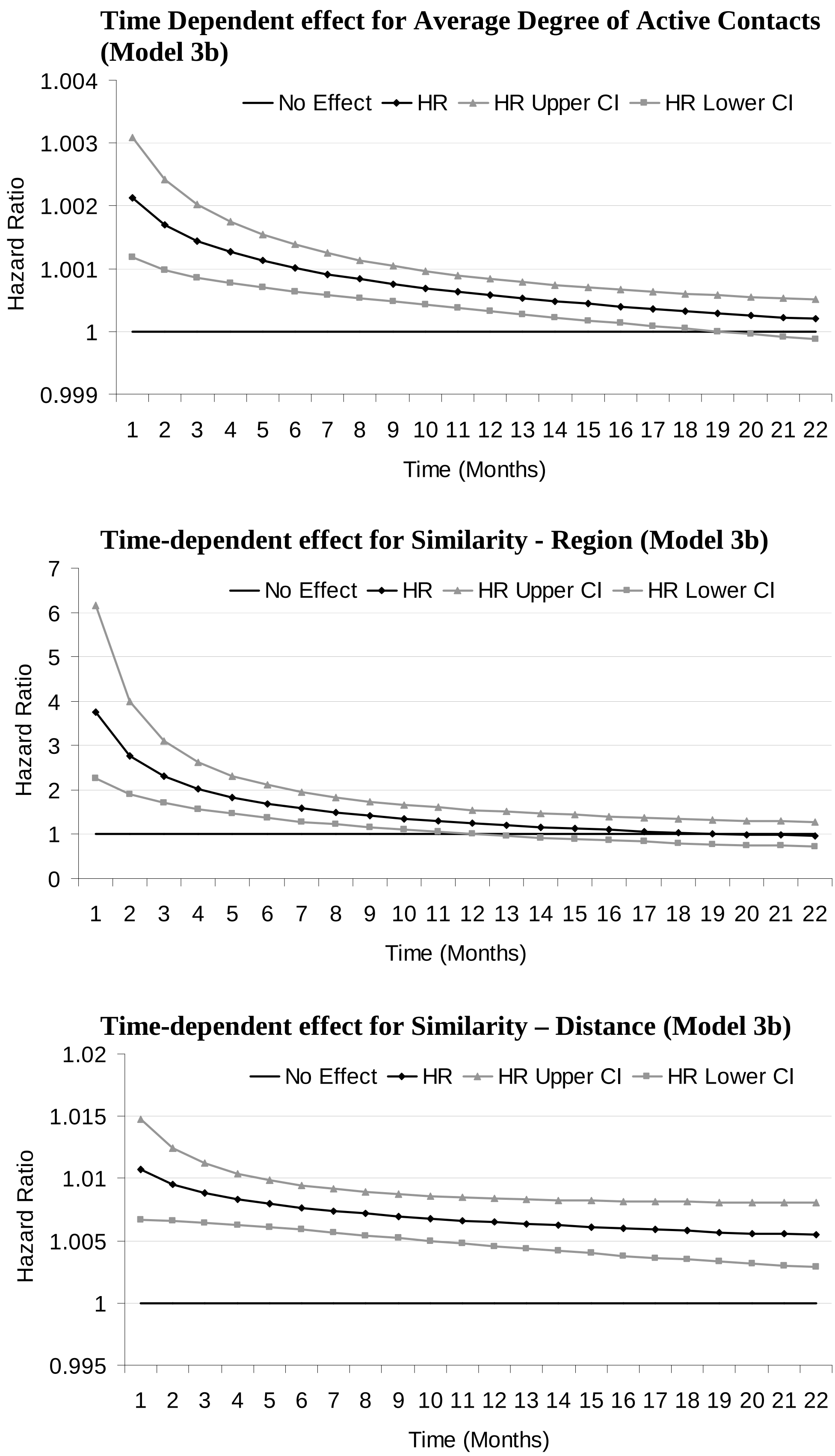


Time-dependent effect for Ego-Network Density (Model 3b)



HR=Hazard Ratio; CI=Confidence Interval (95%)

Figure 19: Study 3 – Time-Dependent Effects of Model 3b



HR=Hazard Ratio; CI=Confidence Interval (95%)

Figure 19: Study 3 – Time-Dependent Effects of Model 3b (continued)

7.4.1 Further Analysis on the Effect of Similarity

The analysis of the Connected User group indicates that the effects of similarity with respect to region and gender are in the opposite direction than hypothesized (H5a and H5c). The hypotheses stated that a higher share of contacts with the same characteristics would lead to higher group identification and higher loyalty intentions, which is supported by social theory (e.g., McPherson, Smith-Lovin, and Cook 2001). However, the specific context of this study might lead to higher loyalty and usage for gender-integrated groups and groups with people from different regions. Therefore, it should be tested if there is a non-linear relationship between similarity in region and in gender and a user's defection. One way of examining the extent of non-linearity is to use indicator (or dummy) variables, where each dummy represents a different group by describing a certain range of values for an independent variable (see for example Collett 2003 for non-linearity testing in survival analysis). For this purpose, three groups of users are defined for each similarity measure: users with high, medium, and low average similarity with their active contacts. These three groups are defined as 0-33% (low), 33-66% (medium) and 66-100% (high) of contacts sharing the same geographical or gender related characteristics. For example, users of the low similarity group for gender have less than 33% of contacts with the same gender as they are. Overall, for gender similarity 39% of the users are in the low, 40% in the medium, and 21% in the high similarity group. With respect to region similarity, 32% are in the low, 9% are in the medium and 59% are in the high similarity group.

An additional Cox PH model (Model 3c) is specified, including two of the described group dummy variables for each similarity measure (region and gender), instead of the continuous similarity variables. The new model is presented in Table 27. The results show that the group of users with low similarity in gender does not show a significantly different hazard rate than users with a good mix of contacts with the same and the opposite gender (medium group). But, having a high share of contacts with the same gender increases the hazard of defection significantly. The latter confirms the findings from Model 3b at least when comparing the effects to the high similarity group in gender.

In the same manner, the low and high similarity group with respect to the user's region are compared to the medium group and show that both high and low geographical similarity leads to a lower hazard of defection, suggesting some kind of u-shaped relationship between geographical similarity and defection. Figure 20 also graphically illustrates the comparison of group effects. The results are discussed below.

Covariates		Model 3c Connected Users - Full HR	Coef.	SE	p
Main Effects of Covariates					
Degree (Active Contacts)		0.971	-0.030	0.008	0.000 ***
Avg. Degree of Active Contacts		1.002	0.002	0.000	0.000 ***
Betweenness (Active Contacts)		0.905	-0.100	0.022	0.000 ***
Ego-Network Density (Active Contacts)		0.804	-0.218	0.124	0.079 *
Similarity (Active Contacts) - Gender - Low Similarity		1.046	0.045	0.103	0.663
Similarity (Active Contacts) - Gender - High Similarity		1.372	0.316	0.102	0.002 ***
Similarity (Active Contacts) - Region - Low Similarity		0.209	-1.568	0.266	0.000 ***
Similarity (Active Contacts) - Region - High Similarity		0.749	-0.289	0.116	0.012 **
Similarity (Active Contacts) - Distance		1.011	0.011	0.002	0.000 ***
Share of Inactive Contacts		2.031	0.709	0.153	0.000 ***
Active Participation Level		0.997	-0.003	0.002	0.106
Active Participation Delta		0.997	-0.003	0.001	0.018 **
Verified Membership		0.651	-0.430	0.100	0.000 ***
Age		1.007	0.007	0.006	0.267
Gender		0.620	-0.479	0.198	0.016 **
Time-dependent Effects of Covariates 1)					
Degree (Active Contacts)	x g(t) 2)	1.010	0.010	0.003	0.000 ***
Avg. Degree of Active Contacts	x g(t)	0.999	-0.001	0.000	0.001 ***
Betweenness (Active Contacts)	x g(t)	1.024	0.024	0.008	0.005 ***
Ego-Network Density (Active Contacts)	x g(t)	-	-	-	-
Similarity (Active Contacts) - Gender - Low Similarity	x g(t)	-	-	-	-
Similarity (Active Contacts) - Gender - High Similarity	x g(t)	-	-	-	-
Similarity (Active Contacts) - Region - Low Similarity	x g(t)	1.476	0.389	0.103	0.000 **
Similarity (Active Contacts) - Region - High Similarity	x g(t)	-	-	-	-
Similarity (Active Contacts) - Distance	x g(t)	0.998	-0.002	0.001	0.047 **
Share of Inactive Contacts	x g(t)	-	-	-	-
Active Participation Level	x g(t)	-	-	-	-
Active Participation Delta	x g(t)	-	-	-	-
Verified Membership	x g(t)	-	-	-	-
Age	x g(t)	-	-	-	-
Gender	x g(t)	1.240	0.215	0.082	0.008 ***
Subjects				1488	
Failures				2874	
Observations				24367	
Log-Likelihood				-5874	
AIC				11791	
BIC				11961	

**** p<.01, ** p<.05, * p<.10*
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; PH assumption tests executed for each model seperately;
time-dependent effects calculated as interaction terms of icovariates with a function of time g(t)
2) g(t)=ln(t)

Table 27: Study 3 – Cox Regression Model with Non-Linear Effects of Similarity

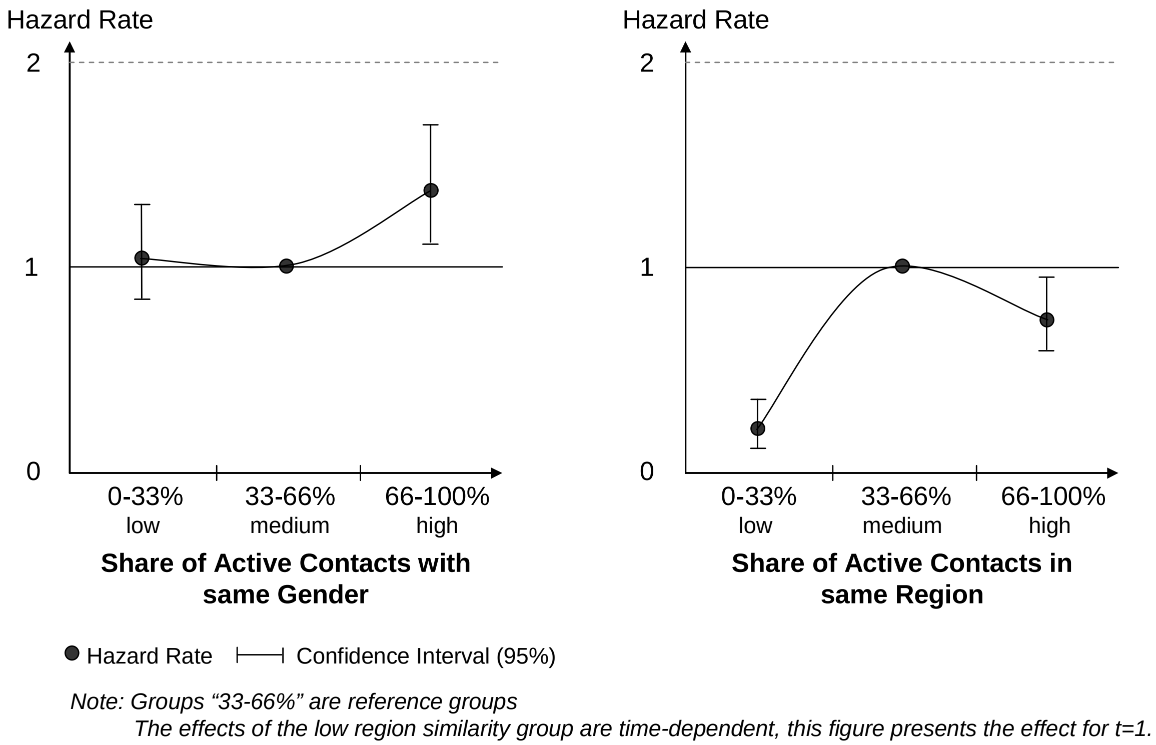


Figure 20: Study 3 – Non-Linear Effects of Similarity

7.4.2 Verification of Results

In order to test the robustness of the results, additional analyses are pursued, which are presented in the following section. The results of all additional analyses in this context are presented in Appendix 16-20. Overall, the findings from the main analyses are confirmed by the verification tests. Although minor differences appear across all robustness tests, most results are similar to the ones gained in the main analyses. In the verification tests, differences in the results for the main variables of the ego-network structure and community engagement are discussed.

Cutoff Threshold. First, the stability of the main results is tested with respect to the selected cutoff threshold for user defection. Remember that a cutoff threshold of three months was chosen because of the best results for classifying active and defected users correctly. Here, the main analyses are executed again with two, four, five, and six months as cutoff thresholds. This means that the identification of user defection is determined differently, according to the last login date before the respective number of months. All users having logged-in after the cutoff date are censored at cutoff. The results of the four additional models using the sample and variables of Models 1, 2 and 3b are compared with the results described in the main analyses above. It is shown that the significance levels and hazard rates are predominantly confirmed by this analysis. The only differences are that in Model 3b the main effect of participation level becomes insignificant for two of the four alternative models, and the time-dependent effect for ego-network density is not existent for three of the four alternative cutoff thresholds. The latter suggests that using a higher number of censored subjects leads to a rather stable proportional hazard rate for ego-network density. However, as the number of uncensored observations increases with the length of the used observation period, the parameter estimates should be more precise in the models with low cutoff values due to the gain in information (Helsen and Schmittlein 1993). Overall, this comparison of different cutoff models confirms the results for the hypotheses tests. See Appendix 16 for the Models 1, 2 and 3b with different cutoff dates.

Function of Time. A second set of analyses is conducted for checking the robustness of the results in cases of different time-functions *g(t)* for inclusion of the time-dependent effects. As the appropriate function of time could not be determined a priori based on theoretical considerations, it is recommended to try differing transformations of time as a test of robustness (Box-Steffensmeier, Reiter, and Zorn 2003). Models 1, 2 and 3b are estimated for the following functions of time: t, $ln(t)$, t^2, and $\sqrt{t}$.[90] The comparison reveals that the models with $ln(t)$ are predominantly better, or at least equally as good, with respect to log-likelihood, AIC and BIC statistics. The results of using different time-functions in the models are similar with respect to the significance levels and directions of the coefficients, though the functional form of the effects of the time-dependent coefficients differs. The only differences are that in Model 3b the effect of active participation level is slightly insignificant for two of the four models, and in Model 2 the time-dependent effect of verified membership is insignificant for two of the three models, though the main effect is still significantly negative. Appendix 17 compares the models with respect to different functions of time. Overall, the results for the hypotheses tests are confirmed by this robustness check.

Validation Sample. In the next test, a second cohort is used to validate the results. All analyses are executed again with users who registered in December 2008 – one month after

[90] Model 2 is only estimated for $g(t)=t$, $ln(t)$, and $\sqrt{t}$, because STATA was not able to estimate the model containing interaction terms of the independent variables and t^2.

the first cohort. Cohort 2 includes 5,497 subjects with 17% censored cases, and the average lifetime across Cohort 2 is 9.24 months. Because the consistency between the two cohorts is high, validation by the second cohort is assumed to be adequate. Appendix 18 shows the results of Cohort 2. Compared to Cohort 1, most results are confirmed. However, there are some differing effects, which are described in the following. In Model 1, the effects for active participation level are not time-dependent (fulfilling the PH assumption), and thus more stable over time. In Model 2, both main and time-dependent effects of participation level are insignificant, while the effect of the delta participation measure on defection is time-dependent. In Model 3b, the effects of similarity in gender and active participation level differ compared to Cohort 1. Similarity in gender is not significant in the validation sample, though the effect is still in the same direction. The effect of the active participation level is positive in the beginning, indicating that higher levels of participation lead to a higher hazard of defection. However, this effect changes over time. In fact, only in the first two months is there a significantly positive effect, which turns significantly negative in month 9 and stays negative afterwards. An explanation for this cohort could be that new users are active and try out the online community in the beginning, but when their needs are not satisfied they leave. Users whose needs are fulfilled stay in the community and the more they invest in terms of participation at later stages, the longer they tend to stay. Therefore, this partly confirms the results for active participation in Cohort 1. Furthermore, the two cohorts differ in some of the time-dependent effects. In Model 3b of Cohort 2, there is a time-dependent effect of verified membership, which is consistent with Models 1 and 2. On the other side, betweenness, average degree of contacts and similarity in region do not violate the PH assumption and are therefore estimated as being stable over time. Overall, because the total effects are very similar, most results are confirmed. However, after taking Cohort 2 into consideration, similarity in gender and participation level do not show consistent results across the two samples.

Stratification by User Group. In order to check the validity of the results of Model 3b, a stratified model with the total sample of users is tested which also includes the network variables. For this purpose, all variables from Model 3b are used in Model 1, but the model is stratified by a binary user group variable, indicating Connected Users or Not-Connected Users, in order to control for the effects of the groups. The stratified model yields similar results as Model 3b. (see Appendix 19)

Active Participation. Additionally, Models 1, 2, and 3b are re-specified with a binary participation variable, which has the value 1 when a user participated at all in the respective month (participation > 0), and value zero otherwise. The effects for the new variable are very similar to the participation level effects. In Model 1, the effect of the binary participation variable is significantly negative over the entire observation period. In Model 2, the main effect is still insignificant (as in the model of the main analysis) and the total effect becomes significantly positive from month 3 onwards.[91] In Model 3b, the effect of active participation (as a binary variable) is insignificant until month 4, but becomes significantly negative (lowering the likelihood of defection) in all subsequent months. This mainly underlines the findings for active participation level. (see Appendix 20)

Degree. Further, it is checked for the robustness of the results by using different covariates. One model is estimated with the logarithmic transformation of degree (ln(degree+1)) rather than the original degree measure, in order to reduce the variance of the variable. However, a transformation of degree leads to a high correlation with the employed betweenness covariate in the main Model 3b, which should be avoided in the regression analysis. Additionally, the hazard rate of degree is easier to interpret when using the untransformed measure. Realizing these disadvantages, the additional analysis is only pursued as a robustness test. Overall, the results from Model 1 and Model 3b can be confirmed. The effect of the transformed degree variable is constant and not time-dependent in Model 3b, therefore it becomes more stable in the full model. (see Appendix 20)

Region. All models are also run with region as an additional control variable. Because the location could have an impact on the behavior of users (e.g., Reinartz and Kumar 2003) due to a higher choice of leisure time events in high population density areas, the Cox regression models are stratified by the user's region to control for such potential effects. All models show similar main effects compared to the main models. The only exceptions are the insignificance of the time-dependent effect of verified membership in model 2 and of the main effect of active participation level in Model 3b. (see Appendix 20)

Influential Users. There appeared to be outliers and influential points in the model data set, but further evaluation of these points revealed no apparent error in calculation or collection of these values. In the main analyses it should be avoided to reject outliers, because the data represents the entire Cohort population in the online community of users who signed-up

[91] Total effects can be calculated from the main effects and time-dependent effects. See also chapter 7.3.3.2.

in November 2008. This also includes users with more extreme values, but represents the true sample. Nevertheless, to test for the robustness, DFBeta analyses (SAS Institute 2008; Therneau and Grambsch 2000) are used to identify influential users in the data set.[92] Users with an influence greater than 10% of the standard error in any covariate of the models are considered as rather influential.[93] An appropriate way to evaluate influential points is to contrast the effects of the main analysis with the data without the identified outliers (e.g., Collett 2003). Removal of these most influential points for each model results in similar effects compared to the main models. The change of active participation (delta measure) becomes insignificant in Models 1 and 3b, and there is no time-dependent effect of active participation level in Model 2. Though all other effects stayed significant and did not change much in size. As the main focus of this study is to investigate the effects of the social context, this is regarded as a minor issue. Overall, the results underline the small influence of outliers and confirm the decision to leave these points in the modeling data set to represent the natural variation that could occur in the online community user base. (see Appendix 20)

7.5 Discussion

As there are a large number of different online social communities in the market and this market becomes more saturated, the task of retaining their users becomes more important for community operators. As in many non-contractual settings, users might defect at some point in time and do not return to the platform without notifying the operator. To better understand which factors influence user defection, different models are used to test for the effects of network position, network configuration and community engagement on the length of membership. Since online communities are a social phenomenon, this study is particularly focused on the social effects within these communities. Past research has already demonstrated the significant role social contacts can take in product adoption (Katona, Zubcsek, and Sarvary 2011) and customer retention (Nitzan and Libai 2011). This research extends past studies by including a more comprehensive set of network variables representing the composition of the users' social networks. It is the first study compiled on social effects and retention in the context of online communities.

---

[92] DFBeta analysis specifies the approximate changes in the parameter estimates $(\hat{\beta} - \hat{\beta}_{(j)})$ when the *j*th observation is omitted (SAS Institute 2008). These variables are useful in assessing local influence.
[93] Therneau and Grambsch (2000) even state that DFBeta values of less than one third of the covariates' standard error are of only little influence.

7.5.1 Network Effects

The inclusion of social network characteristics of the individual users shows a superior model fit compared to the model without social effects. This emphasizes the value of social effects and its impact on customer behavior, which supports recent findings (e.g., Katona, Zubcsek, and Sarvary 2011; Nitzan and Libai 2011). This study demonstrates that the configuration of a user's social network in an online community significantly impacts the length of membership. Network variables with significant effects are degree, average degree of the user's contacts, betweenness centrality, ego-network density, similarity in region, distance and gender, and the share of defected contacts.

Network Position. Overall, the results provide support for the impact of degree, betweenness centrality and ego-network density on the likelihood to defect. Higher levels of those variables suggest a lower hazard to leave the platform. For the user's degree, this is in contrast to the findings of Nitzan and Libai (2011), who show in a contractual context of mobile phone users that a higher number of neighbors increases the hazard of churn. In this study's analysis, a larger, more connected network is related to longer membership duration. Additionally, if the user has a high betweenness, i.e. a central position that connects different users as a "broker", this important position also reduces the hazard of defection. The argument that more social capital leads to higher loyalty explains the influence of degree and betweenness in this context. Having access to more sources of information and social support keeps the users longer on the platform. Further, a closely connected circle of friends, which constitutes strong ties and a cohesive group of people, increases the likelihood to stay. This underlines the network closure argument (e.g., Burt 2000, Coleman 1988), that cohesive subgroups can develop identities and trust which lead to stronger ties towards one's contacts. Contrary to the hypothesis, the degree of active contacts positively effects the defection of users. Despite the decline of the effect over time, it suggests that contacts with a higher degree might have limited time to interact with all of their contacts, which results in limited attention per contact (Katona, Zubcsek, and Sarvary 2011).

Time Effects. An interesting finding of the analysis is that some of the investigated effects are time-dependent and decrease over time. Thereby, the empirical study demonstrates the benefits of exploring non-proportionality in survival analysis. Here, the effects of degree, average degree of contacts, betweenness, and ego-network density depend on time. In the beginning, the effects are stronger, but they decrease over time and in some instances even become insignificant at some point. In the first months, it is important to have many contacts who themselves are not connected to many others and be in a closed group of friends. Getting attention from other users and identifying with a group leads to higher loyalty because the user is able to discover the benefits of the community. At later times of membership, other factors play a more important role. For example, the activity of one's contacts on the platform and interaction with the focal user (e.g., Butler 2001), as well as versatile information from different sources through brokerage (e.g., Burt 2000; Granovetter 1973) could constitute important factors to keep the community attractive at later stages of membership. The latter is supported by the fact that betweenness is significant throughout the observation period. Therefore, the commitment in the first months is much higher when making new friends and communicating over the platform rather than at later stages in time, which could impact the effect to wane over time.

Similarity. The constitution of the network of active contacts, in terms of how similar they are compared to the focal user, reveals interesting results. Social theories often state that higher levels of homophily are associated with stronger ties, which leads to more interaction with other individuals (Granovetter 1973; McPherson, Smith-Lovin, and Cook 2001). Similarity with a defected neighbor also leads to a higher risk of churn for the focal customer (Nitzan and Libai 2011). The analyses in the present study suggest non-linear effects of geographical and gender-related similarity in relation to user defection. The reason might be that high similarity could also induce a negative relation with social capital, because the exchange of new ideas, information and values can be limited in highly similar groups. Less similarity should, therefore, mean greater exposure to a wider range of ideas. At the same time, however, higher similarity may also improve communication, which could also result in curvilinear relationships (e.g., Borgatti, Jones, and Everett 1998).

Users with a high share of contacts of the same gender show a significantly higher hazard of defection compared to users with a more gender balanced social network (medium group). On the other hand, users with a low share of contacts of the same gender do not significantly differ in their hazard to defect compared to the users in the medium group. The reason could be the specific context of the community, where users come together in the community for leisure time events and get to know other people. Because of the topic of the online community in this study, one could assume that a mix of different people in the circle of online friends would be more promising to get versatile information for the focal user. Further, cross-gender friendships are also related to sexual attraction, important in young adulthood (e.g., Halatsis and Christakis 2009; Reeder 2000), which might be a reason for shorter membership duration when interaction only occurs with contacts of the same gender.

In addition, there is also a non-linear effect of similarity in region on user defection. In fact, users with a good mix of contacts from various regions (medium group: 33-66% of contacts from the same region) are more likely to defect than users with few or many contacts from the same region. This is rather surprising from a social capital perspective, as users with both closer and more distant ties would be assumed to have access to different information sources. The hazard rates even indicate that users with few contacts from the same region have a lower hazard of defection than users with many contacts in the same region (though not significantly different). One explanation of this effect could be that users with few friends in the same region use the platform to stay in contact with people they do not frequently meet offline, therefore online communication is the channel of choice to keep in touch with distant ties. Those connections increase the users' social capital as they give access to information sources not permanently available in the offline world, thus having ties to users of other regions might be of specific value. Because users who have many contacts in the same region are less likely to defect as well, it could be supposed that they can identify more with the community, as they see each other offline, strengthening the ties to those people online as well, and build a more cohesive group of friends online, compared to the medium group. This underlines that there exist contrasting social capital benefits of information access and solidarity in a close group (Adler and Kwon 2002). The users in the medium group might be "stuck in the middle", i.e. they neither benefit enough from connecting with distant friends nor build a cohesive network of close friends. However, further research would be required to gain more insights on why certain groups are less likely to defect.

With respect to geographical distance, a larger average distance between the focal user and his contacts has a positive impact on user defection. When considering that users with few contacts from the same region stay longer, the effect of distance would recommend that the distance to users from other regions should be rather low. If the contact is too far away, i.e. the tie to this contact is rather weak, the user does not stay in the community.

Defected Contacts. The results confirm recent findings, that a high number of defected contacts increases customer churn (Nitzan and Libai 2011). Particularly, this study looked at the share of defected contacts in the total group of direct contacts of a user. A higher share of defected contacts is strongly associated with an increase in that user's hazard of defection.

This underlines the importance of social effects and collective behavior (Granovetter 1978), and raises the urgency to retain users, as they strongly impact their neighbors.

7.5.2 Community Engagement Effects

Participation. Both the level of active participation and the change of active participation compared to the preceding months have a negative effect on defection. Although active participation is not significant in some of the (verification test) models, it is often only slightly insignificant, thus still indicating that higher participation is important to keep the users on the platform. This is contrary to the findings of Nitzan and Libai (2011), who showed that in a contractual setting higher usage leads to a higher hazard of defection. However, it underlines previous findings in online community research, that participation positively influences loyalty (e.g., Algesheimer, Dholakia and Herrmann 2005; Casalo et al 2007; Woisetschlaeger, Hartleb and Blut 2008). In this study, interaction with the community reinforces and strengthens ties in the network. In paid services, where interactions generate costs as in telecommunications, the economic benefits of changing providers could be the reason for defection of highly active customers. Further, seen as an indicator for satisfaction (Bolton and Lemon 1999), change in participation level also confirms the positive influence of satisfaction on duration.

In addition, the results show that the effect of participation level is time-dependent in some of the analyzed models. When comparing the user groups, the impact of participation on defection is negative for the group of Connected Users. However, it is slightly positive in the validation sample in the first three months. This suggests that Connected Users who try out the community in the beginning and whose needs are not satisfied leave the community immediately, while established users value the interaction with their contacts, which keeps them using the platform. Because the effect becomes significantly negative after some time, it is then in line with the hypotheses for at least parts of the observation period.

For Not-Connected Users this effect is in the opposite direction, largely insignificant in the beginning, but turning positively significant at a later stage. A reason for the positive effect might be that users without friends try to interact with the community, but without formal friendships there is no social capital, which results in dissatisfaction. Therefore, the users might leave when dissatisfied with the community. This is supported by the effect of change in participation on defection.

Verified Membership. Having a verified membership results in a lower hazard to defect. While for Connected Users this effect is more stable over time, for Not-Connected Users it wanes shortly after registration. A verified membership thus implicates a higher commitment and trust in the platform, especially in the beginning of the membership. Thus, commitment and trust can be important factors for loyalty, but must be nurtured over time, to keep this effect alive.

7.6 Managerial Implications

The knowledge about the social effects on user defection developed in this study can help community operators to identify measures to facilitate user retention. Marketing managers should develop tools integrating the temporal dynamics of users' behavior and users' social networks in order to predict a user's risk of terminating the relationship with the community. Because behavioral and network data are easily available for community operators, those measures provide an efficient way to identify users with a higher risk of leaving the community. Thereby, resources should be spent more effectively on those users who provide high value for the community (e.g., through content contributions or page impressions) and who are more likely to stay or return, rather than users who are not willing to become loyal at all or "whose time has come".

Because users with at least one contact stay longer in the online community, one specific goal is the establishment of social ties with other users, especially in the first month of the membership. Specifically in communities with both user- and firm-generated content, users should be introduced to other members and discussion groups to demonstrate the value of connecting and interacting with other people. For example, this can be done on the landing page after the user logged in to immediately integrate new users into the community. After the users have built online social networks of sufficient size, the operator can help to improve these networks by automatic recommendations on the personalized home page of each user, suggesting users who are outside the close circles of friends to increase betweenness and users who are already friends of one's contacts to increase ego-network density. Further, the system should not recommend users who already have a large number of contacts, as these users are limited in their time available to interact with all of their contacts.

Although this study takes only one community into consideration, there are interesting time-dependent effects which give an indication of which factors are more important along the timeline. Because the effect of ego-network density wanes over time, network closeness should be promoted in the early months of membership to connect with the contacts' friends, while betweenness shows a negative effect throughout the observation period, so that more distant friends should be recommended later in time. In addition, new friendships should be recommended to users with a lower number of contacts, particularly in early stages of the membership. The timeframe of the effects might be different for other communities, depending on the length of the life cycle. Nevertheless, such effects can occur in any similar thematic or regional online community with limited user life cycles. Therefore, community operators should investigate social effects and how they develop over time to be able to identify the right measures at the right time to retain their users.

Another aspect relevant for automatic recommendations to connect with other users is their similarity in region, distance and gender. The findings show that users with few friends in the same regions stay longer. Therefore, users should connect with users from other regions. However, because users in the medium group have a higher hazard of defection, they need to be directed out of this "stuck in the middle" position. Because the results also suggest that a higher average distance between the users and their contacts results in lower membership duration, users from regions close to the focal user's region should be recommended, possibly even to regions the user is already related to through existing contacts of his friend list. With respect to gender, the system should focus on recommendations to connect with users of the opposite gender, especially for users with many same-gender friendships.

To keep the rate of defected users low, it is also important to stimulate user participation. On the personalized home page it could be recommended to reconnect and interact with contacts that have been less active in the past. On one hand, this can increase the overall participation; on the other hand, this could also decrease the risk that the focal user's contacts defect as they are potentially involved in interaction that increases their likelihood to stay. This way, the share of defected users could potentially be reduced. Further, if contacts defected, this should not explicitly be communicated to the user or made visible. If the user does not realize that his friends, or at least part of his friends, defected, the own decision making of staying or leaving would be less influenced and the threshold of collective behavior to leave would be kept artificially at a lower level. However, the exposure to a high number of defectors is an important information community operators should use to approach users at risk to defect.

Further, some users do not ever stay loyal to the community. They register out of specific needs, which might by the consumption of certain information in the community and leave the community thereafter. However, in traditional settings it has been suggested that customers who bought more recently and who already stayed for a longer time with the firm are more likely to become customers again (e.g., Thomas, Blattberg, and Fox 2004). Therefore, users who participated more in the past and more recently should also be more likely to use the community in the future. This is partly supported by the negative impact of active participation level and change of active participation on user defection. However, as there are users who defect, their recapturing should also be on the community operator's agenda. Thus, the operator should focus first on defected users who stayed with the community for a longer time, participated to a higher degree in the past, and who had many friends in the community to convince them to return to the community. Other users, without any participation, low visit frequency and short membership duration, would be less important to spend reacquisition dollars on.

Offline events could also be used to facilitate the identification and connection with more users, with users from other regions, and with users of the opposite gender. Inviting recently defected users and communication with those users could help to recapture them as community members. At least the community operator can try to understand the specific reasons of defected users better when they approach them for reactivation.

Another way to keep not only Connected Users, but also users without contacts using the platform is to provide high quality content and high value interaction with other users. For example, invitations to discussion groups could help to stimulate participation and increase satisfaction with the community. Because 'a short life' can also be due to user dissatisfaction, the community operator should make efforts to satisfy users during the entire life cycle, in the phases of user acquisition, introduction to the community, customer service, and information and interaction provided on the platform.

In addition, promoting verified memberships can help to establish a stronger tie between the user and the community as a whole and increase commitment and trust. Because verified members, particularly of the Connected Users group, have a lower hazard of defection, the operator should provide incentives to become a verified member. This could include access to high quality content, interactive games and raffles, or additional functionality. Because the effect of verified membership becomes insignificant for Not-Connected Users over time, it is important to promote the added value of such a membership in the beginning, but keep the value throughout the membership so that the effect does not wane over time.

Overall, the current study can help to identify measures for community operators to retain current users, convert short-life users into long-life users, and recapture users.

7.7 Limitations and Future Research

This study particularly gained deeper insights on the effects of individual social contexts in online communities on users' membership duration. Although the results are insightful and have interesting theoretical and managerial implications, this research comes with limitations.

This study uses data collected by the community operator on user characteristics and behavior in the online community. However, the intrinsic reasons why users defected are not completely clear. Therefore, information on user motivation to leave the community could provide additional insights on user defection. Because the focus is set on the impact of the social context on user defection, objective measures for the group of Connected Users are highlighted in this study. Attitudinal measures that represent the relational and cognitive dimension of social capital could enhance the understanding of those users defection behavior. Consequently, it is suggested to test a larger set of covariates relevant for all user groups in future studies. Though, it is challenging to collect longitudinal data on the users' attitudes over the course of several months or years.

Although active user participation is included in the current study, passive user behavior could also affect the hazard of defection. Therefore, additional behavioral variables could help to gain more information on the effects of the members' past behavior on defection. For example, the number of logins or page impression could be included in future studies. This might also help to shed some more light on the effects of user behavior of the Not-Connected User group.

Data is used from one specific online community, which is regionally organized and focused on such broad topics as leisure time events and entertainment. In order to test the validity of the social effects found here, different types of online communities could be investigated and compared to each other. In addition, it is unknown how many and which other online communities the members of the observed online community actively use. Being a member of other online communities could significantly influence the decision to churn. Therefore, community external factors could also be relevant in assessing the risk of defection.

8 Conclusion – The Influence of Acquisition, Structure, and Attitudes on User Behavior in Online Social Communities

Online social communities have gained significant attention from marketing scholars and practitioners in recent years. This interest is grounded in the many opportunities these online venues provide for customer research and its' potential to generate economic outcomes. With several hundred million users gathering in online communities, exchanging information, knowledge, ideas, and social support, there is a need to better understand why users participate in these networks. In order to ensure the long-term success, a sufficient number of users are needed, who interact on the platform and actively contribute content to the online social community. Although a large body of literature developed over the past two decades, which investigated the predictors and outcomes of community participation, several research gaps still exist. This dissertation contributes to the existing marketing and online community literature by revealing new insights in how users are influenced in their participation behavior and attitudes towards the community. The social structural context in which relationships and interaction occur is of specific interest. To test the effects of structural, attitudinal and other relevant factors on online community participation and user perceptions, an empirical study was conducted. The empirical study includes three focus areas that each describe a specific point in the membership development process, at which community operators can intervene in order to manage community participation better and increase the attractiveness of the online community. These three stages include the acquisition of users, the activation of users, and the retention of users. In this chapter, the three studies are summarized and the main results are discussed in the overall context of this dissertation.

8.1 Summary of the Empirical Studies

Table 28 provides an overview of the three studies that are embedded in the overall empirical investigation of this dissertation, describing their main characteristics, methodology, and findings.[94] The main results of the three studies are briefly presented in the following.

[94] A detailed presentation and discussion of the studies and its results are presented in chapters 5, 6 and 7.

Study	Dependent Variable(s)	Antecedents	Moderators / Mediators	Methodology
User Acquisition (Study 1)	Satisfaction, Identification, Participation (active, passive, frequency of use), WOM intention, WOM provision (offline, online)	*Acquisition Channels:* a) Marketing Channel - WOM vs. Personal Selling b) WOM Channel - Offline vs. Online WOM	*Mediators:* Social Interaction Value, Information Consumption Value	Survey and Objective Data, ANCOVAs / Multiple-Mediator Regressions
User Activation (Study 2)	Active User Participation	Degree, Betweenness, Ego-Network Density, Share of Real-World Friends, Identification, Reciprocity, Satisfaction, Tenure	*Moderator:* Networkers vs. Non-Networkers (user groups based on social interaction value)	Survey and Objective Data, Zero-Inflated Negative Binomial (Count Data) Model
User Retention (Study 3)	User Defection	Degree, Average Degree of Contacts, Betweenness, Ego-Network Density, Similarity (gender, region, distance), Share of Defected Contacts, Participation Level, Change in Participation, Verified Membership	*Group Comparison:* Users with contacts vs. users without contacts	Objective Data, Survival Analysis - Cox Proportional Hazards Model

Table 28: Overview of the three Empirical Studies of this Dissertation

User Acquisition. The study contributes to existing research by comparing the interpersonal communication channels of WOM and personal selling, and it is the first to examine the difference between online and offline WOM with regards to post-adoption attitudes and behavior. The objective of the study was to answer the following research questions: (1) Do WOM-referred users differ in their attitudes and behavior compared to users acquired by personal selling? (2) Do offline-referred users differ in their attitudes and behavior compared to online-referred users? (3) Do users from different acquisition channels differ in their motivation to use the online community, and do these motivations mediate the effects on user attitudes and behavior?

The study on user acquisition reveals that the type of communication channel through which users are attracted to join the online social community significantly affects the users' post-adoption attitudes and behavior. In particular, the results show that WOM-referred users exhibit significantly higher levels of community identification, participation, WOM intentions and recommendation behavior than those coming from personal selling. Further, offline-referred users show higher satisfaction, participation behavior, WOM intentions and offline recommendations than online-referred users. In addition, it is demonstrated that distinct interpersonal communication channels attract users with different motivations. Users coming from WOM are more driven by the social interaction value of the platform, while users acquired by sales representatives are more driven by the information consumption value. In the same manner, offline-referred users show higher levels of social interaction value. However, there is no difference between offline- and online-referred users in information consumption value. It is shown that these user motivations play a mediating role for some of the observed dependent variables. In both comparisons of acquisition channels, social interaction value mediates (fully or partially) between the channel and all outcome variables. Information consumption value also mediates between the marketing channel (WOM vs. personal selling) for satisfaction, identification, frequency of use, and WOM intentions and behavior. However, it does not mediate any relationship for the two distinct WOM channels. Interestingly, although mediation is considered, most direct effects of the marketing channel on the outcomes stay significant. For the effects of WOM channel on the outcomes, three direct effects remain significant (passive participation, WOM intention, and offline WOM provision).

Overall, users coming from the WOM channel show more favorable attitudes and higher participation towards the online social community compared to users coming from personal selling. In addition, offline-referred users show more positive attitudes and behavior than online-referred users.

User Activation. The aim of the study on user activation was to understand the main influencing factors of active user participation, i.e. the users' interaction and contribution to the community. The study is the first to provide an in-depth analysis of the effects of an individual's position in the network, in combination with the individual's attitudes, on active participation behavior. Thereby, the following research questions are answered: (1) What are the structural drivers for active user participation? (2) How do attitudinal drivers affect active user participation in the presence of structural drivers? (3) How does user motivation affect the relationship of the structural and attitudinal factors on active user participation?

It is demonstrated that network structure – in terms of centrality, ego-network density, and online-offline configuration – significantly affects active participation behavior. A higher number of contacts in the community leads to more active participation. The effect of the ego-network density suggests that within a loosely knit group of friends a user is more likely to actively participate. A higher share of contacts in the online community whom the user knows from the offline world limits the active participation of the users. Further, attitudinal factors yield additional insights, influencing active participation in combination with objective network data. Identification with the community leads to a higher conditional number of activities (if users decided to participate). Reciprocity does not yield overall effects, but higher satisfaction leads to more active participation. Interestingly, when controlling for the number of contacts a higher tenure is negatively related to participation behavior. In addition, the study shows that different user groups are affected on different stages of user participation. Thereby users who show lower levels of social interaction motives are more influenced on whether they participate at all. Users with higher social interaction motives are predominantly influenced in their conditional number of active participations. However, both groups are significantly influenced by their network structures.

User Retention. In the last study, the drivers of why people leave the online community are observed. The study is the first to investigate the impact of social structures in the context of an online community service on the defection of its users. Here, an answer on the following research questions is given: (1) What are the structural drivers for members to defect from the online community? (2) How does community engagement affect the members' defection from the online community? (3) How do the effects on user defection change over time?

The study focuses on the impact of social structures. It shows that the user's network position affects their likelihood of defection. Specifically, a higher degree centrality leads to a lower hazard of defection, but a higher average degree of the users' contacts results in a higher hazard of defection. User centrality, in terms of the users' betweenness, has a negative effect on defection, and also a closely knit ego-network leads to a lower risk of leaving the community. Further, the configuration of the users close network, in terms of similarity with their group of contacts and the share of contacts who already left the community, affects the users' likelihood of defection. Particularly, if a high number of the focal user's contacts are of the same gender, the risk of defection is higher. A lower share of direct contacts from the same region results in a lower risk of leaving the community, though the effect is not linear. If the user's contacts are on average in more distant regions, the hazard of defection is higher. In addition, a higher share of defected contacts of a user's overall contacts leads to a higher risk that he leaves the online social community.

Further, the engagement in the community also plays an important role for the users' loyalty. Users with a higher level of active participation would rather stay in the community, though this effect does not hold for all observed user groups. There is also a positive relation between the change in the users' active participation behavior and the decision to stay in the community, suggesting that an increased activity over the last months reduces the risk of leaving. In addition, members with verified memberships are also less likely to defect.

An interesting finding is that some of the described effects decrease over time. Particularly, of the network measures the impact of degree centrality, the average degree of the users' contacts, ego-network density, and geographical similarity wane over time to become insignificant at a certain point. This means, that in the first months after registration, the effects are higher, and the impact of the respective variables is thus higher for the hazard of defection.

Overall, these results provide valuable insights for researchers and practitioners. It is shown that some of the results can be confirmed across studies. Other factors reveal different effects according to the research focus of the study. A more detailed discussion of the dedicated results with respect to the single studies can be found in the respective chapters 5, 6 and 7. A general discussion on the interrelationship of the findings from all studies is provided in the next section.

8.2 General Discussion

8.2.1 Discussion of the Results for Online Social Community Research

On all three stages of the membership development process, the consideration of the social context of the individual users in the online community helps to explain the users' attitudes, perceptions, and behavior. In this respect, the application of social theories in the context of online social communities affirmed to be beneficial. In particular, the empirical studies show evidence that attitudes and behavior are affected by the social relationship between the actors in interpersonal communications for user acquisition, the position of the users in the online social network, the configuration of the users' online social networks, and the users' perceptions of their relationship to the community. This illustrates that the individual's social context should not be ignored when researching and managing an online social community. As social capital theory emphasizes, there lies specific value in the relationships to other community members. Particularly the structural and relational dimensions of social capital help to make sense of the users' behavior. Consequently, the observation of network structure and relationships between users can yield important insights for marketers in settings with user interaction.

Network Position. The observation of the structural dimension of social capital provides valuable results, which suggest that information on an individual user's position in the social network has significant effects on his behavior. The users' degree centrality is one of the central indicators of higher active participation and loyalty to the online social community. The social capital that resides in the connection to many other users leads to positive behavioral outcomes. Thereby, the number of contacts is a simple but effective measure to consider one facet of the users' network structure.

In addition to the number of contacts, the concepts of network closure and brokerage play an important role in social capital theory (Burt 1992; Coleman 1988). In line with the theoretical argument that brokerage yields information and control benefits, while closure yields solidarity benefits, the results of this study are twofold: they show that a strongly knit (dense) network of contacts has positive effects on membership loyalty, while it has a negative effect on active participation. Burt (1992; 2000) argues that network closure has the advantage of generating trust and belongliness among users, leading to group cohesion, while on the other hand it increases the amount of redundant information circulating in this group. Therefore, the establishment of trust and norms in a closed group lead to more loyalty. However, the functional benefit of having access to non-redundant information from distant groups of people increases the interaction of the users, which favors less connected groups of contacts. It becomes clear that both effects need to be considered. As Adler and Kwon (2002) point out, both concepts contribute to social capital, but their benefits are dependent on the situation and perspective of research. Therefore, researchers as well as practitioners need to understand both perspectives and its effect on behavior in different situations. There is one interesting finding in the empirical study of this thesis that relates to the situational aspect: time-dependent effects of ego-network density suggest that only in the first couple of months a closely connected group of contacts has a significant positive effect on loyalty. This means, that being a new member of the community represents a situation where solidarity benefits are more important. Thus, situation is an important aspect, which should be attributed further consideration in future online community research.

As another measure of centrality, betweenness indicates brokerage opportunities for the focal user. Therefore, higher betweenness should be associated with access to new information and ideas and lead to higher social capital. Indeed, being in a more central position (in terms of betweenness) reduces the risk that a user leaves the community. Here, the access to resources and higher social capital can explain loyalty behavior. However, the empirical study did not show a consistent significant effect of betweenness on active user participation.

These two sides show that betweenness, as a source of social capital, can explain general user behavior to return to the online social community better than active participation behavior. For the latter, degree has been demonstrated to be a better, more robust measure of centrality in explaining user contributions.

Overall, this thesis highlights the relevance of structural indicators, which are missing in most of the past studies on online community participation (see the literature review in chapter 2.5). Because user behavior can be explained better when using structural measures than without them, this should become a standard procedure when investigating social environments. Therefore, online social community research should not miss to include information about the user's position in the network when investigating user behavior.

Network Configuration. But not only the network structure, also the configuration of the users' social networks is an important facet to understand their behavior. Because online and offline social networks are often intertwined (Wellman and Hampton 1999), the connection to the offline world should be regarded explicit recognition, particularly for locally organized online social communities. The results of the empirical studies show that a high share of contacts, who meet the focal user regularly in the offline world, leads to less active participation. Similarly, a low number of contacts in the same region increases the likelihood of staying in the community.[95] In general, these results suggest that users with less offline contacts perceive high social capital because of the access to more diverse resources, while users with more offline contacts perceive a reduced level of online social capital because the resources are also available through offline channels. With respect to the functional benefits of social capital, the empirical findings indicate that online-only contacts supplement social capital available in the offline world and increase participation, but offline contacts substitute sources of social capital and consequently decrease user participation. This highlights that online social communities are specifically valuable when maintaining distant and weaker ties through the online channel. If social capital is available in the offline world, this would result in less social capital benefits online with the same people. Social presence theory supports this view because certain activities, like receiving social support or solving problems, are pursued preferably through media with richer communication and social context cues (if the medium is available). Thereby, online social capital is undermined by the offline contact in

[95] Users with a high number of contacts from the same region also tend to stay longer than those with about equal numbers of contacts from the same and different regions. However, the hazard of defection is lower for the group with a low share of people from the same region compared to those with a high share of people from the same region. For further discussion of this effect see chapter 7.4.1.

situations where online and offline overlap. However, there might be situations where solidarity benefits are valued more, which can arise from being in a more cohesive group in terms of offline interaction. This is demonstrated by the effect that users with a high share of contacts from the same region also tend to have longer memberships in the community (though the likelihood to stay is lower than for the group of people who have a lower share of contacts from the same region). Knowing many people from the same region, whom they potentially meet offline more often, leads to higher loyalty to this group of people in the community. But the reduced amount of new information in such groups results in lower active participation behavior.

Other variables to describe the network configuration help to understand what kind of network configuration is more beneficial. Because users with less friends of the same gender stay longer, this suggests that contacts of the opposite gender are favorable to increase participation. In that case, it is not the homophily argument that explains behavior, but more importantly the access to specific resources which might by of functional or emotional value for the user. In the same manner, knowing how many of the user’s contacts already left the platform indicates the availability of the remaining users for exchange of resources. The collective behavior argument helps to provide additional explanation for the risk of defection from the online social community because of reduced benefits that are accessible. Both examples tested in this thesis support social structure as a source of social capital. According to the specific context of the online community, the situation arises in which more users from the opposite gender are more favorable contacts, while in other communities this could different. Thus, sensitivity to the context of research is important.

Attitudes and Perceptions. Attitudinal factors and user perceptions can yield additional insights for researchers and marketers. In general, it can be confirmed that attitudes, representing the relational dimension of social capital, are important to understand user behavior, although objective data is used in the analysis.

First, the study shows that satisfaction plays a key role. Higher satisfaction levels are associated with more participation. If the expectations of the user are positively confirmed, he participates more actively and stays longer in the community. Second, identification with the community increases the attachment and the interaction among members. This underlines the theoretical contribution of social identity theory for online communities. People who identify with the group show favorable behavior in the group, i.e. they actively participate and stay in the group. Third, reciprocity shows mixed results depending on the stage of the participation process and the user type observed. Reciprocity works for people who usually contribute few content. However, for users who regularly contribute reciprocity does not explain activity. This inconsistency in results is in line with the findings from the literature review (chapter 2.5.4.2) that reciprocity might be a more complex process which relates to other factors. Therefore, social exchange processes do not consistently and directly explain the users' participation behavior. Fourth, motivation is an important source of social capital (Adler and Kwon 2002). The empirical studies show that, based on different motivations, users are differently affected in their behavior. Knowing that, online community operators need to better understand what types of users exist in the network and how they need to communicate to different user segments in order to increase the community's success.

User perceptions and attitudes should be taken into account when studying user behavior in online communities. It complements the structural perspective because it influences the users' motivation to exchange resources.

Social capital as an overarching concept. Altogether, the presented research supports the existence of different dimensions of social capital. Social capital is found to be an adequate overarching concept that is related to other streams of social concepts and which helps to understand user behavior. Under the umbrella of social capital, social network analytical concepts are appropriate to describe the structural dimension, while social identity theory and social exchange theory can be utilized to explain the relational dimension. Concepts of collective behavior and social presence also help to understand how certain aspects of the network configuration relate to human behavior. Although the latter concepts are rather indirectly related to social capital, they can enrich the argumentation on social capital benefits.

This thesis underlines the importance of studying social networks and social capital in an online community context. It extends past research perspectives by looking more closely on the social context of user relationships and the structural dimension of social capital to understand user behavior. A major finding is that both structural and attitudinal factors are relevant for understanding behavioral outcomes and community success. A combination of different types of data, representing different facets of social capital, turned out to be beneficial for online community research and to develop managerial implications. Social capital theorists acknowledged that different sources of social capital can be intertwined (e.g., Adler and Kwon 2002; Nahapiet and Ghoshal 1998; Tsai and Ghoshal 1998). However, further empirical evidence of the interconnections between network structure, network configuration, attitudes and motives and their impact on user behavior in online communities should specify these relationships. In this thesis, it is demonstrated that there is a relationship between motivation and attitudes and behavior, and to some degree there is evidence that the structural, relational and cognitive dimension are related. The results show that effects of attitudes on behavior are affected by the inclusion of structural aspects in the analysis. Also differently motivated groups of users are differently affected in their behavior. Because it was not the focus of this empirical study to explain all possible interrelations of the different dimensions and variables, it is proposed that future research should take a stronger focus on investigating such effects. This need is also addressed in chapter 8.3. This could advance the understanding of which factors potentially take mediator or moderator roles and which dimensions influence each other.

Further, the discussion on offline-online network configurations suggest that social capital in online communities depends on the interrelation between online and offline social networks. It is not only the interrelation between the different dimensions of social capital in the online context, but also the relationship between online and offline contact to one's network of friends which determines user behavior. This indicates a relationship and overlap between online and offline social capital which should be investigated in more detail in future studies. Specifically the utilization of relationships to people in either the offline or the online world for specific purposes could yield additional insights in which context the sources of social capital work better or supplement each other. Therefore, a more holistic investigation of social capital and its effect would include both online and offline connections.

Interpersonal communication channels for user acquisition. In addition to the findings on how social capital and related theories influence user behavior, studying interpersonal communication channels to acquire new users to the community highlights that the social context of the communication channel affects attitudes and behavior. Channels with a stronger relationship to the sender of the marketing message accentuate social motives. This results in positive attitudes to the community and increased social interactions, which also describe to a certain degree how much the users value their relationship with the community. Thereby, channels which bring user with higher social motives to the community, also drive the development of social capital in the community through these users. It is therefore important to understand the distinctive characteristics of these channels. This thesis shows that WOM more likely arouses social motives for community usage than personal selling. In the same way offline-referred users are more driven by social motives than online-referred users. Although distinct user motivations are accentuated in different channels, the direct effects on certain variables suggest that the social relationship between the sender of the marketing message and the receiver also explains the user's attitudes and behavior on the platform to some degree. The differentiation of the effects of personal selling, online and offline WOM supports the theoretical perspective of different social contexts that influence the user's behavior after registering on the platform. Therefore, WOM in general and offline referrals in particular yield more favorable relationships to the community that result into social capital achieved on the platform. This means that the generation of social capital already starts with the acquisition of the "right" users, who are willing to connect to others and the community.

Altogether, the three studies show that a) the social context yields important insights on how users behave in online social communities, b) individual network positions are important drivers of participation, c) the configuration, particularly with respect to the interconnection of online and offline networks, shows a relevant effect on behavior, d) attitudes and perceptions can significantly contribute to a further understanding why users are more active, e) user groups differ in their motivations, leading to different attitudes and behaviors towards the online community, and f) the acquisition channel already influences social capital development on the platform.

What arises from an overall view on the membership development process is that the findings at all stages are interdependent. Social capital represents a valuable umbrella concept which can be enriched by appropriate social theories to comprehensively understand user behavior. However, specific types of user attitudes and participation are affected by different social capital benefits, where there are situations in which one or the other benefit (be it either information, control, or solidarity benefits) is more important.

8.2.2 Discussion of the Results for Online Social Community Management

In addition to the theoretical discussion of the results, community operators are interested in its practical relevance. For community operators, the results of the empirical studies have several implications. In fact, on each of the three stages in the membership development process specific actions can be taken to manage the community more effectively and increase the success of the community. More detailed recommendations for community management are discussed in the dedicated implications sections of the respective empirical studies (chapter 5, 6 and 7). Here, general implications are discussed. Overall, the results suggest that community operators should pay explicit attention to the network structure of their online community, assess the motivation and attitudes of their users on a regular basis, make use of different acquisition channels and integrate their marketing activities on all stages of the membership development process.

A major finding is that a user's position in the network significantly affects the behavior in the online social community. Therefore, community operators need to consider network measures when tracking the state of their online social community. The advantage of objective network data is that they are easily accessible for the operators of online social communities and they help to understand how the users are connected. Further, they indicate which users are more likely to actively participate and which users are more likely to defect. Thus, network measures may help to manage participation through steering the interconnection of the users on the platform. The underlying community system can be used to recommend users to connect to specific other users or interact with certain users from their contact list. Personalized recommendations for connection and interaction with other users should be based on the results of how specific network configurations affect participation (e.g., higher number of users from the opposite gender). Then, they can be integrated, for example, on the personal home page of each user individually. This way, the level of participation could be increased and the risk of defection decreased. Further, on the basis of the users' positions in the network, the integration of new users can be assessed and defected users can be prioritized for win back activities. Highly connected users are important to the platform, which suggests including network measures as performance indicators into marketing and customer management.

In addition to the objective network data, this study has shown that self-reported measures, such as motivations, attitudes, and perceptions, can significantly improve the understanding of what influences participation behavior. Therefore, such factors should be used in community management to assess the level of attachment to the community and to differentiate users with specific attitudes and motivations. In particular, motivations are important discriminators for user behavior and the effects on user participation. A segmentation of users along their motives can provide more targeted user communication as well as better-customized content on the platform. The introduction of a 'mini user survey' on the platform to regularly track user motivations and to lay a basis for segmentation would be a promising tool.

Further, acquisition channels can be used to attract different types of users, with different motives and needs. Having a diversified portfolio of user groups in the online community can help to distribute the risk of being dependent on one specific user group. For example, social networkers use the community to interact and communicate with other people. Although being a highly valuable group for the community, if the interaction goes down and people leave, the value would be lost for such networkers as well. Having also users on the platform who are predominantly interested in the content (be it firm- or user-generated) can stabilize the traffic on the platform. It is not asserted, that there is a right or wrong acquisition channel. Moreover, the community operator needs to know what channels are available and how they affect their current customer base. WOM referrals have been found of specific importance as a cost-efficient channel that attracts more actively participating users. Therefore, referral programs should be used to stimulate WOM recommendations. Specifically offline WOM recommendations attract users with favorable attitudes and behaviors towards the online social community. Triggering offline WOM would therefore be of high relevance when there is a need to attract users who are more willing to socially interact.

Overall, community operators need to take actions and continuously improve on all three stages of the membership development process. They need to attract new users, activate the users to participate actively, and retain the users. In order to prioritize users and focus on users of specific value for the platform, structural and attitudinal data, as well as motivations should be included to design marketing tools. For example, knowing which user comes from which channel and how many contacts a user has can help to integrate them better in the community and also evaluate the risk of defection. Thereby, the effectiveness of certain marketing campaigns can be evaluated, and marketing budget can be spent more efficiently.

8.3 Future Research Directions

The focus of this chapter is to elaborate on some future research directions, which potentially can yield deeper insights on certain aspects of online community research.

Overall, the empirical studies presented in this dissertation are based on data coming from one online social community. The online community used in this research has a broad topic and is thus expected to provide rather generalizable results. Although using one community has the advantage that there is no undesirable variances between users of different communities due to varying functionality and purpose, the effects should be verified in other online community settings. Further, with respect to the effectiveness of different acquisition channels, the potential differences of online and offline WOM should be tested in other contexts. Online social communities provide a specific setting, which is mainly based on the social context of the users. This study particularly acknowledges this social context, because it is focused on understanding its effects in online communities. However, it would be interesting to see how the two types of WOM perform for other products and services.

The empirical studies investigated the direct effects of structure, attitudes, and engagement on the respective outcome variables. In addition, some mediating and moderating effects have been revealed. However, it was not feasible to look into all possible relationships between the variables of interest. Specifically, the results suggest that there might be some relationships between structural and attitudinal factors. The interrelation of measures on the users' position in the network, the configuration of the network, the relationship to the online community, and the individual perceptions of the users is a potential area of future research. Further, mediation and moderation effects of the structural and relational dimensions should be investigated. A combination of the facets of social capital in the online and offline context could support the progress in research on different effects of online social capital.

It has been demonstrated that motivations play an important role and shape the users attitudes and behavior. However, because of the unavailability of data, it was not possible to test how motivation affects defection. This can provide additional insights and help to better understand why users leave their online community.

The research focus in this study was to investigate the individual user's relationship to the whole group of people in the online community. However, dyadic relationships could provide additional information. For example, a user's tie strength and homophily to single other users in the community and its effect on behavior towards these individual users could deepen the understanding of why users interact with specific other users more intensely. Although this was out of the scope of the current dissertation, it could be an interesting future research perspective. Further, the hypotheses in this work are tested on an individual user level. Another approach is to observe sub-groups or sub-communities and their effect on the overall participation level and community success (see for example Toral et al. 2009 for community level observations). Linking structural dimensions of micro- and macro-level effects could provide community operators with knowledge on how the entire online community, and sub-communities therein, need to be structured. Further, an understanding of how the users within sub-communities are affected by their position and relationships towards other users could be achieved.

In this dissertation, longitudinal data has been used to investigate certain effects on the users' participation. However, few researchers have investigated longitudinal effects and life cycle models regarding online community activity. Therefore, there is a need to achieve further empirical evidence on how users develop in terms of their active participation in the online community over time. This can generate knowledge of how social capital is utilized in different situations and different points in time. A better understanding of why user behavior changes and which economic effects this has for the community operator could be of specific interest. In fact, linking user behavior to financial outcomes in online communities has only rarely been observed (e.g., Trusov, Bodapati, and Bucklin 2010). This could help to prioritize marketing budgets and activities, as well as to differentiate and personalize marketing messages towards different user groups.

Further, experiments are a popular instrument to test theorized or empirically discovered hypotheses. Online communities have not been subject to many experimental studies on active user participation so far. With this lack of experimental research it is proposed to test certain aspects of community participation by generating experimental conditions and alter the relevant factors to test and understand the 'real life' effects for individual activity changes.

Finally, as there are many different types of online communities in the market, a comparison of the structural and relational effects on user participation can be pursued across different kinds of online communities. Cross-community studies could enrich the understanding of how users behave and how they are influenced in different communities.

Appendix

Appendix 1: Literature Review of Online Community Participation Studies

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Algesheimer et al. (2010)	Marketing Science	Participation in community *(binary variable: participation / no participation)*	Nationality, Age, Gender, Membership length, Positive feedback, Negative feedback, E-Mail invitation to participate in the community	Customer buying and selling behavior: # Bids, Amount Spent, # Listings, Revenues	eBay Community for “collectibles” product category, data set of 13,735 individual customers
Bagozzi and Dholakia (2002)	Journal of Interactive Mktg.	Desires (to interact) We-Intentions (to interact)	***of Desires:*** Attitudes, Positive anticipated emotions, Negative anticipated emotions, Subjective norms, Perceived behavioral control, Past behavior, Group norms, Social identity *(self categorization, affective commitment, group-based self-esteem)* ***of We-Intentions:*** Desires, Perceived behavioral control, Past behavior, Group norms, Social identity *(self categorization, affective commitment, group-based self-esteem)*	***of Desires:*** We-Intentions	Virtual chat room, 157 usable responses
Casalo, Flavian, and Guinaliu (2007)	Online Information Review	Participation		Trust in free software (honesty, benevolence, competence), Loyalty (tow. free software products)	Several free software virtual communities, 215 usable responses
Casalo, Flavian, and Guinaliu (2008a)	Mgmt. Research News	Commitment to the community (Perception of active participation)	Trust (Honesty, Benevolence, Competence)		Free software virtual communities, 163 usable responses
Casalo, Flavian, and Guinaliu (2008b)	J. of Mktg. Communi-cations	Participation	Trust	Commitment (affective commitment to the brand)	Several virtual brand communities, 215 usable responses

Appendix 1: Literature Review on Online Community Participation Studies

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Casalo et al. (2009)	Industrial Mgmt. and Data Systems	Participation Intentions	Reputation, Satisfaction	Affective commitment to the OSS	Open source software online social network, 215 usable responses
Chen (2007)	Journal of Information Science	Continuance intention	Post-usage social interaction ties, Website use satisfaction		Professional virtual community on programming, 360 usable responses
Chen and Hung (2010)	Information & Mgmt.	Knowledge Contributing Behavior, Knowledge Collecting behavior	Norm of reciprocity, Interpersonal trust, Knowledge-sharing self-efficacy, Perceived relative advantage, Perceived compatibility	Knowledge utilization, Community promotion	Two professional virtual communities, 323 usable responses
Cheung and Lee (2009)	Journal of Information Science	Intention to continue using	Satisfaction, Commitment, Group norms		Virtual community for teachers and educators, 315 usable responses
Cheung and Lee (2010)	Decision Support Systems	We-Intentions to use social networking site	Subjective norm, Group norm, Social identity (cognitive, affective, evaluative)		Facebook, 389 usable responses
Chiu, Hsu, and Wang (2006)	Decision Support Systems	Quantity of knowledge sharing	Social interaction ties, Trust, Norm of reciprocity, Identification, Shared language, Shared vision, Personal outcome expectations, Community-related outcome expectations		IT oriented professional virtual community, 310 usable responses
Chu (2009)	Internet Research	Helping behaviors (Information sharing, Knowledge contribution)	Social Capital (Network, Norm, Belief, Trust), Online community characteristics (Size, Diversity, Ancillary Resource, Role of members), Sense of community		Nine different online communities, 355 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Chung and Buhalis (2008)	Information Techno-logy & Tourism	Participation	Information acquisition benefits, Social-psychological benefits, Hedonic benefits		Online travel community, 217 usable responses
de Valck et al. (2007)	British J. of Mgmt.	Members‘ Visit Frequency	Membership duration, Satisfaction with: - member-to-member interaction, - organizer-to-member interaction, - organizer-to-community interaction, - community site		Virtual community of interest, 3605 usable responses
de Valck, Van Bruggen, Wierenga (2009)	Decision Support Systems	Frequency of visits, Duration of visits, Retrieve information, Supply information, Discuss information		Perceived Influence on: - cooking frequency - recipe knowledge - recipe choice - satisfaction with result	Virtual community dedicated to culinary matters, 1007 usable responses
Dholakia, Bagozzi, and Pearo (2004)	Int. J. of Research in Mktg.	Desires (to interact) / We-Intentions (to interact)	**of We-Intentions:** Group norms, Mutual agreement, Social identity (cognitive, affective, evaluative), Desires **of Desires:** Social identity (cognitive, affective, evaluative), Mutual agreement, Mutual accommodation, Entertainment value	**of We-Intentions:** Participation Behavior **of Desires:** We-Intentions	Diverse (9 Types of online communities) 465 participants (of 264 different communities)
Dholakia et al. (2009)	Journal of Service Research	Helping oneself, Helping others	Functional benefits, Social benefits		Firm-hosted virtual peer-to-peer problem solving (P3) community, two data sets: 2,299 and 204 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Han, Zheng, and Xu (2007)	HICSS '07	Intention to stay, Willingness to help, Willingness to spend time	Social needs, Self expression needs, Trust, Identification, Control variables (Age, Gender, Web experience, Tenure)		Cyworld community, 90 usable responses
Hennig-Thurau et al. (2004)	Journal of Interactive Mktg.	Frequency of visits, # of comments written	Platform assistance, Venting negative feelings, Concern for other consumers, Extraversion/ positive self-enhancement, Social benefits, Economic incentives, Helping the company, Advice seeking		Web-based opinion platforms, 2,063 usable responses of actively participating consumers
Hsu et al. (2007)	Int. J. of Human-Computer Studies	Knowledge sharing behavior	Identification-based trust, Personal outcome expectations, Community-related outcome expectations, Knowledge sharing self-efficacy		Nine different types of virtual communities, 274 usable responses
Jin et al. (2009)	Computers in Human Behavior	Continuance intention (keep members continue using information)	Information usefulness, Satisfaction		University Bulletin Board System, 240 usable responses
Joinson (2008)	CHI 2008 Pro-ceedings	Frequency of visits, Time spent	Control (Sex, Age, Occupation), Social connection, Shared identities, Photographs, Content gratification, Social investigation, Social network surfing, Status updates		Facebook, 241 usable responses
Kang et al. (2007)	Computers in Human Behavior	Social participation	Community commitment, Loyalty		Online community for healthy living and herbal products, 632 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Kankanhalli, Tan, and Wei (2005)	MIS Quarterly	Electronic knowledge repository usage by knowledge contributors	Loss of power, Codification effort, Organizational reward, Image, Reciprocity, Knowledge self-efficacy, Enjoyment in helping others, Generalized trust, Pro-sharing norms, Identification (*for most factors, moderated effects were in focus)* Control Variables (Age, Gender, Education, Work experience, Community size)		Electronic knowledge repositories, 150 usable responses from knowledge contributors
Kim, Choi, and Han (2004)	Int. J. of Internet Mktg. and Advertising	Social participation	Community commitment		Online community for healthy living and herbal products, 1514 usable responses
Kim, Lee, Hiemstra (2004)	Tourism Mgmt.	Loyalty (as relative visit frequency)	Sense of online virtual community: - Membership - Influence and relatedness - Integration and fulfillment of need - Shared emotional connection	# of travel product purchases	Travel related online virtual community, 351 usable responses
Koh and Kim (2004)	Expert Systems with Appli-cations	Knowledge sharing activity (objective data) Community participation (perceptual data)	-	**of Knowledge sharing activity:** Community participation (as positive perception), Community promotion **of Community participation:** Loyalty to the virtual community provider,	77 virtual communities of an Internet community service provider, 641 usable responses from the 77 communities
Koh et al. (2007)	Communi-cations of the ACM	Posting activity, Viewing activity	Leaders involvement, Level of Offline interaction, Usefulness (of content)		77 virtual communities of an Internet community service provider with at least 3 answers per community

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Lampe et al. (2010)	CHI	Perceived frequency of use: - Past use, - Future use, - Future contribution) Contributions: - # messages, - # write-ups	Uses and Gratifications (information seeking, providing information, social enhancement, maintaining interpersonal connectivity, entertainment, self discovery), Social identity (cognitive, evaluative), Organization commitment (sense of belonging, normative commitment), Satisfaction, Control Variables (Education, Age, Community self-efficacy, Internet self-efficacy)		User generated encyclopedia and writing platform, usable responses: 295 from anonymous users, 304 from registered users, 165 from contributors matched with IP address
Lin (2007)	Internet Research	Behavior intention (to return to use the virtual community)	Sense of belonging		20 virtual communities, 165 usable responses
Lin (2006)	Cyber-Psychology & Behavior	Behavioral intention to participate	Attitude (tow. participation in virtual community), Subjective norms, Perceived behavioral control		Different virtual communities, 165 usable responses
Ling et al. (2006)	Journal of Computer-Mediated Communi-cation	Contribution (posts, ratings, log-ins)	Uniqueness, Dissimilarity, Past log-ins, Fewer weeks elapsed since last login, Benefit, Intrinsic motivation, Goal assignment (group, individual), Goal specificity	-	Online movie recommender community, four experiments with 1) 245, 2) 830, 3) 806, 4) 833 participants
Ma and Agarwal (2007)	Information Systems Research	A3: Knowledge Contribution	Satisfaction, Perceived identity verification, Group identification, Information need fulfillment, Tenure, Offline activities		Two online communities (emotional support / car community), 666 usable responses
Nambisan and Baron (2007)	Journal of Interactive Mktg.	Participation (number of postings)	Benefit constructs: - Learning - Social integrative - Personal integrative, - Hedonic Attitude towards the host firm, Community norms, Tenure		Online forums for customers to provide product support services to peers, 152 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Nov and Ye (2008)	ICIS '08	Photo sharing (# of user's public photos)	Enjoyment, Commitment to the community, Self development, Structural embeddedness (# of ties to other actors), Tenure in community		FlickR photo community, 422 usable responses
Pajuniemi (2009)	Future of Consumer Society Con-ference	Active participation, Passive participation	-	**of Active participation:** Community Bonding **of Passive participation:** Community Loyalty	Four different virtual brand communities, 297 usable responses
Pelling and White (2009)	Cyber-Psychology & Behavior	Intention to engage in high-level social networking Web site use High-level social networking Web site use behavior	Age, Gender, Past behavior, Attitude, Subjective norm, Perceived behavioral control, Self identity, Belongliness	**of Intentions to engage:** Behavior	Different social networking Web sites, 233 usable responses, 129 usable responses in follow-up survey
Ridings, Gefen, and Arinze (2002)	Journal of Strategic Information Systems	Desire to Give Info, Desire to Get Info	Trust in Others' Ability, Trust in Others' Benevolence/ Integrity		36 different Bulletin Boards, 663 usable responses
Shang, Chen and Liao (2006)	Internet Research	Participation (Lurking and Posting)	Involvement (in the product - cognitive, affective), Trust	Attitudinal brand loyalty	Brand community (Apple computer users), 316 usable responses
Shen and Khalifa (2008)	Int. J. of Human-Computer Interaction	Community participation (# of postings, # of different threads, # of new threads created)	Extrinsic motivation, Intrinsic motivation, Awareness, Affective social presence, Cognitive social presence		Four similar online forums, 430 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Tiwana and Bush (2005)	IEEE Trans-actions on Engineer-ing Mgmt.	Continuance intention (using knowledge network)	Satisfaction, Reputation, Relational Capital (relationship to other members), Personalization, Control variables (Tenure, Usage in hours per week)		Four expertise sharing networks, 122 usable responses
Wang and Fesenmaier (2004a)	Tourism Mgmt.	General participation, Active contribution	**of General participation:** Participation benefits (functional, social, psychological, hedonic) **of Active contribution:** General participation, Contribution incentives (instrumental, efficacy, quality control, gaining status, expectancy)	**of General participation:** Active contribution **of Active contribution:** -	Virtual travel community, 322 usable responses
Wang and Fesenmaier (2004b)	J. of Travel Research	Online community participation (time spent per week)	Functional needs, Social needs, Psychological needs, Hedonic needs, Membership duration, Age, Gender, Education		Virtual travel community, 322 usable responses
Wasko and Faraj (2005)	MIS Quarterly	Volume of contribution	Reputation, Enjoy helping, Centrality (degree), Self-rated expertise, Tenure in field, Commitment, Reciprocity		Electronic network supporting professional legal association, 173 usable responses; Postings from a four months period
Wiertz and de Ruyter (2007)	Organi-zation Studies	Knowledge contribution quantity	Reciprocity, Commitment to the community, Commitment to the host firm, Informational value (Moderator), Sportsmanship (Moderator), Online interaction propensity (Moderator)		Firm-hosted technical support community, 203 usable responses
Wilson, Fornasier, and White (2009)	Cyber-Psychology & Behavior	Time spent using social networking site (hours per week)	NEO five-factor inventory (openness to experience, conscientiousness, extroversion, agreeableness, neuroticism), Coopersmith self-esteem inventory		Different social networking sites, 201 usable responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Woisetschläger, Hartleb, and Blut (2008)	Journal of Relation-ship Mktg.	Participation	Community identification, Community satisfaction, Degree of influence	Word-of-Mouth recommendation, Brand image, Community loyalty	Brand community (naming rights of virtual football stadium), 1025 usable responses
Xie, Chen, and Wu (2008)	Int. Conf. on Info. Mgmt., Innovation Mgmt. and Industrial Engineer.	Participation intention, Contribution intention	Commitment (affective, continuance, normative)		Different communities, 213 usable responses
Xu, Jones, and Shao (2009)	Information & Mgmt.	Performance *(amount and quality of knowledge contributed to the project, i.e. # of function points made and accepted into the project during the observed time period)*	Involvement Control variables (Education, Experience)		Open source project community about internet communication,172 usable responses
Yoo, Suh, and Lee (2002)	Journal of. Global Information Mgmt.	Participation frequency Visit frequency and time	**of Participation:** Sense of community (fulfillment, membership, influence) **of Visit:** Managing strategy (purpose, rule, events, subgroups), IS Quality (system, information)	**of Visit:** Sense of community **of Participation:** -	8 different non-profit and commercial online communities, 240 used responses

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Authors	Publication	Participation	Antecedents	Consequences	Community Type
Yu, Lu, and Liu (2010)	Computers in Human Behavior	Knowledge sharing behavior	Enjoy helping, Sharing culture, Usefulness/ relevancy (of knowledge sharing)		Three online communities (participants who had knowledge sharing experience via weblogs), 442 usable responses
Zhang (2010)	IEEE Trans-actions on Engineer. Mgmt.	Usage	Sense of Community, Satisfaction		Social Networking Sites, 181 usable responses of university students
Zhang et al. (2010)	Int. J. of Information Mgmt.	Intention to continue sharing knowledge	Psychological safety, Trust, Control Variables (Membership, Knowledge self-efficacy, Satisfaction, Habit)		Two university virtual communities for knowledge exchange and discussion, 144 usable responses
Zhou (2011)	Internet Research	Participation Intention	Social Identity, Subjective Norm, Group Norm	Participation Behavior	Different online communities, 450 usable responses from university students

Note: Table includes only direct effects to and from User Participation
All variables tested in the respective research study presented (incl. Non-significant antecedents and consequences)

Appendix 1: Literature Review on Online Community Participation Studies (continued)

Appendix 2: Original Constructs adapted from Past Research

Constructs *(Authors)*		**Measures**
Satisfaction	1	Overall satisfaction (“very dissatisfied" / “very satisfied”)
Fornell et al. (1996)		The degree to which performance falls short of or exceeds expectations (“falls short of expectations” / “exceeds expectations”)
		Rating of performanee relative to the customer's ideal good or service In the category. (“not very close to ideal provider" / “very close to ideal provider”)
Identification	1	I am very attached to the community.
Algesheimer,	2	Other brand community members and I share the same objectives.
Dholakia,	3	The friendships I have with other brand community members mean a lot to me.
and Herrmann (2005)	4	I see myself as a part of the brand community.
	5[1)]	If brand community members planned something, I’d think of it as something “we” would do rather than something “they” would do.
Reciprocity *Wasko and Faraj*	1	I trust that other members of Community XYZ would contribute something to the community if I contribute something.
(2005)	2	I know that other members of Community XYZ would contribute something, so it is only fair for me to contribute something as well.
Information Value	1	There is unique value in the XYZ forums
Mathwick, Wiertz, and	2	I find the information on this XYZ forum to be valuable
de Ruyter (2008)	3	I think of this XYZ forum as an information resource
Social Value	1	To have something to do with others.
Dholakia, Bagozzi,	2	To stay in touch.
and Pearo (2004)		
WOM Intention	1	I would recommend XYZ for ... to my friends.
Maxham and	2[1)]	How likely are you to spread positive word of mouth about XYZ ?
Netemeyer (2002)	3[1)]	If my friends were looking to purchase ..., I would tell them to try XYZ.

1) Item was not used in user survey of this dissertation.

Appendix 3: Study 1 – Non-Response Bias Tests

	Early Respondents (first 25%, n=172)		Late Respondents (last 25%, n=172)		T-Tests	
Measures	Mean	Std.Dev.	Mean	Std.Dev.	T-value	P-value
T-Tests						
Overall Satisfaction	4.50	1.18	4.60	1.23	-.773	.440
Identification	3.21	1.60	3.43	1.71	-1.244	.222
WOM Intention	3.95	2.12	4.10	2.08	-.642	.521
Active Participation[1)]	1.30	1.80	1.58	1.98	-1.357	.176
Passive Participation[1)]	2.00	1.49	2.19	1.59	-1.130	.259
Frequency of Use	1.97	2.31	2.35	2.41	-1.497	.135
WOM Provision Online[1)]	.40	.70	.50	.85	-1.162	.246
WOM Provision Offline[1)]	.84	.93	.96	1.06	-1.092	.276
Social Interaction Value	3.66	1.88	3.48	1.88	.908	.364
Information Consumption Value	4.90	1.71	4.89	1.81	.041	.967
Age	24.51	9.12	24.42	9.93	.090	.928
Gender	.49	.50	.60	.49	-1.955	.051
Tenure (in months)	13.37	12.01	15.03	12.95	-1.230	.219
Internet Experience	8.44	4.23	8.14	4.05	.677	.499
Internet Usage	4.83	5.13	4.13	3.12	1.512	.131
Chi^2-Tests				Pearson-	Chi^2-Value	P-value
Gender					3.800	.051
Education[2)]					.896	.925

1) log-transformed variables; 2) Chi^2-Tests based on five education groups;
Std.Dev.= Standard Deviation

	Survey Sample (n=689)		Network Sample (n=31,638)		T-Tests	
Measures	Mean	Std.Dev.	Mean	Std.Dev.	T-value	P-value
Active Participation[1)]	1.38	1.88	.79	1.43	8.153	.000
Profile Visits	1135.01	2926.33	1059.19	2360.53	.675	.500
Contacts in Friend List	88.54	232.35	100.31	174.19	-1.739	.082
Age[2)]	23.79	9.23	21.55	5.75	6.340	.000
Gender	.54	.50	.51	.50	1.684	.093
Tenure (in years)[3)]	1.81	1.25	2.17	1.26	-7.532	.000

1) log-transformed variables;
2) n (total population) = 31.336 due to missing values;
3) n (survey) = 685, due to missing values;
Note: Network Sample includes users with last log-in in the last two month before the sending of the survey, who had a valid E-Mail address and at least one contact in the online social community; it does not include users from the survey sample

Appendix 4: Study 1 – Robustness Test of Results from Main and Step-Down Analysis

	Robustness Tests						Regression Assumption Tests		
	Kruskal-Wallis Test Results[1)]		Robust Regression Results[2)]		Bootstrapped Regression Results[3)]		Multi-collinearity[4)]	Heterosce-dasticity[5)]	Normality[6)]
			Main Analysis[7)]	Step-Down Analysis[8)]	Main Analysis	Step-Down Analysis	Variance Inflation Factor	(Breusch-Pagan-Test)	(Smirnov-Kolmogorov-Test)
Dependent Variable (DV)	Chi^2	p	p	p	p	p	VIF	p	p
Marketing Channel[9)]									
Satisfaction	1.33	0.25	0.41	0.55	0.42	0.57	1.28	0.07	0.01
Identification	20.18	0.00	0.00	0.08	0.00	0.08	1.28	0.04	0.00
Active Participation	25.53	0.00	0.00	0.01	0.00	0.01	1.28	0.00	0.00
Passive Participation	40.57	0.00	0.00	0.00	0.00	0.00	1.28	0.05	0.00
Frequency of Use	26.38	0.00	0.00	0.05	0.00	0.05	1.28	0.00	0.00
WOM Intention	3.93	0.05	0.08	0.17	0.08	0.16	1.28	0.55	0.00
WOM Provision Online	18.77	0.00	0.00	0.00	0.00	0.00	1.28	0.00	0.00
WOM Provision Offline	41.33	0.00	0.00	0.00	0.00	0.00	1.28	0.00	0.00
Social Interaction Value (SIV)	50.34	0.00	0.00	-	0.00	-	1.28	0.21	0.00
Information Cons Value (ICV)	26.28	0.00	0.00	-	0.00	-	1.28	0.39	0.00
WOM Channel[10)]									
Satisfaction	3.05	0.08	0.07	0.13	0.08	0.13	1.11	0.90	0.02
Identification	1.49	0.22	0.34	0.83	0.34	0.84	1.11	0.96	0.00
Active Participation	3.72	0.05	0.03	0.23	0.03	0.23	1.11	0.00	0.00
Passive Participation	11.63	0.00	0.00	0.01	0.00	0.01	1.11	0.09	0.00
Frequency of Use	4.56	0.03	0.01	0.09	0.01	0.09	1.11	0.00	0.00
WOM Intention	4.18	0.04	0.04	0.03	0.04	0.04	1.11	0.94	0.00
WOM Provision Online	1.91	0.17	0.40	0.25	0.40	0.26	1.11	0.20	0.00
WOM Provision Offline	15.96	0.00	0.00	0.00	0.00	0.00	1.11	0.00	0.00
Social Interaction Value (SIV)	3.58	0.72	0.05	-	0.05	-	1.11	0.43	0.00
Information Cons Value (ICV)	0.13	0.06	0.40	-	0.39	-	1.11	0.31	0.00

** p<.10; ** p<.05; *** p<.01*
1) Kruskal-Wallis equality-of-populations rank test; p-value for Marketing / WOM Channel as independent variable on DV
2) Regression results with robust standard errors
3) Regression results with bootstrapping (5,000 bootstrap samples)
4) highest VIF from regressions on DV reported, with independent variables: Marketing / WOM Channel, age, gender, tenure, SIV, ICV
5) test executed with 'hettest' in Stata on regression on DV, with independent variables: Marketing / WOM Channel, age, gender, tenure
6) test executed with 'sktest' in Stata on regression on DV, with independent variables: Marketing / WOM Channel, age, gender, tenure
7) p-value for Marketing Channel / WOM Channel on DV; other covariates included: age, gender, tenure
8) p-value for Marketing Channel / WOM Channel on DV; other covariates included: age, gender, tenure, SIV, ICV
9) Marketing Channel = Word-of-mouth (=1) vs. Personal Selling (without received WOM; =0); n=398
10) WOM Channel = Offline (=1) vs. Online WOM (=0); n=310

Appendix 5: Study 1 – SEM for Mediation of Acquisition Channels via Motivation

			Acquisition Channel					
			Marketing Channel[1)]			WOM Channel[2)]		
			Coeff.	p		Coeff.	p	
Direct Effects								
X --> M								
Acquisition Channel	→	Social Interaction	1,575	,000	***	,589	,030	**
Acquisition Channel	→	Information Consumption	-,950	,000	***	-,129	,647	
M --> Y								
Social Interaction	→	Satisfaction	,208	,000	***	,286	,000	***
Social Interaction	→	Identification	,396	,000	***	,470	,000	***
Social Interaction	→	WOM Intention	,438	,001	***	,522	,000	***
Social Interaction	→	Active Participation	,441	,000	***	,577	,000	***
Social Interaction	→	WOM Provision Online	,075	,005	***	,072	,003	***
Social Interaction	→	WOM Provision Offline	,181	,000	***	,158	,000	***
Information Consumption	→	Satisfaction	,303	,000	***	,224	,000	***
Information Consumption	→	Identification	,234	,001	***	,175	,001	***
Information Consumption	→	WOM Intention	,548	,000	***	,551	,000	***
Information Consumption	→	Active Participation	,071	,200		-,054	,314	
Information Consumption	→	WOM Provision Online	,066	,008	***	,038	,105	
Information Consumption	→	WOM Provision Offline	,127	,000	***	,116	,002	***
X --> Y								
Acquisition Channel	→	Satisfaction	,131	,372		,236	,144	
Acquisition Channel	→	Identification	,317	,113		,060	,793	
Acquisition Channel	→	WOM Intention	,268	,189		,404	,087	*
Acquisition Channel	→	Active Participation	,560	,002	***	,206	,437	
Acquisition Channel	→	WOM Provision Online	,343	,000	***	-,133	,176	
Acquisition Channel	→	WOM Provision Offline	,497	,000	***	,539	,001	***

1) WOM vs. PS (WOM=1; PS=0); 2) Offline vs. Online WOM (Offline WOM=1; Online WOM=0)
X = Independent Variable; M = Mediator; Y = Dependent Variable
Maximum Likelihood Bootstrapping with 5,000 Bootstrap Samples
Bias-corrected estimates and p-values
**** p<.01; ** p<.05; * p<.10*

Appendix 6: Study 2 – Non-Response Bias Tests

	Early Respondents (first 25%, n=121)		Late Respondents (last 25%, n=121)		T-Tests	
Measures	Mean	Standard Deviation	Mean	Standard Deviation	T-value	P-value
Active Participation[1)]	6.32	6.81	7.27	7.38	-1.041	.299
Degree[2)]	2.91	1.68	3.09	1.79	-.807	.421
Betweenness[2)]	11.11	6.29	11.96	5.77	-1.088	.278
Ego-Network Density	.18	.20	.23	.22	-1.591	.113
Share of Real-World Friends	.39	.31	.43	.33	-.979	.328
Identification	3.32	1.56	3.52	1.69	-.935	.350
Reciprocity	3.26	1.62	3.10	1.77	.720	.472
Overall Satisfaction	4.60	1.20	4.56	1.27	.277	.782
Social Interaction Value	3.88	1.82	3.57	1.95	1.284	.201
Tenure (in years)	1.86	1.25	1.94	1.32	-.469	.639
Age	23.22	8.50	23.60	9.04	-.330	.742
Gender (1=female)	.49	.50	.60	.50	-1.680	.094

1) categorized measure
2) log-transformed measure

	Survey Sample (n=486)		Network Sample (n=30,180[1)])		T-Tests	
Measures	Mean	Std Dev	Mean	Std Dev	T-value	P-value
Active Participation[3)]	6.50	7.09	3.16	5.55	-10.33	.000
Degree[4)]	3.07	1.73	3.18	1.68	1.53	.126
Betweenness[4)]	11.69	5.90	11.58	5.56	-0.43	.670
Ego-Network Density	0.22	0.23	0.21	0.21	-1.14	.256
Tenure	1.89	1.28	2.16	1.26	4.74	.000
Age[2)]	22.48	7.97	21.25	5.66	-3.40	.001
Gender (1=female)	0.54	0.50	0.52	0.50	-0.79	.433

1) Network Sample does not include users from survey sample
2) Due to missing values in network sample n=29,945
3) categorized measure
4) log-transformed measure

Appendix 7: Study 2 – Verification Test with different Weightings of Participation

	Model 1[1)]				Model 2c[2)]			
	Reweighted Participation 1[3)]		Reweighted Participation 2[4)]		Reweighted Participation 1		Reweighted Participation 2	
	Logit[5)]	**Neg.Bin.**	**Logit**	**Neg.Bin.**	**Logit**	**Neg.Bin.**	**Logit**	**Neg.Bin.**
Intercept	-2.156 ***	1.490 ***	-2.156 ***	1.512 ***	-3.010 ***	1.284 ***	-3.013 ***	1.324 ***
Degree	0.815 ***	0.322 ***	0.814 ***	0.324 ***	0.798 ***	0.158 ***	0.798 ***	0.162 ***
Betweenness	-0.027 ***	-0.028 ***	-0.027 ***	-0.029 ***	0.034	0.000	0.035	-0.002
Ego-Network Density	-0.849 ***	-0.369 ***	-0.857 ***	-0.312 ***	-1.167 **	-0.289	-1.169 **	-0.265
Share Real Friends	-	-	-	-	-1.494 ***	-0.382 ***	-1.497 ***	-0.349 ***
Identification	-	-	-	-	0.087	0.107 ***	0.088	0.108 ***
Reciprocity	-	-	-	-	0.124	-0.011	0.123	-0.007
Overall Satisfaction	-	-	-	-	0.273 **	0.057 *	0.273 **	0.054 *
Tenure	-0.554 ***	-0.192 ***	-0.554 ***	-0.185 ***	-0.480 ***	-0.069 **	-0.480 ***	-0.066 **
Age	0.047 ***	0.009 ***	0.046 ***	0.009 ***	0.048 **	0.010 **	0.048 **	0.009 **
Gender	-0.093 ***	-0.033 **	-0.095 ***	-0.018	-0.198	-0.017	-0.198	-0.021
Alpha	0.5433 ***		0.5029 ***		0.2096 ***		0.2073 ***	
Log-Likelihood	-55486		-55854		-1199		-1208	
AIC	111002		111738		2445		2463	

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized)

**** p<.01 ** p<.05 * p<.10;*

1) n=30,431; 2) n=486

*3) Reweighted Participation 1: weighting messages*2; guestbook entries*2; gifts*2; comments*1; ratings*1*

*4) Reweighted Participation 2: weighting messages*2; guestbook entries*5; gifts*2; comments*5; ratings*1*

5) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros). Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"

Appendix 8: Study 2 – Zero-inflated Negative Binomial Model Specification

In a zero-inflated negative binomial model the value of the dependent variable is specified as:

$y_i \sim 0$ with probability φ_i

y_i ~ Negative Binomial with probability $1-\varphi_i$,

where y_i represents the expected number of contributions made by user *i*. φ_i is parameterized as a logistic function of the observable vector of covariates z_i and can be obtained through (Cameron and Trivedi 1998):

$$\varphi_i = \frac{\exp(z_i'\gamma)}{1+\exp(z_i'\gamma)} \ ,$$

where γ represents the parameter estimates for the independent variables z_i' that represents the (0/1) outcome of the binary model, whether or not activity can be observed.

The observed random variable y_i is then generated through

$$y_i = z_i y_i'$$

where z_i represents the (0/1) outcome of the binary model, and y_i' is distributed as negative binomial (μ_i, α^{-1}), given that z_i =1.

Probabilities for certain outcomes can be obtained through (Cameron and Trivedi 1998):

$$\Pr[y_i = 0] = \varphi_i + (1-\varphi_i) f(0)$$

$$\Pr[y_i = r] = (1-\varphi_i) f(r), \qquad r=1,2,\ldots,$$

where f (·) is the negative binomial probability distribution for y_i', that is:

$$f(y_i|\mu_i,\alpha) = \frac{\Gamma(y_i+\alpha^{-1})}{\Gamma(y_i+1)\Gamma(\alpha^{-1})}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}\left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i}$$

Here, the most common implementation of the negative binomial is a NB2 model (Cameron and Trivedi 1998). When $\alpha = 0$, it is the Poisson distribution. The function $\Gamma(\cdot)$ is the gamma function (for more details see for example Cameron and Trivedi 1998; Greene 2002).

The mean and variance in the ZINB model are:

$$E[y_i|x_i, z_i] = \mu_i(1-\varphi_i)$$

$$V[y_i|x_i, z_i] = \mu_i(1-\varphi_i)(1+\mu_i(\varphi_i+\alpha)),$$

which demonstrate over-dispersion: $V[y_i|x_i, z_i] > E[y_i|x_i, z_i]$.

Appendix 9: Study 2 – Verification Test with Second Observation Period for Model 1

	Model 1[1)]		Test Model[2)]		Consistent
	Logit[3)]	NB	Logit	NB	Results[4)]
Intercept	-2.137 ***	1.133 ***	-2.450 ***	1.290 ***	-
Degree	0.829 ***	0.381 ***	0.915 ***	0.404 ***	Yes/Yes
Betweenness	-0.026 ***	-0.030 ***	-0.028 ***	-0.041 ***	Yes/Yes
Ego-Network Density	-0.848 ***	-0.474 ***	-0.903 ***	-0.342 ***	Yes/Yes
Tenure	-0.568 ***	-0.237 ***	-0.580 ***	-0.203 ***	Yes/Yes
Age	0.053 ***	0.011 ***	0.050 ***	0.008 ***	Yes/Yes
Gender	-0.078 **	-0.076 ***	-0.072 **	-0.052 ***	Yes/Yes
Alpha	0.8663 ***		0.6756 ***		
Log-Likelihood	-53515		-61017		
AIC	107061		122064		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities)
**** p<.01 ** p<.05 * p<.10; NB=negative binomial;*
1) model used in main analyses: n=30,431
2) model for verification test includes all measures one period before model 1, i.e. participation is measured in T1,
the network variables at the end of T0; n=35,101
3) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
4) results are considered consistent, when the tested model shows the same results with respect to significance of the effects

Appendix 10: Study 2 – Different Categorizations of Participation for Verification Tests

 # Categories and Intervals of Transformation of Participation

Values within Intervals

40	30	25	20	15
0	0	0	0	0
1	1	1	1	1-2
2	2	2	2	
3	3	3	3	3-4
4	4	4	4-5	
5	5	5		5-7
6	6	6	6-7	
7	7	7		
8	8	8	8-9	8-10
9	9	9		
10	10	10-14	10-12	
11	11			11-14
12	12			
13	13		13-15	
14	14			
15	15-19	15-19		15-24
16			16-19	
17				
18				
19				
20-24	20-24	20-24	20-29	
25-29	25-29	25-29		25-39
30-39	30-39	30-39	30-39	
40-49	40-49	40-59	40-59	40-69
50-59	50-59			
60-79	60-79	60-79	60-79	
				70-99
80-99	80-99	80-99	80-99	
100-149	100-199	100-199	100-199	100-199
150-199				
200-249	200-299	200-299	200-299	200-299
250-299				
300-399	300-499	300-499	300-499	300-499
400-499				
500-749	500-999	500-999	500-999	500-999
750-999				
1000-1999	1000-1999	1000-1999	1000-9999	1000-9999
2000-2999	2000-3999	2000-3999		
3000-3999				
4000-4999	4000-9999	4000-9999		
5000-9999				

Appendix 11: Study 2 – Verification Test with different Transformations of Participation

	# of Categories for Transformation of the Dependent Variable												Consistent Results[3]
	25[1]		15		20		30		40		Original Scale		
	Logit[2]	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	
Intercept	-2.137 ***	1.133 ***	-2.081 ***	0.406 ***	-2.096 ***	0.963 ***	-2.180 ***	1.216 ***	-2.208 ***	1.255 ***	-2.117 ***	1.252 ***	-
Degree	0.829 ***	0.381 ***	0.799 ***	0.417 ***	0.808 ***	0.374 ***	0.858 ***	0.416 ***	0.879 ***	0.451 ***	1.018 ***	0.915 ***	Yes/Yes
Betweenness	-0.026 ***	-0.030 ***	-0.021 ***	-0.032 ***	-0.025 ***	-0.029 ***	-0.027 ***	-0.032 ***	-0.027 ***	-0.035 ***	-0.012	-0.093 ***	Yes/Yes
Ego-Network Density	-0.848 ***	-0.474 ***	-0.781 ***	-0.513 ***	-0.826 ***	-0.446 ***	-0.877 ***	-0.523 ***	-0.896 ***	-0.570 ***	-0.936 ***	-1.338 ***	Yes/Yes
Tenure	-0.568 ***	-0.237 ***	-0.552 ***	-0.266 ***	-0.551 ***	-0.235 ***	-0.592 ***	-0.257 ***	-0.609 ***	-0.279 ***	-0.691 ***	-0.618 ***	Yes/Yes
Age	0.053 ***	0.011 ***	0.061 ***	0.010 ***	0.050 ***	0.011 ***	0.059 ***	0.011 ***	0.064 ***	0.012 ***	0.082 ***	0.054 ***	Yes/Yes
Gender	-0.078 **	-0.076 ***	-0.051	-0.086 ***	-0.078 **	-0.071 ***	-0.074 **	-0.088 ***	-0.069 *	-0.102 ***	0.043	-0.369 ***	Yes/Yes
Alpha	0.866 ***		0.653 ***		0.646 ***		1.227 ***		1.519 ***		4.927 ***		
Log-Likelihood	-53515		-45004		-50777		-55931		-57466		-67195		
AIC	107061		90039		101583		111893		114963		134419		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized); n=30,431
***** p<.01 ** p<.05 * p<.10; NB=negative binomial**
1) model with 25 categories used in main analyses
2) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
3) results are considered consistent, when a minimum of 50% of the tested models show the same results with respect to significance of the effects

Appendix 11: Study 2 – Verification Test with different Transformations of Participation – Model 1

	# of Categories for Transformation of the Dependent Variable												
	25[1)]		15		20		30		40		118		Consistent
	Logit[2)]	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Results[3)]
Intercept	-2.970 ***	0.991 ***	-2.833 ***	0.297	-2.950 ***	0.805 ***	-2.961 ***	0.993 ***	-2.953 ***	0.989 ***	-2.922 ***	0.917 **	-
Degree	0.800 ***	0.185 ***	0.791 ***	0.197 ***	0.796 ***	0.180 ***	0.807 ***	0.206 ***	0.813 ***	0.234 ***	0.851 ***	0.348 ***	Yes/Yes
Betweenness	0.035	0.000	0.035	0.001	0.034	0.001	0.034	0.002	0.035	0.000	0.034	0.006	Yes/Yes
Ego-Network Density	-1.163 **	-0.354	-1.162 **	-0.292	-1.158 **	-0.317	-1.187 **	-0.331	-1.207 **	-0.343	-1.347 **	-0.360	Yes/Yes
Share Real Friends	-1.490 ***	-0.414 ***	-1.449 ***	-0.462 ***	-1.481 ***	-0.422 ***	-1.498 ***	-0.455 ***	-1.505 ***	-0.500 ***	-1.513 ***	-0.901 ***	Yes/Yes
Identification	0.079	0.116 ***	0.053	0.118 ***	0.077	0.116 ***	0.076	0.129 ***	0.073	0.144 ***	0.041	0.237 ***	Yes/Yes
Reciprocity	0.124	-0.007	0.124	-0.003	0.124	-0.007	0.125	-0.006	0.125	-0.008	0.138	-0.032	Yes/Yes
Overall Satisfaction	0.275 **	0.067 *	0.279 **	0.069 **	0.274 **	0.068 **	0.279 **	0.075 *	0.284 **	0.078 *	0.322 **	0.089	Yes/Yes
Tenure	-0.482 ***	-0.083 **	-0.480 ***	-0.096 ***	-0.480 ***	-0.085 ***	-0.487 ***	-0.092 **	-0.490 ***	-0.109 **	-0.510 ***	-0.154 **	Yes/Yes
Age	0.049 **	0.011 **	0.050 **	0.011 ***	0.049 **	0.010 ***	0.050 **	0.012 **	0.050 **	0.014 **	0.054 *	0.019 **	Yes/Yes
Gender	-0.198	-0.061	-0.198	-0.061	-0.195	-0.064	-0.207	-0.057	-0.216	-0.055	-0.228	-0.240 *	Yes/Yes
Alpha	0.282 ***		0.140 ***		0.200 ***		0.434 ***		0.561 ***		1.214 *		
Log-Likelihood	-1177		-963		-1098		-1259		-1314		-1498		
AIC	2400		1971		2242		2564		2675		3042		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized); n=486
**** p<.01 ** p<.05 * p<.10; NB=negative binomial*
1) model with 25 categories used in main analyses
2) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
3) results are considered consistent, when a minimum of 50% of the tested models show the same results with respect to significance of the effects

Appendix 11: Study 2 – Verification Test with different Transformations of Participation – Model 2c

	# of Categories for Transformation of the Dependent Variable												
	25[1]		15		20		30		40		118		Consistent
	Logit[2]	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Results[3]
Intercept	-2.508 **	-0.276	-1.671 **	-1.326 **	-2.521 ***	-0.328	-2.290 *	-0.569	-2.132	-0.709	-1.379	-1.399	-
Degree	1.073 ***	0.108	1.277 ***	0.095	1.068 ***	0.102	1.091 ***	0.130	1.106 ***	0.146	1.347 ***	0.120	Yes/Yes
Betweenness	0.012	0.042	-0.065	0.067 **	0.011	0.038	0.007	0.049 *	0.003	0.054	-0.079	0.120 ***	Yes/No
Ego-Network Density	-2.178 ***	-0.347	-2.880 **	-0.203	-2.184 ***	-0.269	-2.260 ***	-0.303	-2.322 ***	-0.295	-2.852 **	-0.711	Yes/Yes
Share Real Friends	-1.044 *	-0.377	-0.762	-0.491	-1.033 *	-0.360	-1.065	-0.397	-1.083	-0.415	-0.838	-0.787	No/Yes
Identification	-0.072	0.139 *	-0.269	0.162 **	-0.069	0.128 *	-0.091	0.146 *	-0.109	0.158 *	-0.322	0.271 **	Yes/Yes
Reciprocity	0.325 **	0.035	0.433 *	0.044	0.325 **	0.035	0.325 *	0.047	0.327 *	0.052	0.479 *	0.019	Yes/Yes
Overall Satisfaction	0.373 **	0.082	0.405	0.097	0.363 *	0.082	0.392 **	0.097	0.409 **	0.099	0.426	0.190 *	Yes/Yes
Tenure	-0.620 ***	0.002	-0.725 ***	-0.009	-0.619 ***	0.004	-0.632 ***	0.004	-0.641 ***	0.002	-0.766 ***	0.032	Yes/Yes
Age	0.026	0.022 *	0.059	0.017	0.027	0.019	0.023	0.027 *	0.021	0.030 *	0.061	0.019	Yes/No
Gender	-0.916 **	0.205	-1.552 **	0.262	-0.902 **	0.180	-0.986 **	0.242	-1.050 **	0.276	-1.660 **	0.301	Yes/Yes
Alpha	0.579 ***		0.393 ***		0.409 ***		0.815		0.997		1.900 ***		
Log-Likelihood	-503		-408		-471		-527		-542		-589		
AIC	1052		862		988		1101		1131		1223		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized); n=260
***** p<.01 ** p<.05 * p<.10; NB=negative binomial*
1) model with 25 categories used in main analyses
2) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
3) results are considered consistent, when a minimum of 50% of the tested models show the same results with respect to significance of the effects

Appendix 11: Study 2 – Verification Test with different Transformations of Participation – Model 3 (Non-Networker)

	# of Categories for Transformation of the Dependent Variable												
	25[1]		15		20		30		40		118		Consistent
	Logit[2]	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Results[3]
Intercept	-2.879 *	2.097 ***	-2.921 *	1.421 ***	-2.883 *	1.856 ***	-2.882 *	2.284 ***	-2.886 *	2.370 ***	-2.968 *	2.446 ***	-
Degree	0.716 ***	0.183 ***	0.701 ***	0.201 ***	0.715 ***	0.179 ***	0.717 ***	0.200 ***	0.716 ***	0.230 ***	0.721 ***	0.417 ***	Yes/Yes
Betweenness	0.018	-0.016	0.020	-0.018	0.018	-0.014	0.018	-0.017	0.018	-0.020	0.021	-0.041	Yes/Yes
Ego-Network Density	0.985	-0.540 **	1.150	-0.499 **	1.003	-0.490 **	0.991	-0.591 **	1.007	-0.655 **	1.168	-0.661	Yes/Yes
Share Real Friends	-2.686 ***	-0.394 ***	-2.676 ***	-0.469 ***	-2.684 ***	-0.422 ***	-2.687 ***	-0.424 **	-2.689 ***	-0.468 **	-2.723 ***	-0.992 ***	Yes/Yes
Identification	0.182	0.054 *	0.180	0.054 **	0.181	0.058 **	0.181	0.062 *	0.180	0.072 *	0.158	0.177 ***	Yes/Yes
Reciprocity	-0.103	-0.048 **	-0.106	-0.047 **	-0.103	-0.048 **	-0.104	-0.055 *	-0.104	-0.066 **	-0.113	-0.140 ***	Yes/Yes
Overall Satisfaction	0.262	0.056 *	0.264	0.056 *	0.263	0.054 *	0.263	0.061	0.263	0.067	0.271	0.094	Yes/No
Tenure	-0.457 ***	-0.138 ***	-0.444 **	-0.150 ***	-0.455 **	-0.140 ***	-0.457 **	-0.157 ***	-0.456 **	-0.182 ***	-0.462 **	-0.288 ***	Yes/Yes
Age	0.074 **	0.006 *	0.075 **	0.007 **	0.074 **	0.006 *	0.075 **	0.006	0.075 **	0.008	0.084 *	0.017 **	Yes/Yes
Gender	0.407	-0.169 **	0.449	-0.156 **	0.412	-0.164 **	0.409	-0.175 **	0.413	-0.179 *	0.458	-0.274 *	Yes/Yes
Alpha	0.102 ***		0.019 ***		0.069 ***		0.189 ***		0.269 ***		0.729 **		
Log-Likelihood	-632		-515		-589		-690		-731		-875		
AIC	1310		1077		1224		1426		1509		1796		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized); n=226
**** p<.01 ** p<.05 * p<.10; NB=negative binomial*
1) model with 25 categories used in main analyses
2) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
3) results are considered consistent, when a minimum of 50% of the tested models show the same results with respect to significance of the effects

Appendix 11: Study 2 – Verification Test with different Transformations of Participation – Model 3 (Networker)

Appendix 12: Study 2 – Verification Test with Non-transformed Betweenness Measure

	Model 1[1)]				Model 2a[2)]				Model 2c[2)]				
	Used Model[3)]		Test Model[4)]		Used Model		Test Model		Used Model		Test Model		Consistent
	Logit[5)]	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Logit	NB	Results[6)]
Intercept	-2.137 ***	1.133 ***	-2.151 ***	1.026 ***	-1.795 ***	1.468 ***	-1.753 ***	1.407 ***	-2.970 ***	0.991 ***	-2.773 ***	0.997 ***	-
Degree	0.829 ***	0.381 ***	0.712 ***	0.313 ***	0.901 ***	0.246 ***	0.775 ***	0.234 ***	0.800 ***	0.185 ***	0.759 ***	0.186 ***	Yes/Yes
Betweenness[7)]	-0.026 ***	-0.030 ***	0.000 ***	-0.000 *	0.012	-0.008	0.000 ***	-0.000	0.035	0.000	0.000 ***	-0.000	No/Yes
Ego-Network Density	-0.848 ***	-0.474 ***	-0.684 ***	-0.379 ***	-1.676 ***	-0.394	-1.301 **	-0.418 *	-1.163 **	-0.355	-0.845	-0.358	Yes/Yes
Share Real Friends	-	-	-	-	-	-	-	-	-1.490 ***	-0.414 ***	-1.489 ***	-0.413 ***	Yes/Yes
Identification	-	-	-	-	-	-	-	-	0.080	0.117 ***	0.075	0.116 ***	Yes/Yes
Reciprocity	-	-	-	-	-	-	-	-	0.124	-0.007	0.105	-0.007	Yes/Yes
Overall Satisfaction	-	-	-	-	-	-	-	-	0.275 **	0.067 *	0.261 **	0.067 *	Yes/Yes
Tenure	-0.568 ***	-0.237 ***	-0.562 ***	-0.238 ***	-0.573 ***	-0.120 ***	-0.522 ***	-0.120 ***	-0.482 ***	-0.083 **	-0.441 ***	-0.083 **	Yes/Yes
Age	0.053 ***	0.011 ***	0.053 ***	0.008 ***	0.064 ***	0.015 ***	0.064 ***	0.016 ***	0.049 **	0.011 **	0.050 **	0.011 **	Yes/Yes
Gender	-0.078 **	-0.076 ***	-0.102 ***	-0.085 ***	-0.290	-0.126	-0.308	-0.122	-0.198	-0.061	-0.204	-0.060	Yes/Yes
Alpha	0.8663 ***		0.9023 ***		0.3365 ***		0.3343 ***		0.2820 ***		0.2802 ***		
Log-Likelihood	-53515		-53546		-1207		-1201		-1177		-1172		
AIC	107061		107120		2445		2429		2400		2387		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized)
***** p<.01 ** p<.05 * p<.10; NB=negative binomial*
1) n=30,431; 2) n=486
3) model used in main analyses
4) model for verification test includes the not log-transformed betweenness variable to reduce the correlation with degree
5) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros).
Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
6) results are considered consistent, when a minimum of 50% of the tested models show the same results with respect to significance of the effects
7) Betweenness measure not log-transformed to represent the original scale range; multiplied with the inverse ratio of the lowest non-zero value to determine the lowest betweenness value at value 1

Appendix 12: Study 2 – Verification Test with Non-transformed Betweenness Measure

Appendix 13: Study 2 – Verification Test with Control Variable for Friend Activities

	Model 2c[1)]		Model 2d[2)]		Consistent
	Logit[3)]	**NB**	**Logit**	**NB**	**Results[4)]**
Intercept	-2.970 ***	0.991 ***	-2.845 ***	0.976 ***	-
Degree	0.800 ***	0.185 ***	0.284	0.101 **	No/Yes
Betweenness	0.035	0.000	0.009	-0.005	Yes/Yes
Ego-Network Density	-1.163 **	-0.355	-1.234 **	-0.469 **	Yes/No
Share Real Friends	-1.490 ***	-0.414 ***	-1.171 ***	-0.368 **	Yes/Yes
Friends' Activity[5)]	-	-	0.324 ***	0.066 ***	-
Identification	0.080	0.117 ***	0.062	0.115 ***	Yes/Yes
Reciprocity	0.124	-0.007	0.118	-0.011	Yes/Yes
Overall Satisfaction	0.275 **	0.067 *	0.215 *	0.058 *	Yes/Yes
Tenure	-0.482 ***	-0.083 **	-0.365 ***	-0.063 *	Yes/Yes
Age	0.049 **	0.011 **	0.040 *	0.008 *	Yes/Yes
Gender	-0.198	-0.061	-0.255	-0.091	Yes/Yes
Alpha	0.2820 ***		0.2716 ***		
Log-Likelihood	-1177		-1161		
AIC	2400		2372		

Model: Zero-Inflated Negative Binomial Models; Dependent Variable: Active Participation (# of activities; categorized)
**** p<.01 ** p<.05 * p<.10; NB=negative binomial*
1) model used in main analyses: n=486
2) model for verification test includes an additional measure of the activity of the actor's contacts; n=486
3) signs in the logit-part of the model are reversed to indicate the effect on "contributions" instead of "not contributing" (certain zeros). Positive effects are thus interpreted as "the greater the independent variable, the more likely the users were to contribute"
4) results are considered consistent, when the tested model shows the same results with respect to significance of the effects
5) Friends' Activity includes all active participation of the focal users' contacts in T2. The aggregated variable is log-transformed.

Appendix 14: Study 3 – Correlations of Independent Variables

	Covariates	1	2	3	4	5	6	7	8	9	10	11	12	13	**VIF[2)]**
1	Degree (Active Contacts)	1.00													1.92
2	Betweenness (Active Contacts)[1)]	0.40 **	1.00												2.75
3	Ego-Network Density (Active Contacts)	0.08 **	0.35 **	1.00											1.57
4	Similarity (Active Contacts) - Region	0.24 **	0.67 **	0.56 **	1.00										3.68
5	Similarity (Active Contacts) - Distance	0.12 **	0.32 **	0.11 **	-0.04 **	1.00									1.61
6	Similarity (Active Contacts) - Gender	0.18 **	0.56 **	0.51 **	0.71 **	0.21 **	1.00								2.34
7	Avg. Degree of Active Contacts	0.29 **	0.57 **	0.28 **	0.42 **	0.42 **	0.41 **	1.00							1.72
8	Share of Inactive Contacts	0.06 **	0.17 **	0.08 **	0.14 **	0.11 **	0.13 **	0.13 **	1.00						1.04
9	Active Participation Level	0.58 **	0.22 **	0.04 **	0.13 **	0.07 **	0.10 **	0.15 **	0.01 **	1.00					2.17
10	Active Participation Delta	0.04 **	0.01	0.00	0.01	0.00	0.01	0.00	-0.01 **	0.47 **	1.00				1.43
11	Verified Membership	0.13 **	0.25 **	0.12 **	0.19 **	0.13 **	0.14 **	0.18 **	0.09 **	0.06 **	0.00	1.00			1.11
12	Age	-0.04 **	-0.20 **	-0.11 **	-0.19 **	0.01	-0.17 **	-0.14 **	-0.03 **	0.01	0.01	0.06 **	1.00		1.09
13	Gender	0.00	0.00	-0.03 **	-0.03 **	0.00	0.05 **	0.02 **	-0.02 **	0.01 **	0.00	-0.15 **	-0.15 **	1.00	1.06

n=55,488 observations

*** p<.01; * p<.05*

All statistics based on the data set with cutoff threshold = 3 months

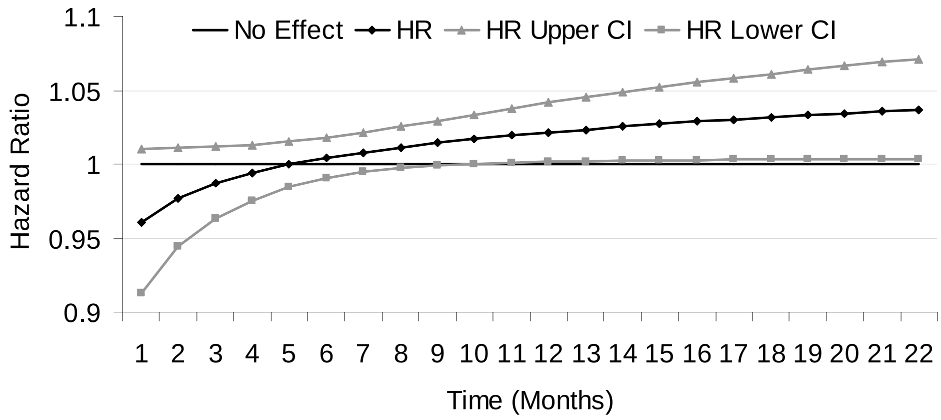
1) Transformed betweenness measure

2) VIF (Variance Inflation Factor) for independent variables based on regression on Defection Status

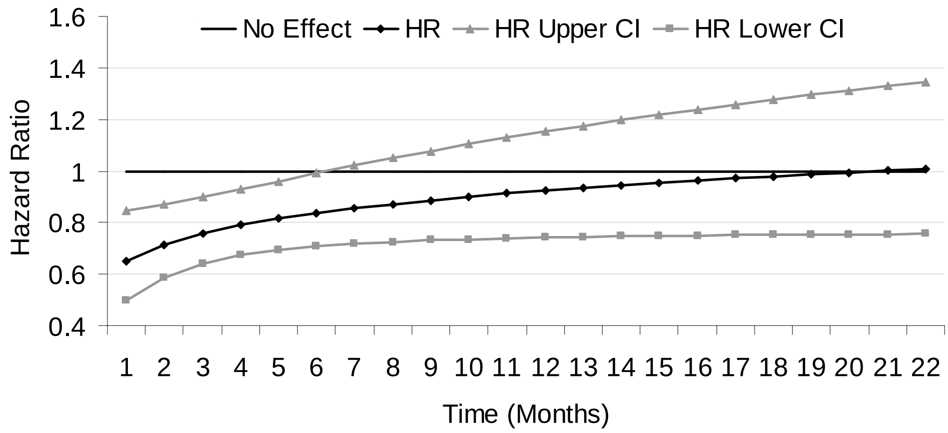
Appendix 14: Study 3 – Correlations of Independent Variables

Appendix 15: Study 3 – Time-Dependent Effects of Model 2

Time-dependent effect for Active Participation Level (Model 2)



Time-dependent effect for Verified Membership (Model 2)



HR=Hazard Ratio; CI=Confidence Interval (95%)

Appendix 16: Study 3 – Verification Test with Alternative Cutoff Thresholds

Model 1					Main Model												
Defection Identification Threshold:		2 Months			3 Months			4 Months			5 Months			6 Months			
Covariates		**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	
Main Effects of Covariates																	
Degree (Active Contacts)		-0.116	0.009	0.000 ***	-0.133	0.010	0.000 ***	-0.154	0.011	0.000 ***	-0.164	0.012	0.000 ***	-0.183	0.013	0.000 ***	
Active Participation Level		-0.035	0.012	0.003 ***	-0.034	0.012	0.004 ***	-0.033	0.012	0.008 ***	-0.033	0.013	0.010 ***	-0.032	0.013	0.016 **	
Active Participation Delta		-0.004	0.001	0.016 **	-0.004	0.001	0.002 ***	-0.004	0.001	0.002 ***	-0.004	0.002	0.014 **	-0.004	0.002	0.017 **	
Verified Membership		-0.828	0.124	0.000 ***	-0.814	0.125	0.000 ***	-0.809	0.126	0.000 ***	-0.801	0.127	0.000 ***	-0.774	0.127	0.000 ***	
Age		0.009	0.002	0.000 ***	0.009	0.003	0.001 ***	0.008	0.003	0.002 ***	0.007	0.003	0.004 ***	0.007	0.003	0.009 ***	
Gender		-0.167	0.047	0.000 ***	-0.175	0.048	0.000 ***	-0.171	0.048	0.000 ***	-0.170	0.048	0.000 ***	-0.161	0.048	0.001 ***	
Time-dependent Effects of Covariates [1)]																	
Degree (Active Contacts)	x g(t) [2)]	0.037	0.003	0.000 ***	0.043	0.003	0.000 ***	0.051	0.004	0.000 ***	0.055	0.004	0.000 ***	0.062	0.004	0.000 ***	
Active Participation Level	x g(t)	0.012	0.004	0.004 ***	0.011	0.004	0.006 ***	0.011	0.004	0.011 **	0.011	0.005	0.015 **	0.011	0.005	0.026 **	
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Verified Membership	x g(t)	0.158	0.056	0.005 ***	0.155	0.058	0.008 ***	0.160	0.060	0.008 ***	0.158	0.062	0.011 **	0.135	0.064	0.033 **	
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Gender	x g(t)	0.085	0.026	0.001 ***	0.096	0.026	0.000 ***	0.089	0.027	0.001 ***	0.085	0.028	0.002 ***	0.071	0.028	0.013 **	
Subjects			5752			5752			5752			5752			5752		
Failures			4923			4775			4658			4579			4483		
Observations			56465			55488			54394			53221			51952		
Log-Likelihood			-38796			-37784			-36963			-36406			-35717		
AIC			77611			75588			73946			72832			71453		
BIC			77701			75677			74035			72920			71542		

**** p<.01, ** p<.05, * p<.10*
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)
2) All models estimated for g(t) = ln(t)

Appendix 16a: Study 3 – Verification Test with Alternative Cutoff Thresholds for Model 1

Model 2					Main Model											
Defection Identification Threshold:		2 Months			3 Months			4 Months			5 Months			6 Months		
Covariates		Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p
Main Effects of Covariates																
Active Participation Level		-0.041	0.026	0.117	-0.040	0.026	0.121	-0.039	0.026	0.125	-0.038	0.026	0.134	-0.038	0.026	0.138
Active Participation Delta		-0.012	0.006	0.040 **	-0.012	0.006	0.040 **	-0.012	0.006	0.035 **	-0.012	0.006	0.034 **	-0.012	0.006	0.032 **
Verified Membership		-0.436	0.136	0.001 ***	-0.434	0.136	0.001 ***	-0.424	0.137	0.002 ***	-0.441	0.138	0.001 ***	-0.263	0.084	0.002 ***
Age		0.002	0.003	0.412	0.002	0.003	0.538	0.008	0.004	0.065 *	0.008	0.004	0.059 *	0.008	0.004	0.062 *
Gender		-0.152	0.049	0.002 ***	-0.157	0.049	0.002 ***	-0.092	0.033	0.005 ***	-0.098	0.033	0.003 ***	-0.104	0.033	0.002 ***
Time-dependent Effects of Covariates [1)]																
Active Participation Level	x g(t)	0.025	0.013	0.055 *	0.025	0.013	0.058 *	0.025	0.013	0.059 *	0.024	0.013	0.065 *	0.024	0.013	0.071 *
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	0.144	0.073	0.047 **	0.143	0.075	0.057 *	0.127	0.077	0.099 *	0.150	0.078	0.055 *	-	-	-
Age	x g(t)	-	-	-	-	-	-	-0.005	0.003	0.045 **	-0.006	0.003	0.032 **	-0.006	0.003	0.026 **
Gender	x g(t)	0.059	0.029	0.045 **	0.065	0.030	0.031 **	-	-	-	-	-	-	-	-	-
Subjects			4264			4264			4264			4264			4264	
Failures			3978			3901			3835			3798			3755	
Observations			31484			31121			30692			30226			29717	
Log-Likelihood			-30028			-29583			-29187			-28961			-28695	
AIC			60072			59182			58391			57938			57405	
BIC			60139			59248			58457			58004			57463	

**** p<.01, ** p<.05, * p<.10*
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)
2) All models estimated for g(t) = ln(t)

Appendix 16b: Study 3 – Verification Test with Alternative Cutoff Thresholds for Model 2

Model 3b					Main Model												
Defection Identification Threshold:		2 Months			3 Months			4 Months			5 Months			6 Months			
Covariates		Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	
Main Effects of Covariates																	
Degree (Active Contacts)		-0.026	0.007	0.000 ***	-0.035	0.008	0.000 ***	-0.044	0.009	0.000 ***	-0.043	0.010	0.000 ***	-0.051	0.011	0.000 ***	
Avg. Degree of Active Contacts		0.002	0.000	0.000 ***	0.002	0.000	0.000 ***	0.002	0.001	0.000 ***	0.002	0.001	0.000 ***	0.002	0.000	0.000 ***	
Betweenness (Active Contacts)		-0.082	0.021	0.000 ***	-0.076	0.021	0.000 ***	-0.073	0.022	0.001 ***	-0.075	0.022	0.001 ***	-0.059	0.022	0.007 ***	
Ego-Network Density (Active Contacts)		-0.893	0.347	0.010 **	-0.788	0.346	0.023 **	-0.236	0.126	0.061 *	-0.253	0.129	0.051 *	-0.314	0.134	0.019 **	
Similarity (Active Contacts) - Region		1.380	0.251	0.000 ***	1.320	0.255	0.000 ***	1.298	0.255	0.000 ***	1.332	0.258	0.000 ***	1.172	0.232	0.000 ***	
Similarity (Active Contacts) - Distance		0.011	0.002	0.000 ***	0.011	0.002	0.000 ***	0.011	0.002	0.000 ***	0.011	0.002	0.000 ***	0.007	0.001	0.000 ***	
Similarity (Active Contacts) - Gender		0.302	0.098	0.002 ***	0.315	0.101	0.002 ***	0.355	0.103	0.001 ***	0.386	0.105	0.000 ***	0.412	0.107	0.000 ***	
Share of Inactive Contacts		0.676	0.145	0.000 ***	0.723	0.153	0.000 ***	0.769	0.158	0.000 ***	0.820	0.160	0.000 ***	0.863	0.165	0.000 ***	
Active Participation Level		-0.004	0.002	0.090 *	-0.004	0.002	0.097 *	-0.003	0.002	0.118	-0.004	0.002	0.101	-0.004	0.002	0.086 *	
Active Participation Delta		-0.003	0.001	0.032 **	-0.003	0.001	0.020 **	-0.003	0.001	0.019 **	-0.003	0.001	0.029 **	-0.003	0.001	0.029 **	
Verified Membership		-0.434	0.096	0.000 ***	-0.434	0.100	0.000 ***	-0.409	0.103	0.000 ***	-0.427	0.107	0.000 ***	-0.433	0.111	0.000 ***	
Age		0.008	0.006	0.186	0.008	0.006	0.199	0.007	0.006	0.234	0.006	0.006	0.374	0.003	0.007	0.602	
Gender		-0.421	0.195	0.031 **	-0.464	0.198	0.019 *	-0.511	0.200	0.011 **	-0.544	0.202	0.007 ***	-0.475	0.203	0.019 **	
Time-dependent Effects of Covariates 1)																	
Degree (Active Contacts)	x g(t) 2)	0.008	0.002	0.001 ***	0.011	0.003	0.000 ***	0.015	0.003	0.000 ***	0.014	0.003	0.000 ***	0.018	0.004	0.000 ***	
Avg. Degree of Active Contacts	x g(t)	-0.001	0.000	0.001 ***	-0.001	0.000	0.001 ***	-0.001	0.000	0.002 ***	-0.001	0.000	0.002 ***	-0.001	0.000	0.001 ***	
Betweenness (Active Contacts)	x g(t)	0.018	0.008	0.025 **	0.016	0.008	0.055 *	0.016	0.008	0.053 *	0.018	0.009	0.046 **	0.010	0.009	0.261	
Ego-Network Density (Active Contacts)	x g(t)	0.326	0.141	0.021 **	0.261	0.145	0.072 *	-	-	-	-	-	-	-	-	-	
Similarity (Active Contacts) - Region	x g(t)	-0.492	0.105	0.000 ***	-0.442	0.109	0.000 ***	-0.451	0.110	0.000 ***	-0.485	0.113	0.000 ***	-0.395	0.102	0.000 ***	
Similarity (Active Contacts) - Distance	x g(t)	-0.002	0.001	0.006 ***	-0.002	0.001	0.060 *	-0.002	0.001	0.080 *	-0.002	0.001	0.079 *	-	-	-	
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Share of Inactive Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Active Participation Level	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Verified Membership	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Gender	x g(t)	0.185	0.078	0.018 **	0.214	0.082	0.009 ***	0.242	0.084	0.004 ***	0.261	0.087	0.003 ***	0.209	0.089	0.019 **	
Subjects			1488			1488			1488			1488			1488		
Failures			945			874			823			781			728		
Observations			24981			24367			23702			22995			22235		
Log-Likelihood			-6328			-5882			-5554			-5278			-4931		
AIC			12696			11805			11146			10594			9898		
BIC			12859			11967			11299			10746			10042		

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for g(t) = ln(t)

Appendix 16c: Study 3 – Verification Test with Alternative Cutoff Thresholds for Model 3b

Appendix 17: Study 3 – Verification Test with Alternative Functions of Time

Model 1		Main Model											
Function of time:		g(t)=ln(t)			g(t)=t			$g(t)=t^2$			g(t)=sgrt(t)		
Covariates		**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**
Main Effects of Covariates													
Degree (Active Contacts)		-0.133	0.010	0.000 ***	-0.062	0.005	0.000 ***	-0.038	0.003	0.000 ***	-0.107	0.008	0.000 ***
Active Participation Level		-0.034	0.012	0.004 ***	-0.026	0.008	0.001 ***	-0.020	0.006	0.001 ***	-0.037	0.012	0.002 ***
Active Participation Delta		-0.004	0.001	0.002 ***	-0.005	0.001	0.000 ***	-0.005	0.001	0.000 ***	-0.005	0.001	0.001 ***
Verified Membership		-0.814	0.125	0.000 ***	-0.765	0.104	0.000 ***	-0.709	0.083	0.000 ***	-0.902	0.152	0.000 ***
Age		0.009	0.003	0.001 ***	0.009	0.003	0.000 ***	0.010	0.002	0.000 ***	0.009	0.003	0.000 ***
Gender		-0.175	0.048	0.000 ***	-0.171	0.043	0.000 ***	-0.125	0.036	0.001 ***	-0.261	0.063	0.000 ***
Time-dependent Effects of Covariates [1)]													
Degree (Active Contacts)	x g(t)	0.043	0.003	0.000 ***	0.003	0.000	0.000 ***	0.000	0.000	0.000 ***	0.023	0.002	0.000 ***
Active Participation Level	x g(t)	0.011	0.004	0.006 ***	0.001	0.000	0.001 ***	0.000	0.000	0.001 ***	0.008	0.003	0.002 ***
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	0.155	0.058	0.008 ***	0.023	0.009	0.009 ***	0.001	0.000	0.009 ***	0.129	0.049	0.008 ***
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.096	0.026	0.000 ***	0.019	0.005	0.000 ***	0.001	0.000	0.000 ***	0.094	0.024	0.000 ***
Subjects			5752			5752			5752			5752	
Failures			4775			4775			4775			4775	
Observations			55488			55488			55488			55488	
Log-Likelihood			-37784			-37831			-37868			-37808	
AIC			75588			75682			75755			75635	
BIC			75677			75771			75844			75725	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests for main model; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for sample with three months cutoff threshold

Appendix 17a: Study 3 – Verification Test with Alternative Functions of Time for Model 1

Model 2		Main Model											
Function of time:		g(t)=ln(t)			g(t)=t			$g(t)=t^2$			g(t)=sqrt(t)		
Covariates		Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p
Main Effects of Covariates													
Active Participation Level		-0.040	0.026	0.121	-0.033	0.024	0.163	-	-	-	-0.057	0.036	0.111
Active Participation Delta		-0.012	0.006	0.040 **	-0.012	0.006	0.043 **	-	-	-	-0.012	0.006	0.040 **
Verified Membership		-0.434	0.136	0.001 ***	-0.337	0.120	0.005 ***	-	-	-	-0.462	0.178	0.009 ***
Age		0.002	0.003	0.538	0.002	0.003	0.530	-	-	-	0.002	0.003	0.533
Gender		-0.157	0.049	0.002 ***	-0.174	0.046	0.000 ***	-	-	-	-0.234	0.069	0.001 ***
Time-dependent Effects of Covariates [1)]													
Active Participation Level	x g(t)	0.025	0.013	0.058 *	0.005	0.003	0.071 *	-	-	-	0.024	0.013	0.066 *
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	0.143	0.075	0.057 *	0.016	0.013	0.229	-	-	-	0.099	0.068	0.141
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.065	0.030	0.031 **	0.017	0.006	0.003 ***	-	-	-	0.074	0.028	0.009 ***
Subjects			4264			4264						4264	
Failures			3901			3901						3901	
Observations			31121			31121			MODEL ESTIMATION IN			31121	
Log-Likelihood			-29583			-29582			STATA			-29583	
AIC			59182			59180			NOT SUCCESSFUL			59181	
BIC			59248			59247						59248	

**** p<.01, ** p<.05, * p<.10*
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests for main model; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)
2) All models estimated for sample with three months cutoff threshold

Appendix 17b: Study 3 – Verification Test with Alternative Functions of Time for Model 2

Model 3b		Main Model											
Function of time:		g(t)=ln(t)			g(t)=t			g(t)=t²			g(t)=sgrt(t)		
Covariates		Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p
Main Effects of Covariates													
Degree (Active Contacts)		-0.035	0.008	0.000 ***	-0.017	0.004	0.000 ***	-0.010	0.002	0.000 ***	-0.030	0.007	0.000 ***
Avg. Degree of Active Contacts		0.002	0.000	0.000 ***	0.002	0.000	0.000 ***	0.001	0.000	0.000 ***	0.003	0.001	0.000 ***
Betweenness (Active Contacts)		-0.076	0.021	0.000 ***	-0.070	0.014	0.000 ***	-0.061	0.011	0.000 ***	-0.086	0.022	0.000 ***
Ego-Network Density (Active Contacts)		-0.788	0.346	0.023 **	-0.722	0.260	0.006 ***	-0.562	0.197	0.004 ***	-0.991	0.393	0.012 **
Similarity (Active Contacts) - Region		1.320	0.255	0.000 ***	0.929	0.194	0.000 ***	0.649	0.151	0.000 ***	1.401	0.289	0.000 ***
Similarity (Active Contacts) - Distance		0.011	0.002	0.000 ***	0.010	0.002	0.000 ***	0.009	0.001	0.000 ***	0.011	0.002	0.000 ***
Similarity (Active Contacts) - Gender		0.315	0.101	0.002 ***	0.312	0.101	0.002 ***	0.315	0.101	0.002 ***	0.312	0.101	0.002 ***
Share of Inactive Contacts		0.723	0.153	0.000 ***	0.736	0.154	0.000 ***	0.746	0.154	0.000 ***	0.729	0.154	0.000 ***
Active Participation Level		-0.004	0.002	0.097 *	-0.003	0.002	0.103	-0.004	0.002	0.085 *	-0.003	0.002	0.106
Active Participation Delta		-0.003	0.001	0.020 **	-0.003	0.001	0.014 **	-0.003	0.001	0.013 **	-0.003	0.001	0.017 **
Verified Membership		-0.434	0.100	0.000 ***	-0.434	0.100	0.000 ***	-0.435	0.100	0.000 ***	-0.433	0.100	0.000 ***
Age		0.008	0.006	0.199	0.008	0.006	0.177	0.008	0.006	0.161	0.008	0.006	0.189
Gender		-0.464	0.198	0.019 *	-0.241	0.147	0.102	-0.122	0.112	0.275	-0.460	0.221	0.037 **
Time-dependent Effects of Covariates [1)]													
Degree (Active Contacts)	x g(t)	0.011	0.003	0.000 ***	0.001	0.000	0.000 ***	0.000	0.000	0.000 ***	0.007	0.002	0.000 ***
Avg. Degree of Active Contacts	x g(t)	-0.001	0.000	0.001 ***	-0.000	0.000	0.000 ***	-0.000	0.000	0.000 ***	-0.001	0.000	0.000 ***
Betweenness (Active Contacts)	x g(t)	0.016	0.008	0.055 *	0.002	0.001	0.015 **	0.000	0.000	0.015 **	0.014	0.006	0.023 **
Ego-Network Density (Active Contacts)	x g(t)	0.261	0.145	0.072 *	0.042	0.019	0.025 **	0.002	0.001	0.021 **	0.238	0.113	0.035 **
Similarity (Active Contacts) - Region	x g(t)	-0.442	0.109	0.000 ***	-0.049	0.015	0.001 ***	-0.002	0.001	0.012 **	-0.323	0.085	0.000 ***
Similarity (Active Contacts) - Distance	x g(t)	-0.002	0.001	0.060 *	-0.000	0.000	0.062 *	-0.000	0.000	0.069 *	-0.001	0.001	0.060 *
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Share of Inactive Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Active Participation Level	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.214	0.082	0.009 ***	0.021	0.010	0.047 **	0.001	0.000	0.122	0.143	0.063	0.022 **
Subjects			1488			1488			1488			1488	
Failures			874			874			874			874	
Observations			24367			24367			24367			24367	
Log-Likelihood			-5882			-5884			-5890			-5882	
AIC			11805			11809			11819			11805	
BIC			11967			11971			11981			11967	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests for main model; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for sample with three months cutoff threshold

Appendix 17c: Study 3 – Verification Test with Alternative Functions of Time for Model 3b

COHORT 2		Model 1				Model 2				Model 3a				Model 3b			
		All Users				Not-Connected Users				Connected Users - Base				Connected Users - Full			
Covariates		**HR**	**Coef.**	**SE**	**p**	**HR**	**Coef.**	**SE**	**p**	**HR**	**Coef.**	**SE**	**p**	**HR**	**Coef.**	**SE**	**p**
Main Effects of Covariates																	
Degree (Active Contacts)		0.894	-0.112	0.009	0.000 ***	-	-	-	-	0.972	-0.028	0.007	0.000 ***	0.979	-0.021	0.006	0.001 ***
Avg. Degree of Active Contacts		-	-	-	-	-	-	-	-	-	-	-	-	1.001	0.001	0.000	0.000 ***
Betweenness (Active Contacts)		-	-	-	-	-	-	-	-	-	-	-	-	0.962	-0.038	0.007	0.000 ***
Ego-Network Density (Active Contacts)		-	-	-	-	-	-	-	-	-	-	-	-	0.339	-1.083	0.413	0.009 ***
Similarity (Active Contacts) - Region		-	-	-	-	-	-	-	-	-	-	-	-	2.055	0.720	0.118	0.000 ***
Similarity (Active Contacts) - Distance		-	-	-	-	-	-	-	-	-	-	-	-	1.015	0.015	0.002	0.000 ***
Similarity (Active Contacts) - Gender		-	-	-	-	-	-	-	-	-	-	-	-	1.155	0.144	0.109	0.189
Share of Inactive Contacts		-	-	-	-	-	-	-	-	-	-	-	-	2.222	0.798	0.175	0.000 ***
Active Participation Level		0.996	-0.004	0.002	0.076 *	0.967	-0.033	0.026	0.198	1.009	0.009	0.003	0.001 ***	1.009	0.009	0.003	0.001 ***
Active Participation Change (Delta)		0.995	-0.005	0.001	0.000 ***	0.965	-0.036	0.015	0.019 **	0.996	-0.004	0.001	0.000 ***	0.996	-0.004	0.001	0.000 ***
Verified Membership		0.371	-0.991	0.132	0.000 ***	0.559	-0.581	0.143	0.000 ***	0.329	-1.113	0.346	0.001 ***	0.352	-1.044	0.345	0.002 ***
Age		1.014	0.013	0.002	0.000 ***	1.008	0.008	0.003	0.003 ***	1.009	0.009	0.006	0.147	1.002	0.002	0.006	0.744
Gender		0.974	-0.026	0.030	0.381	0.961	-0.040	0.033	0.230	0.897	-0.109	0.072	0.133	0.862	-0.149	0.074	0.045 **
Time-dependent Effects of Covariates [1]																	
Degree (Active Contacts)	x g(t) [2]	1.037	0.036	0.003	0.000 ***	-	-	-	-	1.008	0.008	0.002	0.001 ***	1.006	0.006	0.002	0.006 ***
Avg. Degree of Active Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Betweenness (Active Contacts)	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ego-Network Density (Active Contacts)	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	1.462	0.380	0.172	0.028 **
Similarity (Active Contacts) - Region	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Similarity (Active Contacts) - Distance	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	0.998	-0.002	0.001	0.071 *
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Share of Inactive Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Active Participation Level	x g(t)	-	-	-	-	1.014	0.014	0.013	0.311	0.993	-0.007	0.002	0.000 ***	0.993	-0.007	0.002	0.000 ***
Active Participation Delta	x g(t)	-	-	-	-	1.014	0.014	0.007	0.050 *	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	1.332	0.287	0.061	0.000 ***	1.222	0.200	0.077	0.010 **	1.433	0.360	0.138	0.009 ***	1.412	0.345	0.138	0.012 **
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Subjects				5497				4132				1365				1365	
Failures				4553				3764				789				789	
Observations				50806				29320				21486				21486	
Log-Likelihood				-35861				-28443				-5316				-5237	
AIC				71738				56902				10650				10509	
BIC				71809				56968				10722				10653	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; PH assumption tests executed for each model seperately; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) g(t)=ln(t)

Appendix 18: Study 3 – Verification Test with Cohort 2

Appendix 19: Study 3 – Verification Test with Stratified Model by User Group

COHORT 1		**Model 1b**				
		All Users - Full Model				
Covariates		**HR**	**Coef.**	**SE**	**p**	
Main Effects of Covariates						
Degree (Active Contacts)		0.966	-0.035	0.008	0.000	***
Avg. Degree of Active Contacts		1.002	0.002	0.000	0.000	***
Betweenness (Active Contacts)		0.926	-0.077	0.021	0.000	***
Ego-Network Density (Active Contacts)		0.448	-0.803	0.346	0.020	**
Similarity (Active Contacts) - Region		3.764	1.326	0.255	0.000	***
Similarity (Active Contacts) - Distance		1.010	0.010	0.002	0.000	***
Similarity (Active Contacts) - Gender		1.367	0.313	0.100	0.002	***
Share of Inactive Contacts		2.045	0.715	0.153	0.000	***
Active Participation Level		0.996	-0.004	0.002	0.088	*
Active Participation Change (Delta)		0.996	-0.004	0.001	0.004	***
Verified Membership		0.726	-0.320	0.063	0.000	***
Age		1.003	0.003	0.003	0.275	
Gender		0.839	-0.175	0.048	0.000	***
Time-dependent Effects of Covariates [1)]						
Degree (Active Contacts)	x g(t) [2)]	1.012	0.011	0.003	0.000	***
Avg. Degree of Active Contacts	x g(t)	0.999	-0.001	0.000	0.001	***
Betweenness (Active Contacts)	x g(t)	1.016	0.016	0.008	0.054	*
Ego-Network Density (Active Contacts)	x g(t)	1.302	0.264	0.145	0.069	*
Similarity (Active Contacts) - Region	x g(t)	0.643	-0.442	0.109	0.000	***
Similarity (Active Contacts) - Distance	x g(t)	0.998	-0.002	0.001	0.086	*
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	
Share of Inactive Contacts	x g(t)	-	-	-	-	
Active Participation Level	x g(t)	-	-	-	-	
Active Participation Delta	x g(t)	-	-	-	-	
Verified Membership	x g(t)	-	-	-	-	
Age	x g(t)	-	-	-	-	
Gender	x g(t)	1.085	0.082	0.082	0.002	***
Subjects				5752		
Failures				4775		
Observations				55488		
Log-Likelihood				-35474		
AIC				70987		
BIC				71166		
		Stratified by Friend Group				

**** p<.01, ** p<.05, * p<.10*
Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value
1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests;
PH assumption tests executed for each model seperately;
time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)
2) g(t)=ln(t)

Model 1	Model changed by:	Binary Participation[3]			Transf. Degree[4]			Stratified by Region[5]			Without Outliers[6]		
Covariates		**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**
Main Effects of Covariates													
Degree (Active Contacts)		-0.124	0.009	0.000 ***	-1.233	0.076	0.000 ***	-0.116	0.010	0.000 ***	-0.141	0.011	0.000 ***
Active Participation Level		-0.290	0.086	0.001 ***	-0.019	0.011	0.078 *	-0.034	0.012	0.004 ***	-0.074	0.018	0.000 ***
Active Participation Delta		-0.005	0.001	0.000 ***	-0.004	0.001	0.001 ***	-0.004	0.001	0.003 ***	0.001	0.003	0.847
Verified Membership		-0.818	0.125	0.000 ***	-0.662	0.126	0.000 ***	-0.803	0.126	0.000 ***	-0.794	0.125	0.000 ***
Age		0.009	0.003	0.001 ***	0.005	0.003	0.079 *	0.008	0.003	0.001 ***	0.008	0.003	0.002 ***
Gender		-0.178	0.048	0.000 ***	-0.179	0.048	0.000 ***	-0.167	0.048	0.000 ***	-0.177	0.048	0.000 ***
Time-dependent Effects of Covariates [1]													
Degree (Active Contacts)	x g(t) [2]	0.040	0.003	0.000 ***	0.347	0.029	0.000 ***	0.038	0.003	0.000 ***	0.045	0.004	0.000 ***
Active Participation Level	x g(t)	-0.086	0.050	0.087 *	0.006	0.004	0.113	0.011	0.004	0.005 ***	0.023	0.006	0.000 ***
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	0.167	0.059	0.004 ***	0.151	0.059	0.010 **	0.145	0.059	0.014 **	0.160	0.059	0.006 ***
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.096	0.026	0.000 ***	0.092	0.026	0.000 ***	0.079	0.027	0.003 ***	0.097	0.026	0.000 ***
Subjects			5752			5752			5752			5746	
Failures			4775			4775			4775			4769	
Observations			55488			55488			55488			55426	
Log-Likelihood			-37766			-37610			-26773			-37686	
AIC			75553			75240			53566			75393	
BIC			75642			75329			53656			75482	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for g(t) = ln(t)

3) Active Participation Level is replaced by a binary variable (participated in month t: yes/no) in 'Binary Participation' model

4) Degree is log-transformed in 'Transformed Degree' model

5) Model is stratified by the variabel 'Region' in that the user is registered

6) Outliers identified by DFBeta analysis and omitted in "Without Outliers" model

Appendix 20a: Study 3 – Verification Test with Alternative Independent Variables for Model 1

Model 2

Model changed by:		Binary Participation[3]			Stratified by Region[4]			Without Outliers[5]		
Covariates		**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**	**Coef.**	**SE**	**p**
Main Effects of Covariates										
Active Participation Level		-0.128	0.099	0.196	-0.039	0.026	0.127	-0.038	0.026	0.141
Active Participation Delta		-0.012	0.005	0.027 **	-0.010	0.006	0.084 *	-0.025	0.013	0.058 *
Verified Membership		-0.439	0.136	0.001 ***	-0.427	0.137	0.002 ***	-0.434	0.136	0.001 ***
Age		0.002	0.003	0.532	0.003	0.003	0.257	0.002	0.003	0.528
Gender		-0.159	0.049	0.001 ***	-0.077	0.033	0.020 **	-0.156	0.049	0.002 ***
Time-dependent Effects of Covariates [1]										
Active Participation Level	x g(t) [2]	0.282	0.082	0.001 ***	0.024	0.013	0.069 *	-	-	-
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	0.146	0.075	0.052 *	0.111	0.076	0.144	0.143	0.075	0.057 *
Age	x g(t)	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.066	0.030	0.029 **	-	-	-	0.065	0.030	0.032 **
Subjects			4264			4264			4257	
Failures			3901			3901			3894	
Observations			31121			31121			31075	
Log-Likelihood			-29579			-20845			-29525	
AIC			59174			41703			59063	
BIC			59241			41762			59122	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for g(t) = ln(t)

3) Active Participation Level is replaced by a binary variable (participated in month t: yes/no) in 'Binary Participation' model

4) Model is stratified by the variabel 'Region' in that the user is registered

5) Outliers identified by DFBeta analysis and omitted in "Without Outliers" model

Appendix 20b: Study 3 – Verification Test with Alternative Independent Variables for Model 2

Model 3b													
Model changed by:		Binary Participation[3)]			Transf. Degree[4)]			Stratified by Region[5)]			Without Outliers[6)]		
Covariates		Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p
Main Effects of Covariates													
Degree (Active Contacts)		-0.032	0.008	0.000 ***	-0.328	0.052	0.000 ***	-0.028	0.008	0.001 ***	-0.061	0.017	0.000 ***
Avg. Degree of Active Contacts		0.002	0.000	0.000 ***	0.002	0.000	0.000 ***	0.000	0.000	0.001 ***	0.003	0.001	0.000 ***
Betweenness (Active Contacts)		-0.088	0.022	0.000 ***	-0.079	0.020	0.000 ***	-0.087	0.022	0.000 ***	-0.105	0.028	0.000 ***
Ego-Network Density (Active Contacts)		-0.780	0.344	0.024 **	-0.580	0.343	0.091 *	-0.949	0.349	0.006 ***	-2.595	0.551	0.000 ***
Similarity (Active Contacts) - Region		1.321	0.256	0.000 ***	1.372	0.252	0.000 ***	1.947	0.335	0.000 ***	3.244	0.421	0.000 ***
Similarity (Active Contacts) - Distance		0.011	0.002	0.000 ***	0.011	0.002	0.000 ***	0.032	0.005	0.000 ***	0.037	0.005	0.000 ***
Similarity (Active Contacts) - Gender		0.319	0.101	0.002 ***	0.351	0.101	0.001 ***	0.235	0.103	0.022 ***	0.263	0.103	0.011 **
Share of Inactive Contacts		0.691	0.154	0.000 ***	0.634	0.154	0.000 ***	0.535	0.164	0.001 ***	0.788	0.170	0.000 ***
Active Participation Level		0.447	0.234	0.056 *	-0.025	0.014	0.069 *	-0.003	0.002	0.109	-0.052	0.013	0.000 ***
Active Participation Delta		-0.003	0.001	0.002 ***	-0.004	0.001	0.006 ***	-0.003	0.001	0.027 **	0.007	0.006	0.280
Verified Membership		-0.420	0.100	0.000 ***	-0.422	0.100	0.000 ***	-0.453	0.101	0.000 ***	-0.373	0.104	0.000 ***
Age		0.009	0.006	0.123	0.005	0.006	0.377	0.007	0.006	0.248	0.005	0.006	0.480
Gender		-0.438	0.198	0.027 **	-0.487	0.198	0.014 **	-0.373	0.201	0.063 *	-0.572	0.240	0.017 **
Time-dependent Effects of Covariates [1)]													
Degree (Active Contacts)	x g(t) [2)]	0.011	0.003	0.000 ***	-	-	-	0.009	0.003	0.001 ***	0.019	0.006	0.001 ***
Avg. Degree of Active Contacts	x g(t)	-0.001	0.000	0.001 ***	-0.001	0.000	0.001 ***	-	-	-	-0.001	0.000	0.007 ***
Betweenness (Active Contacts)	x g(t)	0.025	0.008	0.004 ***	0.034	0.008	0.000 ***	0.020	0.009	0.020 **	0.033	0.011	0.003 ***
Ego-Network Density (Active Contacts)	x g(t)	0.250	0.145	0.084 *	0.248	0.143	0.082 *	0.336	0.146	0.022 **	0.935	0.216	0.000 ***
Similarity (Active Contacts) - Region	x g(t)	-0.443	0.109	0.000 ***	-0.411	0.108	0.000 ***	-0.557	0.142	0.000 ***	-1.153	0.168	0.000 ***
Similarity (Active Contacts) - Distance	x g(t)	-0.002	0.001	0.049 **	-0.002	0.001	0.065 *	-0.009	0.002	0.000 ***	-0.011	0.002	0.000 ***
Similarity (Active Contacts) - Gender	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Share of Inactive Contacts	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Active Participation Level	x g(t)	-0.447	0.102	0.000 ***	0.008	0.005	0.090 *	-	-	-	-	-	-
Active Participation Delta	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Verified Membership	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Age	x g(t)	-	-	-	-	-	-	-	-	-	-	-	-
Gender	x g(t)	0.198	0.082	0.015 **	0.214	0.082	0.009 ***	0.165	0.083	0.047 **	0.255	0.097	0.008 ***
Subjects			1488			1488			1488			1390	
Failures			874			874			874			801	
Observations			24367			24367			24367			23353	
Log-Likelihood			-5862			-5874			-3830			-5215	
AIC			11767			11788			7698			10470	
BIC			11937			11950			7852			10631	

**** p<.01, ** p<.05, * p<.10*

Note: HR=Hazard Ratio; Coef.=Coefficient; SE=Standard Error; p=p-value

1) time-dependent effects are only included for variables with non-proportional hazards, indicated by PH assumption tests; time-dependent effects calculated as interaction terms of independent variables with a function of time g(t)

2) All models estimated for g(t) = ln(t)

3) Active Participation Level is replaced by a binary variable (participated in month t: yes/no) in 'Binary Participation' model

4) Degree is log-transformed in 'Transformed Degree' model

5) Model is stratified by the variabel 'Region' in that the user is registered

6) Outliers identified by DFBeta analysis and omitted in "Without Outliers" model

Appendix 20c: Study 3 – Verification Test with Alternative Independent Variables for Model 3b

List of Appendixes

References

Achrol, Ravi S. and Philip Kotler (1999), "Marketing in the Network Economy," *Journal of Marketing*, 63 (Special Issue), 146-63.

Adjei, Mavis T., Stephanie M. Noble, and Charles H. Noble (2010), "The influence of C2C communications in online brand communities on customer purchase behavior," *Journal of the Academy of Marketing Science*, 38 (5), 634-653.

Adler, Paul S. and Seok-Woo Kwon (2002), "Social Capital: Prospects for a New Concept," *Academy of Management Review*, 27 (January), 17-40.

Ahuja, Manju K., Dennis F. Galletta, and Kathleen M. Carley (2003), "Individual Centrality and Performance in Virtual R&D Groups: An Empirical Study," *Management Science*, 49 (January), 21-38.

Ajzen, Icek (1991), "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processess*, 50 (2), 179-211.

Alarcon-del-Amo, Maria-del-Carmen, Carlota Lorenzo-Romero, and Miguel-Angel Gomez-Borja (2011), "Classifying and Profiling Social Networking Site Users: A Latent Segmentation Approach," *Cyberpsychology, Behavior, and Social Networking*, 14 (9), 547-553.

Alexa Internet Inc. (2011), "Top Sites in United States," (accessed November 3, 2011), [available at http://www.alexa.com/topsites/countries/US].

Algesheimer, René, Sharad Borle, Utpal M. Dholakia, and Siddharth S. Singh (2010), "The Impact of Customer Community Participation on Customer Behaviors: An Empirical Investigation," *Marketing Science*, 29 (July-August), 756-769.

Algesheimer, René, Utpal M. Dholakia, and Andreas Herrmann (2005), "The Social Influence of Brand Community: Evidence from European Car Clubs," *Journal of Marketing*, 69 (July), 19-34.

Algesheimer, René and Florian Wangenheim (2006), "A Network Based Approach to Customer Equity Management," *Journal of Relationship Marketing*, 5 (April), 39-57.

Allison, Paul D. (2010), *Survival Analysis Using SAS®: A Practical Guide*. Cary, NC: SAS Institute Inc., second edition.

Andersen, P. K. and R. D. Gill (1982), "Cox's Regression Model for Counting Processes: A Large Sample Study," *The Annals of Statistics*, 10 (December), 1100-1120.

Anderson, Eugene and Claes Fornell (1994), "A Customer Satisfaction Research Prospectus," in *Service quality: New directions in theory and practice*, Roland T. Rust and Richard L. Oliver, eds. Thousand Oaks, CA: Sage, 241-268.

Armstrong, Arthur and John Hagel III (1996), "The Real Value of On-Line Communities," *Harvard Business Review*, 74 (May-June), 134-141.

Armstrong, J. Scott and Terry S. Overton (1977), “Estimating Non-Response Bias in Mail Surveys,” *Journal of Marketing Research*, 14 (August), 396–402.

Arndt, Johan (1967). “Role of Product-Related Conversations in the Diffusion of a New Product,” *Journal of Marketing Research*, 4 (August), 291-295.

Assael, Henry (1998), *Consumer Behavior and Marketing Action*. Cincinnati: South-Western College Publishing, sixth edition.

Ba, Sulin (2001), “Establishing Online Trust through a Community Responsibility System,” *Decision Support Systems*, 31 (3), 323-336.

Bagozzi, Richard P. (1994), “Structural equation model in marketing research: Basic principles,” in *Principles of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell Publishers, 317-385.

Bagozzi, Richard P. and Utpal M. Dholakia (2002), “Intentional social action in virtual communities,” *Journal of Interactive Marketing*, 16 (Spring), 2-21.

Bagozzi, Richard P. and Youjae Yi (1988), “On the Evaluation of Structural Equation Models,” *Journal of the Academy of Marketing Science*, 16 (Spring), 74-94.

Bagozzi, Richard P. and Youjae Yi (1989), “On the Use of Structural Equation Models in Experimental Designs,” *Journal of Marketing Research*, 26 (August), 271-284.

Balasubramanian, Sridhar and Vijay Mahajan (2001), “The Economic Leverage of the Virtual Community,” *International Joumal of Electrotnic Commerce*, 5 (3), 103-138.

Bandura, Albert (1982), “Self-Efficacy Mechanism in Human Agency,” *American Psychologist*, 37 (February), 122-147.

Bansal, Harvir S. and Voyer, Peter A. (2000), “Word-of-Mouth Processes within a Services Purchase Decision Context,” *Journal of Service Research*, 3 (November), 166-177.

Barlow, Alexis K. J., Noreen Q. Siddiqui, and Mike Mannion (2004), “Developments in information and communication technologies for retail marketing channels,” *International Journal of Retail & Distribution Management*, 32 (3), 157-163.

Barnes, J. A. (1954), “Class and Committees in a Norwegian Island Parish”, *Human Relations*, 7 (1), 39-58.

Baron, Reuben M. and David A. Kenny (1986), “The Moderator–Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations,” *Journal of Personality and Social Psychology*, 51 (6), 1173–1182.

Batagalj, Vladimir and Andrej Mrvar (2010), “Pajek - Program for Analysis and Visualization of Large Networks - Reference Manual List of commands with short explanation version 2.00,” (accessed June 16, 2011), [available at http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.pdf].

Bernoff, Josh (2010), “Social Technographics: Conversationalists get onto the ladder,” (accessed June 14, 2011), [available at http://forrester.typepad.com/groundswell/2010/01/conversationalists-get-onto-the-ladder.html].

Bettencourt, Lance A. (1997), “Customer voluntary performance: Customers as partners in service delivery,” *Journal of Retailing*, 73 (3), 383-406.

Bhattacharya, C. B. (1998), “When Customers Are Members: Customer Retention in Paid Membership Contexts,” *Journal of the Academy of Marketing Science*, 26 (1), 31-44.

Bhattacharya, C. B., Hayagreeva Rao, and Mary Ann Glynn (1995), “Understanding the Bond of Identification: An Investigation of Its Correlates among Art Museum Members,” *Journal of Marketing*, 59 (October), 46-57.

Bhattacherjee, Anol (2001), “Understanding Information Systems Continuance: An Expectation-Confirmation Model,” *MIS Quarterly*, 25 (3), 351-370.

Blattberg, Robert C. and John Deighton (1996), “Manage Marketing by the Customer Equity Test,” *Harvard Business Review*, 74 (July-August), 136-144.

Blau, Peter M. (1964), *Exchange and Power in Social Life*. New York: John Wiley & Sons.

Bolton, Ruth N. (1998), “A Dynamic Model of the Duration of the Customer’s Relationship with a Continuous Service Provider: The Role of Satisfaction”, *Marketing Science*, 17 (1), 45-65.

Bolton, Ruth N. and Katherine N. Lemon (1999), “A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction,” *Journal of Marketing Research*, 36 (May), 171-186.

Bolton, Ruth N., Katherine N. Lemon, and Peter C. Verhoef (2004), “The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research,” *Journal of the Academy of Marketing Science*, 32 (3), 1-20.

Bone, Paula Fitzgerald (1995), “Word-of-mouth effects on short-term and long-term product judgments,” *Journal of Business Research*, 32 (March), 213-223.

Borden, Neil H. (1964), “The Concept of the Marketing Mix,” *Journal of Advertising Research*, 4 (June), 2-7.

Borgatti, Stephen P. (2005), “Centrality and network flow,” *Social Networks*, 27 (1), 55-71.

Borgatti, Stephen P., Candace Jones, and Martin G. Everett (1998), “Network Measures of Social Capital,” *Connections*, 21 (2), 27-36.

Borgatti, Stephen P., Ajay Mehra, Daniel J. Brass, Giuseppe Labianca (2009), “Network Analysis in the Social Sciences,” *Science*, 323 (5916), 892-895.

Bourdieu, Pierre (1986), “The Forms of Capital,” in *Handbook of Theory and Research for the Sociology of Education*, John G. Richardson, ed. New York: Greenwood, 241-58.

Box-Steffensmeier, Janet M., Dan Reiter, and Christopher J. W. Zorn (2003), “Nonproportional Hazards and Event History Analysis in International Relations,” *Journal of Conflict Resolution*, 47 (February), 33-53.

Box-Steffensmeier, Janet M. and Christopher J. W. Zorn (2001), “Duration Models and Proportional Hazards in Political Science,” *American Journal of Political Science*, 45 (October), 972-988.

Boyd, Dana M. and Nicole B. Ellison (2008), “Social Network Sites: Definition, History and Scholarship,” *Journal of Computer-Mediated Communication*, 13 (1), 210-230.

Brass, Daniel J. (1984), “Being in the Right Place: A Structural Analysis of Individual Influence in an Organization,” *Administrative Science Quarterly*, 29 (December), 518-539.

Brass, Daniel J. and Marlene E. Burkhardt (1993), “Potential Power and Power Use: An Investigation of Structure and Behavior,” *The Academy of Management Journal*, 36 (June), 441-470.

Breslow, N. (1974), “Covariance Analysis of Censored Survival Data,” *Biometrics*, 30 (March), 89-99.

Brown, Jacqueline Johnson and Peter H. Reingen (1987), “Social Ties and Word-of-Mouth Referral Behavior,” *Journal of Consumer Research*, 14 (3), 350-362.

Brown, Jo, Amanda J. Broderick, and Nick Lee (2007), “Word of mouth communication within online communities: Conceptualizing the online social network,” *Journal of Interactive Marketing*, 21 (3), 2-20.

Brown, Rupert (2000), “Social Identity Theory: past achievements, current problems and future challenges,” *European Journal of Social Psychology*, 30 (6), 745-778.

Burt, Ronald S. (1980), “Models of Network Structure,” *Annual Review of Sociology*, 6, 79-141.

Burt, Ronald S. (1992), “The Social Structure of Competition,” in *Networks and Organizations: Structure, Form, and Action*, Nitin Nohria and Robert G. Eccles, eds. Boston: Harvard Business School Press, 57-91.

Burt, Ronald S. (1997), “The Contingent Value of Social Capital,” *Administrative Science Quarterly*, 42 (June), 339-365.

Burt, Ronald S. (2000), “The Network Structure of Social Capital,” in *Research in Organizational Behavior*, Barry M. Staw and Robert I. Sutton, eds. Amsterdam, London and New York: Elsevier Science, 345-423.

Butler, Brian S. (2001), “Membership Size, Communication Activity, Sustainability: A Resource-Based Model of Online Social Structures”, *Information Systems Research*, 12 (4) (December), 346-362.

Byrne, Barbara M. (2010), *Structural Equation Modeling with AMOS: Basic Concpets, Applications, and Programming*. New York: Routledge, second edition.

Cameron, Adrian C. and Pravin K. Trivedi (1998), *Regression Analysis of Count Data*. Cambridge: University Press.

Cappelli, Peter (2000), “A Market-Driven Approach to Retaining Talent,” *Harvard Business Review*, 78 (1), 103-111.

Casalo, Luis V., Carlos Flavian, and Miguel Guinaliu (2007), “The impact of participation in virtual brand communities on consumer trust and loyalty: The case of free software,” *Online Information Review*, 31 (6), 775-792.

Casalo, Luis V., Carlos Flavian, and Miguel Guinaliu (2008a), “Fundaments of trust management in the development of virtual communities,” *Management Research News*, 31 (5), 324-338.

Casalo, Luis V., Carlos Flavian, and Miguel Guinaliu (2008b), “Promoting Consumer's Participation in Virtual Brand Communities: A New Paradigm in Branding Strategy,” *Journal of Marketing Communications*, 14 (1), 19-36.

Casalo, Luis V., Jesus Cisneros, Carlos Flavian, and Miguel Guinaliu (2009), “Determinants of success in open source software networks,” *Industrial Management & Data Systems*, 109 (4), 532-549.

Castilla, Emilio J. (2005), “Social Networks and Employee Performance in a Call Center,” *American Journal of Sociology*, 110 (5), 1243–1283.

Chan, Darius K.-S. and Grand H.-L. Cheng (2004), “A Comparison of Offline and Online Friendship Qualities at Different Stages of Relationship Development,” *Journal of Social and Personal Relationships*, 21 (June), 305-320.

Chen, Chih-Jou and Shiu-Wan Hung (2010), “To give or to receive? Factors influencing members’ knowledge sharing and community promotion in professional virtual communities,” *Information & Management*, 47 (May), 226-236.

Chen, Irene Y. L. (2007), “The factors influencing members' continuance intentions in professional virtual communities,” *Journal of Information Science*, 33 (4), 451-467.

Chen, Yubo and Jinhong Xie (2008), “Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix,” *Management Science*, 54 (March), 477-491.

Cheung, Christy M.K. and Matthew K.O. Lee (2009), “Understanding the sustainability of a virtual community: model development and empirical test,” *Journal of Information Science*, 35 (3), 279-298.

Cheung, Christy M.K. and Matthew K.O. Lee (2010), "A theoretical model of intentional social action in online social networks," *Decision Support Systems*, 49 (1), 24-30.

Chevalier, Judith A. and Dina Mayzlin (2006), "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research*, 43 (August), 345-354.

Chiu, Chao-Min, Meng-Hsiang Hsu, and Eric T.G. Wang (2006), "Understanding Knowledge Sharing in Virtual Communities: An Integration of Social Capital and Social Cognitive Theories," *Decision Support Systems*, 42 (December), 1872-1888.

Chou, Nan-Ting and David Steenhard (2009), "A Flexible Count Data Regression Model Using SAS® PROC NLMIXED," *SAS Global Forum 2009 Statistics and Data Analysis, Paper 250-2009* (accessed May 30, 2011), [available at http://support.sas.com/resources/papers/proceedings09/250-2009.pdf].

Christakis, Nicholas A. and James H. Fowler (2008), "The Collective Dynamics of Smoking in a Large Social Network", *New England Journal of Medicine*, 358 (21), 2249-2258.

Chu, Kuo-Ming (2009), "A study of members' helping behaviors in online community," *Internet Research*, 19 (3), 279-292.

Chung, Jin Young and Dimitrios Buhalis (2008), "Information Needs in Online Social Networks," *Information Technology & Tourism*, 10 (4), 267-281.

Churchill, Gilbert A., Neil M. Ford, Steven W. Hartley, and Orville C. Walker (1985), "The Determinants of Salesperson Performance: A Meta-Analysis," *Journal of Marketing Research*, 22 (May), 103-118.

Clark, Margaret S., Judson Mills, and Martha C. Powell (1986), "Keeping Track of Needs in Communal and Exchange Relationships," Journal *of Personality and Social Psychology*, 51 (August), 333-338.

Clemons, Eric K. (2009), "The complex problem of monetizing virtual electronic social networks," *Decision Support Systems*, 48 (December), 46-56.

Coleman, James (1988), "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94 (Suppl.), 95-120.

Coleman, James (1990), *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Collett, David (2003), *Modelling Survival Data in Medical Research*. Boca Raton/London: Chapman & Hall/CRC, second edition.

Cook, K. S. and J. M. Whitmeyer (1992), "Two Approaches to Social Structure: Exchange Theory and Network Analysis," *Annual Review of Sociology*, 18, 109-127.

Cothrel, Joseph P. (2000), "Measuring the success of an online community," *Strategy & Leadership*, 28 (2), 17-21.

Cox, D. R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 34 (2), 187-220.

Crosby, Lawrence A., Kenneth R. Evans, and Deborah Cowles (1990), "Relationship Quality in Services Selling: An Interpersonal Influence Perspective," *Journal of Marketing*, 54 (July), 68-81.

Cross, Rob and Jonathon N. Cummings (2004), "Tie and Network Correlates of Individual Performance in Knowledge-Intensive Work," *The Academy of Management Journal*, 47 (December), 928-937.

Cummings, Jonathon N., Brian Butler, and Robert Kraut (2002), "The quality of online social relationships," *Communications of the ACM*, 45 (July), 103-108.

Dash, Satyabhusan and K. B. Saji (2007), "The Role of Consumer Self-Efficacy and Website Social-Presence in Customers' Adoption of B2C Online Shopping: An Empirical Study in the Indian Context," *Journal of International Consumer Marketing*, 20 (2), 33-48.

Davis, Fred D. (1989), "Perceived Usefulness, Perceived Ease of Use, and end User Acceptance of Information Technology," *MIS Quarterly*, 13 (3), 319-340.

Davis, Fred D., Richard P. Bagozzi, and Paul R. Warshaw (1989), "User Acceptance of Computer Technology: a Comparison of Two Theoretical Models," *Management Science*, 35 (8), 982-1003.

de Bruyn, Arnaud and Gary L. Lilien (2008), "A Multi-Stage Model of Word of Mouth Through Viral Marketing," *International Journal of Research in Marketing*, 25 (September), 151-163.

de Matos, Celso Augusto and Carlos Alberto Vargas Rossi (2008), "Word-of-mouth communications in marketing: a meta-analytic review of the antecedents and moderators," *Journal of the Academy of Marketing Science*, 36 (4), 578-596.

de Nooy, Wouter, Andrej Mrvar, and Vladimir Batagelj (2005), *Exploratory Social Network Analysis with Pajek*. Cambridge: University Press.

de Valck, Kristine, Fred Langerak, Peter C. Verhoef, and Peeter W. J. Verlegh (2007), "Satisfaction with Virtual Communities of Interest: Effect on Members' Visit Frequency," *British Journal of Management*, 18 (September), 241-256.

de Valck, Kristine, Gerrit H. van Bruggen, and Berend Wierenga (2009), "Virtual communities: A marketing perspective," *Decision Support Systems*, 47 (June), 185-203.

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

Delone, William H. and Ephraim R. McLean (2003), "The DeLone and McLean Model of Information Systems Success: A Ten-Year Update," *Journal of Management Information Systems*, 19 (4), 9-30.

Dess, Gregory G. and Jason D. Shaw (2001), "Voluntary Turnover, Social Capital, and Organizational Performance," *The Academy of Management Review*, 26 (July), 446-456.

Dholakia, Utpal M., Richard P. Bagozzi, and Lisa Klein Pearo (2004), "A Social Influence Model of Consumer Participation in Network- and Small-Group-Based Virtual Communities", *International Journal of Research in Marketing*, 21 (3), 241-263.

Dholakia, Utpal, Vera Blazevic, Caroline Wiertz, and Rene Algesheimer (2009), "Communal Service Delivery: How Customers Benefit From Participation in Firm-Hosted Virtual P3 Communities," *Journal of Service Research*, 12 (2), 208-226.

Dover, Howard F. and B. P. S. Murthi (2006), "Asymmetric Effects of Dynamic Usage Behavior on Duration in Subscriptionbased Online Services," *Journal of Interactive Marketing*, 20 (3-4), 5-15.

Duncan, Tom and Sandra E. Moriarty (1998), "A Communication-Based Marketing Model for Managing Relationships," *Journal of Marketing*, 62 (April), 1-13.

Efron, Bradley (1977), "The Efficiency of Cox's Likelihood Function for Censored Data," *Journal of the American Statistical Association*, 72 (September), 557-565.

Ellemers, Naomi, Paulien Kortekaas, and Jaap W. Ouwerkerk (1999), "Self-Categorization, Commitment to the Group and Group Self-Esteem as Related but Distinct Aspects of Social Identity," *European Journal of Social Psychology*, 29 (2-3), 371-389.

eMarketer (2011), "Facebook Revenues to Reach $4.27 Billion in 2011," (accessed November 7, 2011), [available at http://www.emarketer.com/PressRelease.aspx?R=1008601].

Emerson, Richard M. (1976), "Social Exchange Theory," *Annual Review of Sociology*, 2, 335-362.

Enders, Albrecht, Harald Hungenberg, Hans-Peter Denker, and Sebastian Mauch (2008), "The long tail of social networking. Revenue models of social networking sites," *European Management Journal*, 26 (3), 199-211.

Engel, James E., Robert J. Kegerreis, and Roger D. Blackwell (1969), "Word-of-Mouth Communication by the Innovator," *Journal of Marketing*, 33 (July), 15-19.

Erdman, Donald, Laura Jackson, and Arthur Sinko (2008), "Zero-Inflated Poisson and Zero-Inflated Negative Binomial Models Using the COUNTREG Procedure," *SAS Global Forum 2008, Paper 322-2008* (accessed May 30, 2011), [available at http://www2.sas.com/proceedings/forum2008/322-2008.pdf].

Evans, Franklin B. (1963), "Selling as a Dyadic Relationship – a New Approach," *The American Behavioral Scientist*, 6 (May) 76-79.

Facebook (2011), "Statistics", (accessed March 22, 2011), [available at http://www.facebook.com/press/info.php?statistics].

Field, Andy (2005), *Discovering Statistics Using SPSS*. London: Sage Publications, second edition.

Fisher, Lloyd D. and D. Y. Lin (1999), "Time-Dependent Covariates in the Cox Proportional-Hazards Regression Model," *Annual Review of Public Health*, 20, 145-157.

Flanagin, Andrew J. and Miriam J. Metzger (2001), "Internet use in the contemporary media environment," *Human Communication Research*, 27 (January), 153-81.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39-50.

Fornell, Claes, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara Everitt Bryant (1996), "The American Customer Satisfaction Index: Nature, Purpose, and Findings," *Journal of Marketing*, 60 (October), 7-18.

Freeman, Linton C. (1978/79), "Centrality in social networks conceptual clarification," *Social Networks*, 1 (3), 215-239.

Freeman, Linton C. (2004), *The Development of Social Network Analysis: A Study in the Sociology of Science*. North Charleston, South Carolina: BookSurge.

Fulk, Janet, Charles W. Steinfield, Joseph Schmitz, and J. Gerard Power (1987), "A Social Information Processing Model of Media Use in Organizations," *Communication Research*, 14 (5), 529-552.

Garbarino, Ellen and Mark S. Johnson (1999), "The Different Roles of Satisfaction, Trust, and Commitment in Customer Relationships," *Journal of Marketing*, 63 (April), 70-87.

Gartner Inc. (2011), "Gartner Survey Highlights Consumer Fatigue with Social Media," (accessed August 29, 2011), [available at http://www.gartner.com/it/page.jsp?id=1766814].

Gilly, Mary C., John L. Graham, Mary Finley Wolfinbarger, and Laura J. Yale (1998), "A Dyadic Study of Interpersonal Information Search," *Journal of the Academy of Marketing Science*, 26 (Spring), 83-100.

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23 (4), 545-61.

Goldenberg, Jacob, Sangman Han, Donald R. Lehmann, and Jae Weon Hong (2009), "The Role of Hubs in the Adoption Process," *Journal of Marketing*, 73 (March), 1-13.

Golub, Jonathan and Bernard Steunenberg (2007), "How Time Affects EU Decision-Making," *European Union Politics*, 8 (4), 555-566.

Gouldner, Alvin W. (1960), "The Norm of Reciprocity: A Preliminary Statement," *American Sociological Review*, 25 (2), 161-78.

Grambsch, Patricia M. and Terry M. Therneau (1994), “Proportional Hazards Tests and Diagnostics Based on Weighted Residuals,” *Biometrika*, 81 (August), 515-526.

Granovetter, Mark S. (1973), “The Strength of Weak Ties,” *American Journal of Sociology*, 78 (May), 1360-1380.

Granovetter, Mark S. (1978), “Threshold Models of Collective Behavior”, *American Journal of Sociology*, 83 (6), 1420-1443.

Granovetter, Mark S. (1983), “The Strength of Weak Ties: A Network Theory Revisited,” *Sociological Theory*, 1, 201-233.

Granovetter, Mark S. (1985), “Economic Action and Social Structure: The Problem of Embeddedness”, *American Journal of Sociology*, 91 (3), 481-510.

Greene, William H. (2002), *Econometric Analysis*. New Jersey: Pearson Education - Prentice Hall, fifth edition.

Gremler, Dwayne D., Kevin P. Gwinner, and Stephen W. Brown (2001), “Generating positive word-of-mouth communication through customer-employee relationships,” *International Journal of Service Industry Management*, 12 (1), 44-59.

Gross, Ralph and Alessandro Acquisti (2005), “Information Revelation and Privacy in Online Social Networks”, *WPES ’05 Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*.

Gruen, Thomas W., John O. Summers, and Frank Acito (2000), “Relationship Marketing Activities, Commitment, and Membership Behaviors in Professional Associations,” *Journal of Marketing*, 64 (July), 34-49.

Gustafsson, Anders, Michael D. Johnson, and Inger Roos (2005), “The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention,” *Journal of Marketing*, 69 (October), 210-218.

Hagel, John III (1999), “Net Gain: Expanding Markets Through Virtual Communities,” *Journal of Interactive Marketing*, 13 (1), 55-65.

Hagel, John III and Arthur G. Armstrong (1997), *Net Gain: Expanding Markets through virtual Communities*. Boston, MA: Harvard Business School Press.

Hagel, John III and Arthur G. Armstrong (2006), *Net Gain - Profit im Netz: Märkte erobern mit virtuellen Communitys*. Heidelberg: Redline Wirtschaft, Sonderausgabe.

Halatsis, Panayotis and Nicolas Christakis (2009), “The challenge of sexual attraction within heterosexuals’ cross-sex friendship,” *Journal of Social and Personal Relationships*, 26 (6-7), 919–937.

Hampton, Keith N., Lauren Sessions Goulet, Lee Rainie, and Kristen Purcell (2011), “Social networking sites and our lives,” Pew Research Center’s Internet & American Life Project, (June 16), (accessed September 29, 2011), [available at http://pewinternet.org/~/media//Files/Reports/2011/PIP%20-%20Social%20networking%20sites%20and%20our%20lives.pdf].

Han, Juan Juan, Raymond Jun Zheng, and Yunjie Xu (2007), “The Effect of Indicidual Needs, Trust and Identification in Explaining Participation Intentions in Virtual Communities,” *Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*, 179c.

Hansotia, Behram J. and Paul Wang (1997), “Analytical Challenges in Customer Acquisition,” *Journal of Direct Marketing*, 11 (Spring), 7-19.

Harrell, F. E. (1986), “The PHGLM Procedure,” *SUGI Supplemental Library Guide, Version 5 Edition*, Cary, NC: SAS Institute Inc.

Harrison-Walker, L. Jean (2001), “The Measurement of Word-of-Mouth Communication and an Investigation of Service Quality and Customer Commitment as Potential Antecedents,” *Journal of Service Research*, 4 (August), 60-75.

Hauben, Michael, Hauben, Ronda (1997), *Netizens - On the History and Impact of Usenet and the Internet*. Los Alamitos, CA: IEEE Computer Society Press.

Helsen, Kristiaan and David C. Schmittlein (1993), “Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models,” *Marketing Science*, 11 (4), 395-414.

Henkel, Joachim and Jan G. Sander (2007), “Identifikation innovativer Nutzer in virtuellen Communities,” in *Management der frühen Innovationsphasen*, Cornelius Herstatt and Birgit Verworn, eds. Wiesbaden: Gabler, second edition, 77-107.

Hennig-Thurau, Thorsten, Kevin P. Gwinner, and Dwayne D. Gremler (2002), “Understanding Relationship Marketing Outcomes: An Integration of Relational Benefits and Relationship Quality,” *Journal of Service Research*, 4 (3), 230-247.

Hennig-Thurau, Thorsten, Kevin P. Gwinner, Gianfranco Walsh, and Dwayne D. Gremler (2004), “Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?,” *Journal of Interactive Marketing*, 18 (Winter), 38-52.

Hennig-Thurau, Thorsten and Gianfranco Walsh (2004), “Electronic Word of Mouth: Motives for and Consequences of Reading Customer Articulations on the Internet,” *International Journal of Electronic Commerce*, 8 (2), 51-74.

Herr, Paul M., Frank R. Kardes, and John Kim (1991), “Effects of word-of-mouth and product attribute information on persuasion: An accessibility–diagnosticity perspective,” *Journal of Consumer Research*, 17 (March), 454−462.

Hilbe, Joseph M. (2007), *Negative Binomial Regression*. Cambridge: Cambridge University Press.

Hoffman, Donna L. and Thomas P. Novak (1997), "A New Marketing Paradigm for Electronic Commerce," *The Information Society*, 13 (1), 43-54.

Hogg, Michael A. (2006), "Social Identity Theory," in *Contemporary Social Psychological Theories*, Peter J. Burke, ed. Standford, CA: University Press, 111-136.

Hogg, Michael A., Deborah J. Terry, and Katherine M. White (1995), "A Tale of Two Theories: A Critical Comparison of Identity Theory with Social Identity Theory," *Social Psychology Quarterly*, 58 (4), 255-269.

Hosmer, David W., Stanley Lemeshow, and Susanne May (2008), *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. Hoboken, New Jersey: John Wiley & Sons, second edition.

Hsu, Meng-Hsiang, Teresa L. Ju, Chia-Hui Yen, and Chin-Ming Chang (2007), "Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations," *International Journal of Human-Computer Studies*, 65 (2), 153-169.

Hu, Li-tze and Peter M. Bentler (1999), "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural Equation Modeling*, 6 (1), 1-55.

Huang, Chien-Chih, Tung-Ching Lin, and Kuei-Ju Lin (2009), "Factors affecting pass-along email intentions (PAEIs): Integrating the social capital and social cognition theories," *Electronic Commerce Research and Applications*, 8 (May-June), 160-169.

Hung, Kineta H. and Stella Yiyan Li (2007), "The Influence of eWOM on Virtual Consumer Communities: Social Capital, Consumer Learning, and Behavioral Outcomes," *Journal of Advertising Research*, 47 (4), 485-95.

IAB Interactive Advertising Bureau (2011), "IAB Internet Advertising Revenue Report," (accessed September 29, 2011), [available at http://www.iab.net/media/file/IAB_Full_year_2010_0413_Final.pdf].

Jain, Dipak and Siddhartha S. Singh (2002), "Customer Lifetime Value Research in Marketing: A Review and Future Directions," *Journal of Interactive Marketing*, 16 (2), 34-46.

Jain, Dipak and Naufel J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (Winter), 1-23.

Jeppesen, Lars Bo and Lars Frederiksen (2006), "Why Do Users Contribute to Firm-Hosted User Communities? The Case of Computer-Controlled Music Instruments," *Organization Science*, 17 (1), 45-63.

Jin, Xiao-Ling, Christy M.K. Cheung, Matthew K.O. Lee, and Hua-Ping Chen (2009), "How to keep members using the information in a computer-supported social network," *Computers in Human Behavior*, 25 (5), 1172-1181.

Joinson, Adam N. (2008), "'Looking at', 'Looking up' or 'Keeping up with' People? Motives and Uses of Facebook," *CHI 2008 Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 1027-1036.

Kalbfleisch, John D. and Ross L. Prentice (1980), *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.

Kang, Inwon, Kun Chang Lee, Sangjae Lee, and Jiho Choi (2007), "Investigation of online community voluntary behavior using cognitive map," *Computers in Human Behavior*, 23 (1), 111-126.

Kankanhalli, Atreyi, Bernard C. Y. Tan, and Kwok-Kee Wei (2005), "Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation," *MIS Quarterly*, 29 (1), 113-143.

Katona, Zsolt, Peter Pal Zubcsek, and Miklos Sarvary (2011), "Network Effects and Personal Influences: The Diffusion of an Online Social Network," *Journal of Marketing Research*, 48 (June), 425-443.

Katz, Elihu and Paul F. Lazarsfeld (1955), *Personal influence: The Part Played by People in the Flow of Mass Communications*. Glencoe, IL: The Free Press.

Keller, Ed and Jon Berry (2006), "Word-of-mouth: The real action is offline," Advertising Age, (accessed March 22, 2011), [available at http://www.kellerfay.com/news/Ad%20Age%2012-4-06.pdf].

Kiecker, Pamela and Cathy L. Hartman (1994), "Predicting Buyers' Selection of Interpersonal Sources: The Role of Strong Ties and Weak Ties," in *Advances in Consumer Research*, Volume 21, Chris T. Allen and Deborah. R. John, eds. Provo, UT: Association for Consumer Research, 464-469.

Kim, Amy Jo (2000), *Community Building on the Web*. Berkeley. CA: Peachpit Press

Kim, Jae Wook, Jiho Choi, and Kyesook Han (2004), "Relationship support activities, community commitment, and voluntary behaviour in online communities," *International Journal of Internet Marketing and Advertising*, 1 (4), 431-449.

Kim, Woo Gon, Chang Lee, and Stephen J. Hiemstra (2004), "Effects of an online virtual community on customer loyalty and travel product purchases," *Tourism Management*, 25 (June), 343-355.

King, Gary (1989), "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator," *American Journal of Political Science*, 33 (3), 762-784.

Kleinbaum, David G. and Mitchel Klein (2005), *Survival Analysis: A Self-Learning Text*. New York: Springer, second edition.

Koh, Joon and Young-Gul Kim (2003/04), "Sense of Virtual Community: A Conceptual Framework and Empirical Validation," *International Journal of Electronic Commerce*, 8 (Winter), 75-93.

Koh, Joon and Young-Gul Kim (2004), "Knowledge sharing in virtual communities: an e-business perspective," *Expert Systems with Applications*, 26 (February),155-166.

Koh, Joon, Young-Gul Kim, Brian Butler, and Gee-Woo Bock (2007), "Encouraging Participation in Virtual Communities," *Communications of the ACM*, 50 (February), 69-73.

Kozinets, Robert V. (1999), "E-tribalized marketing?: the strategic implications of virtual communities of consumption," *European Management Journal*, 17 (3), 252-264.

Kozinets, Robert V. (2002), "The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities," *Journal of Marketing Research*, 39 (February), 61-72.

Kozinets, Robert V., Kristine de Valck, Andrea C. Wojnicki, and Sarah J. S. Wilner (2010), "Networked Narriatives: Understanding Word-of-Mouth Marketing in Online Communities," *Journal of Marketing*, 74 (March), 71-89.

Krackhardt, David (1992), "The Strength of Strong Ties: The Importance of Philos in Organizations," in *Networks and Organizations: Structure, Form and Action*, Nitin Nohria and Robert G. Eccles, eds. Boston, MA: Harvard Business School Press, 216-239.

Lacey, Russell, Jaebeom Suh, and Robert M. Morgan (2007), "Differential Effects of Preferential Treatment Levels on Relational Outcomes," *Journal of Service Research*, 9 (February), 241-256.

Laine, Mikko O. J. (2009), "Virtual Communities: A Bibliometric Analysis," *Proceedings of the 42nd Hawaii International Conference on System Sciences - 2009*, (accessed November 16, 2011), [available at http://csdl2.computer.org/comp/proceedings/hicss/2009/3450/00/07-08-03.pdf].

Lambert, Diane (1992), "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing," *Technometrics*, 34 (February), 1-14.

Lampe, Cliff, Rick Wash, Alcides Velasquez, and Elif Ozkaya (2010), "Motivations to participate in online communities," *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, New York, USA: ACM Press.

Lantz, Garold and Sandra Loeb (1998), "An Examination of the Community Identity and Purchase Preferences Using the Social Identity Approach," in *Advances in Consumer Research*, Volume 25, Joseph W. Alba and J. Wesley Hutchinson, eds. Provo, UT : Association for Consumer Research, 486-491.

Laumann, Edward O. (1966), *Prestige and Association in an Urban Community*. Indianapolis: Bobbs-Merrill.

Lazarsfeld Paul F. and Robert K. Merton (1954), “Friendship as a social process: a substantive and methodological analysis,” in *Freedom and Control in Modern Society*, M. Berger, ed. New York: Van Nostrand, 18-66.

Lewis, Michael (2006), “Customer Acquisition Promotions and Customer Asset Value,” *Journal of Marketing Research*, 43 (May), 195-203.

Li, Shaomin (1995), “Survival Analysis,” *Marketing Research*, 7 (4), 16-23.

Lin, Hsiu-Fen (2006), “Understanding Behavioral Intention to Participation in Virtual Communities,” *CyberPsychology & Behavior*, 9 (5), 540-547.

Lin, Hsiu-Fen (2007), “The role of online and offline features in sustaining virtual communities: an empirical study,” *Internet Research*, 17 (2), 119-138.

Lin, Hsiu-Fen and Gwo-Guang Lee (2006), “Determinants of success for online communities: an empirical study,” *Behaviour & Information Technology*, 25 (6), 479-488.

Linda L. Price, Lawrence F. Feick (1984), “The Role of Interpersonal Sources in External Search: An Informational Perspective”, in *Advances in Consumer Research*, Volume 11, Thomas C. Kinnear, ed. Provo, UT : Association for Consumer Research, 250-255.

Ling, Kimberly, Gerard Beenen, Pamela Ludford, Xiaoqing Wang, Klarissa Chang, Xin Li, Dan Cosley, Dan Frankowski, Loren Terveen, Al Mamunur Rashid, Paul Resnick, and Robert Kraut (2005), “Using Social Psychology to Motivate Contributions to Online Communities,” *Journal of Computer-Mediated Communication*, 10 (4), (accessessed May 24, 2011), [available at http://jcmc.indiana.edu.eaccess.ub.tum.de/vol10/issue4/ling.html].

Liu, WenSui and Jimmy Cela (2008), “Count Data Models in SAS,” *Proceedings SAS Global Forum 2008, paper 371-2008,* (accessed June 14, 2011), [available at http://www2.sas.com/proceedings/forum2008/371-2008.pdf].

Liu, Yong (2006), “Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue,” *Journal of Marketing*, 70 (July), 74–89.

Ma, Meng and Ritu Agarwal (2007), “Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities,” *Information Systems Research*, 18 (March), 42-67.

Madupu, Vivek and Delonia O. Cooley (2010), “Antecedents and Consequences of Online Brand Community Participation: A Conceptual Framework,” *Journal of Internet Commerce*, 9 (2), 127-147.

Markus, M. Lynne (1987), “Towards a ‘Critical Mass’ Theory of Interactive Media,” *Communication Research*, 14 (5), 491-511.

Markus, M. Lynne (1994), “Finding a happy medium: explaining the negative effects of electronic communication on social life at work,” *ACM Transactions on Information Systems*, 12 (April), 119-149.

Mathwick, Charla, Caroline Wiertz, and Ko de Ruyter (2008), "Social Capital Production in a Virtual P3 Community," *Journal of Consumer Research*, 34 (April), 832-849.

Maxham, James G. III and Richard G. Netemeyer (2002), "A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts," *Journal of Marketing*, 66 (October), 57-71.

Mayer, Adalbert (2009), "Online social networks in economics," *Decision Support Systems*, 47 (3), 169-184.

McAlexander, James H., John W. Schouten, and Harold F. Koenig (2002), "Building Brand Community," *Journal of Marketing*, 66 (January), 38-54.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, 27, 415-444.

Meffert, Heribert (2000), *Marketing – Grundlagen marktorientierter Unternehmensführung – Konzepte-Instrumente-Praxisbeispiele*. Wiesbaden: Gabler.

Mehra, Ajay, Martin Kilduff, and Daniel J. Brass (2001), "The Social Networks of High and Low Self-Monitors: Implications for Workplace Performance," *Administrative Science Quarterly*, 46 (March), 121-146.

Meyer, John P. and Natalie J. Allen (1991), "A three-component conceptualization of organizational commitment," *Human Resource Management Review*, 1 (1), 61-89.

Moe, Wendy W. and Michael Trusov (2011), "The Value of Social Dynamics in Online Product Ratings Forums," *Journal of Marketing Research*, 48 (June), 444-456.

Molm, Linda D. (2006), "The Social Exchange Framework," in *Contemporary Social Psychological Theories*, Peter J. Burke, ed. Standford, CA: University Press, 24-45.

Moorman, Christine, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314-328.

Moreno, Jacob L. (1934), *Who Shall Survive? - Foundations of Sociometry, Group Psychotherapy, and Sociodrama*. Washington, DC: Nervous and Mental Disease Publishing Company. Reprinted 1953 (second edition) by Beacon House Inc. Beacon, NY.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment-Trust Theory of Relationship Marketing", *Journal of Marketing*, 58 (July), 20-38.

Morris, Merrill and Christine Ogan (1996), "The Internet as Mass Medium," *Journal of Computer-Mediated Communication*, 1 (4), (accessed November 7, 2011), [available at http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1996.tb00174.x/full].

Muniz, Albert M. Jr. and Thomas C. O'Guinn (2001), "Brand Community," *The Journal of Consumer Research*, 27 (March), 412-432.

Murray, Keith (1991), “A Test of Services Marketing Theory: Consumer Information Acquisition Activities,” *Journal of Marketing*, 55 (January), 10-25.

Nahapiet, Janine and Sumantra Ghoshal (1998), “Social Capital, Intellectual Capital, and the Organizational Advantage,” *The Academy of Management Review*, 23 (April), 242-266.

Nambisan, Satish and Robert A. Baron (2007), “Interactions in virtual customer environments: Implications for product support and customer relationship management,” *Journal of Interactive Marketing*, 21 (Spring), 42-62.

Ng’andu, Nicholas H. (1997), “An Empirical Comparison of Statistical Tests for Assessing the Proportional Hazards Assumption of Cox’s Model,” *Statistics in Medicine*, 16 (6), 611-626.

Nitzan, Irit and Barak Libai (2011), “Social Effects on Customer Retention,” *Journal of Marketing*, forthcoming, (accessed August 22, 2011), [available at http://www.marketingpower.com/AboutAMA/Documents/JM_Forthcoming/social_effects_on_customer_retention.pdf].

Nonnecke, Blair and Jenny Preece (2000), “Lurker demographics: Counting the silent,” in *CHI '00 Proceedings of the SIGCHI conference on Human factors in computing systems*, New York: ACM, 73-80.

Nonnecke, Blair and Jenny Preece (2001), “Why lurkers lurk,” *Americas Conference on Information Systems 2001*, (accessed September 12, 2011), [available at http://www.socs.uoguelph.ca/~nonnecke/research/whylurk.pdf].

Nov, Oded and Chen Ye (2008), “Community Photo Sharing: Motivational and Structural Antecedents,” *ICIS Proceedings, Twenty ninth Intenational Conference on Information Sysems Paris 2008*, Association for Information Systems, Paper 91.

Nunnally, Jum C. (1978), *Psychometric Theory*. New York: McGraw-Hill.

Oliver, Richard L. (1997), *Satisfaction: A Behavioral Perspective on the Cons*umer. Boston, MA: Richard D. Irwin/McGraw-Hill.

Pajuniemi, Janne (2009), “The Revolution of Brand Marketing: The Era of Virtual Consumer Communities,” in *Proceedings of the Conference ‘Future of the Consumer Society*, Marileena Koskela and Markus Vinnari, eds. 28-29 May 2009, Tampere, Finland, 71-80.

Palmatier, Robert W., Rajiv P. Dant, Dhruv Grewal, and Kenneth R. Evans (2006), “Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis,” *Journal of Marketing*, 70 (October), 136-153.

Papacharissi, Zizi and Alan M. Rubin (2000), “Predictors of Internet Use”, *Journal of Broadcasting & Electronic Media*, 44 (2), 175-196.

Parasuraman, A. (1997), “Reflections on Gaining Competitive Advantage Through Customer Value,” *Journal of the Academy of Marketing Science*, 25 (2), 154-61.

Park, Namsu, Kerk F. Kee, and Sebastián Valenzuela (2009), “Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes,” *CyberPsychology & Behavior*, 12 (December), 729-33.

Parks, Malcolm R. and Kory Floyd (1996), “Making Friends in Cyberspace,” *Journal of Communication*, 46 (Winter), 80-97.

Parks, Malcolm R. and Lynne D. Roberts (1998), “`Making Moosic’: The Development of Personal Relationships on Line and a Comparison to their Off-Line Counterparts,” *Journal of Social and Personal Relationships*, 15 (August), 517-537.

Pelling, Emma L. and Katherine M. White (2009), “The Theory of Planned Behavior Applied to Young People’s Use of Social Networking Web Sites,” *CyberPsychology & Behavior*, 12 (6), 755-759.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003), “Common method biases in behavioral research: A critical review of the literature and recommended remedies,” *Journal of Applied Psychology*, 88 (October), 879-903.

Porter, Constance Elise (2004), “A Typology of Virtual Communities: A Multi-Disciplinary Foundation for Future Research,” *Journal of Computer-Mediated Communication*, 10 (1), Article 3, (accessed November 7, 2011), [available at http://jcmc.indiana.edu/vol10/issue1/porter.html].

Preacher, Kristopher J. and Andrew F. Hayes (2004), “SPSS and SAS Procedures for Estimating Indirect Effects in Simple Mediation Models,” *Behavior Research Methods Instruments & Computers*, 36 (4), 717-731.

Preacher, Kristopher J. and Andrew F. Hayes (2008), “Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models,” *Behavior Research Methods*, 40 (3), 879-891.

Preece, Jenny (2000), *Online Communities - Designing Usability, Supporting Sociability*. Chichester: John Wiley & Sons.

Preece, Jenny (2001), “Sociability and usability in online communities: Determining and measuring success,” *Behavior and Information Technology*, 20 (5), 347-356.

Putnam, Robert D. (1995), “Bowling Alone: America’s Declining Social Capital,” *Journal of Democracy,* 6 (1), 65-78.

Putter, Hein, M. Sasako, H. H. Hartgrink, C. J. H. van de Velde, and J. C. van Houwelingen (2005), “Long-term survival with non-proportional hazards: results from the Dutch Gastric Cancer Trial,” *Statistics in Medicine,* 24 (18), 2807-2821.

Reeder, Heidi M. (2000), “'I Like you... as a Friend': The Role of Attraction in Cross-Sex Friendship,” *Journal of Social and Personal Relationships*, 17 (3), 329-348.

Reichheld, Frederick F., Sasser, W. Earl (1990), “Zero Defections: Quality Comes to Services”, *Harvard Business Review*, 68 (5), 105-111.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17-35.

Reinartz, Werner J. and V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77-99.

Reinartz, Werner, Jacquelyn S. Thomas, and V. Kumar (2005), „Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63-79.

Reingen, Peter H. and Jerome B. Kernan (1986), "Analysis of Referral Networks in Marketing: Methods and Illustration," *Journal of Marketing Research*, 23 (November), 370-379.

Reynolds, Kristy E. and Sharon E. Beatty (1999), "Customer benefits and company consequences of customer-salesperson relationships in retailing," *Journal of Retailing*, 75 (Spring), 11-32.

Rheingold, Howard (1993), *The Virtual Community*. Reading, MA: Addison-Wesley.

Rice, Roneld E. (1993), "Media Appropriateness - Using Social Presence Theory to Compare Traditional and New Organizational Media," *Human Communication Research*, 19 (4), 451-484

Ridings, Catherine M. and David Gefen (2004), "Virtual Community Attraction: Why People Hang Out Online," *Journal of Computer-Mediated Communication*, 10 (November), (accessed March 22, 2011), [available at http://jcmc.indiana.edu/vol10/issue1/ridings_gefen.html].

Ridings, Catherine M., David Gefen, and Bay Anrize (2006), "Psychological Barriers: Lurker and Poster Motivation and Behavior in Online Communities," *Communications of the Association for Information Systems*, 18, 329-354.

Ridings, Catherine M., David Gefen, and Bay Arinze (2002), "Some antecedents and effects of trust in virtual communities," *The Journal of Strategic Information Systems*, 11 (December), 271-295.

Rosenberg, Larry J. and John A. Czepiel (1984) "A Marketing Approach For Customer Retention," *Journal of Consumer Marketing*, 1 (2), 45-51.

Rothaermel, Frank T. and Stephen Sugiyama (2001), "Virtual internet communities and commercial success: individual and community-level theory grounded in the a typical case of TimeZone.com," *Journal of Management*, 27 (June), 297-312.

Ruef, Martin, Howard E. Aldrich, and Nancy M. Carter (2003), "The Structure of Founding Teams: Homophily, Strong Ties and Isolation Among U.S. Entrepreneurs," *American Sociological Review*, 68 (2), 195-222.

Ryu, Gangseog and Lawrence Feick (2007), "A Penny for Your Thoughts: Referral Reward Programs and Referral Likelihood," *Journal of Marketing*, 71 (January), 84-94.

Sabidussi, Gert (1966), “The Centrality index of a Graph,” *Psychemetrika*, 31 (4), 581-603.

Sangwan, Sunanda (2005), “Virtual Community Success: A Uses and Gratifications Perspective,” *Proceedings of the 38th Annual Hawaii International Conference on System Sciences – HICSS ‘05*, 193c.

SAS Institute Inc. (2008), *SAS/STAT® 9.2 User’s Guide*. Cary, NC: SAS Institute Inc.

Schemper, Michael (1992), “Cox Analysis of Survival Data with Non-Proportional Hazard Functions,” *Journal of the Royal Statistical Society: Series D (The Statistician),* 41 (4), 455-465.

Schmitt, Philipp, Bernd Skiera, and Christophe Van den Bulte (2011), “Referral Programs and Customer Value,” *Journal of Marketing*, 75 (January), 46-59.

Schneller, Johannes (2011), “Das Zusammenspiel von Offline und Online-Werbemedien: Print - Online - Mobile – Apps,” Institut für Demoskopie Allensbach, Presentation on October 6, 2011 in Hamburg, (accessed December 18, 2011), [available at http://www.acta-online.de/praesentationen/acta_2011/acta_2011-offline_online.pdf].

Schoenfeld, David (1982), “Partial residuals for the proportional hazards regression model,” *Biometrika*, 69 (1), 239-241.

Scott, John (2000), *Social Network Analysis - A Handbook*. London: Sage Publications, second edition.

Shang, Rong-An, Yu-Chen Chen, and Hsueh-Jung Liao (2006), “The Value of Participation in Virtual Consumer Communities on Brand Loyalty”, *Internet Research*, 16 (4), 398-418.

Shankara, Venkatesh, Glen L. Urban, and Fareena Sultan (2002), “Online trust: a stakeholder perspective, concepts, implications, and future directions,” *Journal of Strategic Information Systems*, 11 (3-4), 325-344.

Shen, Kathy Ning and Mohamed Khalifa (2008), “Exploring Multidimensional Conceptualization of Social Presence in the Context of Online Communities,” *International Journal of Human-Computer Interaction*, 24 (7), 722-748.

Short, John, Ederyn Williams, and Bruce Christie (1976), *The Social Psychology of Telecommunications*. London: John Wiley & Sons.

Solomon, Michael R., Carol Surprenant, John A. Czepiel, and Evelyn G. Gutman (1985), “A Role Theory Perspective on Dyadic Interactions: The Service Encounter,” *Journal of Marketing,* 49 (Winter), 99-111.

Sonnenbichler, Andreas C. (2010), “A Community Membership Life Cycle Model,” *Presented at the International Network For Social Network Analysis (INSNA): Sunbelt Conference 2009,* San Diego, California, USA, (accessed October 20, 2011), [available at http://arxiv.org/PS_cache/arxiv/pdf/1006/1006.4271v1.pdf].

Sparrowe, Raymond T., Robert C. Liden, Sandy J. Wayne, and Maria L. Kraimer (2001), "Social Networks and the Performance of Individuals and Groups," *The Academy of Management Journal*, 44 (April), 316-325.

Sproull, Lee and Sara Kiesler (1986), "Reducing Social Context Cues: Electronic Mail in Organizational Communication," *Management Science,* 32 (November), 1492-1512.

Stephen, Andrew T. and Olivier Toubia (2010), "Deriving Value from Social Commerce Networks," *Journal of Marketing Research*, 47 (April), 215-228.

Szymanski, David M. (1988), "Determinants of Selling Effectiveness: The Importance of Declarative Knowledge to the Personal Selling Concept," *Journal of Marketing*, 52 (January), 64-77.

Szymanski, David M. and David H. Henard (2001), "Customer Satisfaction: A Meta-Analysis of the Empirical Evidence," *Journal of the Academy of Marketing Science*, 29 (Winter), 16-35.

Tajfel, Henri (1974), "Social Identity and Intergroup Behaviour," *Social Science Information*, 13 (2), 65-93.

Tajfel, Henri and John Turner (1979), "An Integrative Theory of Intergroup Conflict," in *The Social Psychology of Intergroup Relations*, W. G. Austin and S. Worchel, eds. Monterey: Brooks-Cole, 33-47.

Tarasi, Crina O., Ruth N. Bolton, Michael D. Hutt, and Beth A. Walker (2011), "Balancing Risk and Return in a Customer Portfolio," *Journal of Marketing*, 75 (May), 1-17.

Therneau, Terry M. (1996), "Extending the Cox Model," Technical Report Series - Section of Biostatistics, Technical Report Number 58, Mayo Clinic (November 1996), (accessed July 7, 2011), [available at http://mayoresearch.mayo.edu/mayo/research/biostat/upload/58.pdf].

Therneau, Terry M. and Patricia M. Grambsch (2000), *Modeling Survival Data: Extending the Cox Model*. New York: Springer.

Thomas, Jacquelyn S., Robert C. Blattberg, and Edward J. Fox (2004), "Recapturing Lost Customers," *Journal of Marketing Research*, 41 (Februar), 31-45.

Tiwana, Amrit and Ashley A. Bush (2005), "Continuance in expertise-sharing networks: a social perspective," *IEEE Transactions on Engineering Management*, 52 (February), 85-101.

Toral, Sergio L., M. Rocio Martínez-Torres, and Federico J. Barrero (2009), "Virtual communities as a resource for the development of OSS projects: the case of Linux ports to embedded processors," *Behaviour & Information Technology*, 28 (5), 405-419.

Toral, Sergio L., M. Rocío Martínez-Torres, Federico J. Barrero, and Francisco Cortés (2009), "An empirical study of the driving forces behind online communities," *Internet Research*, 19 (4), 378-392.

Trusov, Michael, Anand V. Bodapati, and Randolph E. Bucklin (2010), "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research*, 47 (August), 643-658.

Trusov, Michael, Randolph E. Bucklin, and Koen Pauwels (2009), "Effects of Word-of-Mouth versus Traditional Marketing: Findings from an Internet Social Networking Site," *Journal of Marketing*, 73 (September), 90-102.

Tsai, Wenpin and Sumantra Ghoshal (1998), "Social Capital and Value Creation: The Role of Intrafirm Networks," *The Academy of Management Journal*, 41 (August), 464-476.

Valente, Thomas W. (1996), "Social network thresholds in the diffusion of innovations," *Social Networks*, 18 (1), 69-89.

Van den Poel, Dirk and Bart Larivière (2004), "Customer attrition analysis for financial services using proportional hazard models," *European Journal of Operational Research*, 157 (1), 196-217.

Verhoef, Peter C. and Bas Donkers (2005), "The effect of acquisition channels on customer loyalty and cross-buying," *Journal of Interactive Marketing*, 19 (Spring), 31-43.

Villanueva, Julian, Shijin Yoo, and Dominique M. Hanssens (2008), "The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth," *Journal of Marketing Research*, 45 (February), 48-59.

Vuong, Quang H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57 (March), 307-333.

Walker, Michael E., Stanley Wasserman, and Barry Wellman (1993), "Statistical Models for Social Support Networks," *Sociological Methods & Research*, 22 (1), 71-98.

Walther, Joseph B. (1992), "Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective," *Communication Research*, 19 (1), 52-90.

Wang, Youcheng and Daniel R. Fesenmaier (2004a), "Towards understanding members' general participation in and active contribution to an online travel community," *Tourism Managmeent*, 25 (6), 709-722.

Wang, Youcheng and Daniel R. Fesenmaier (2004b), "Modeling Participation in an Online Travel Community," *Journal of Travel Research*, 42 (February), 261-270.

Wangenheim, Florian and Tomás Bayón (2004a), "Satisfaction, Loyalty and Word of Mouth within the Customer Base of a Utility Provider: Differences between Stayers, Switchers and Referral Switchers," *Journal of Consumer Behaviour*, 3 (March), 211-220.

Wangenheim, Florian and Tomás Bayón (2004b), "The effect of word of mouth on services switching: Measurement and moderating variables," *European Journal of Marketing*, 38 (9/10), 1173-1185.

Wangenheim, Florian and Tomás Bayón (2007), "The chain from customer satisfaction via word-of-mouth referrals to new customer acquisition," *Journal of the Academy of Marketing Science*, 35 (2), 233-249.

Wasko, Molly McLure and Samer Faraj (2000), "It Is What One Does: Why People Participate and Help Others in Electronic Communities of Practice," *Journal of Strategic Information Systems*, 9 (September), 155-173.

Wasko, Molly McLure and Samer Faraj (2005), "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice," *MIS Quarterly*, 29 (March), 35-57.

Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*. Cambridge: University Press.

Webster, Frederick E. Jr. (1968), "Interpersonal Communication and Salesman Effectiveness," *Journal of Marketing*, 32 (July), 7-13.

Weiber, Rolf and Jörg Meyer (2002), "Virtual Communities," in *Handbuch Electronic Business*, Rolf Weiber, ed. Wiesbaden: Gabler, second edition, 343-361.

Weiss, Allen M., Nicholas H. Lurie, and Deborah J. Macinnis (2008), "Listening to Strangers:Whose Responses Are Valuable, How Valuable Are They, and Why?," *Journal of Marketing Research*, 45 (4), 425-436.

Weitz, Barton A. and Kevin D. Bradford (1999), „Personal Selling and Sales Management: A Relationship Marketing Perspective," *Journal of the Academy of Marketing Science*, 27 (Spring), 241-254.

Wellman, Barry (2001), "Physical Place and Cyberplace: The Rise of Personalized Networking," *International Journal of Urban and Regional Research*, 25 (June), 227-252.

Wellman, Barry and Keith Hampton (1999), "Living Networked On and Offline," *Contemporary Sociology*, 28 (November), 648-654.

Wellman, Barry, Anabel Quan Haase, James Witte, and Keith Hampton (2001), "Does the Internet Increase, Decrease, or Supplement Social Capital? Social Networks, Participation, and Community Commitment," *American Behavioral Scientist*, 45 (November), 436-455.

Wellman, Barry, Janet Salaff, Dimitrina Dimitrova, Laura Garton, Milena Gulia, and Caroline Haythornthwaite (1996), "Computer Networks as Social Networks: Collaborative Work, Telework, and Virtual Community," *Annual Review of Sociology*, 22, 213–38.

Wiertz, Caroline and Ko de Ruyter (2007), "Beyond the Call of Duty: Why Customers Contribute to Firm-hosted Commercial Online Communities," *Organization Studies*, 28 (3), 347-376.

Wilson, Kathryn, Stephanie Fornasier, and Katherine M. White (2010), "Psychological Predictors of Young Adults' Use of Social Networking Sites," *Cyberpsychology, Behavior, and Social Networking*, 13 (2), 173-177.

Winkelmann, Rainer (2003), *Econometric Analysis of Count Data*. Berlin - Heidelberg: Springer, fourth ed.

Woisetschlaeger, David M., Vivian Hartleb, and Markus Blut (2008), "How to Make Brand Communities Work: Antecedents and Consequences of Consumer Participation," *Journal of Relationship Marketing*, 7 (October), 237-256.

Wu, Joshua B., Peter W. Hom, Lois E. Tetrick, Lynn M. Shore, Liangding Jia, Chaoping Li, and Lynda Jiwen Song (2006), "The Norm of Reciprocity: Scale Development and Validation in the Chinese Context," *Management and Organization Review*, 2 (November), 377-402.

Wübben, Markus and Florian Wangenheim (2008), "Instant Customer Base Analysis: Managerial Heuristics Often 'Get It Right'," *Journal of Marketing*, 72 (May), 82-93.

Xie, Zexian, Huaping Chen, and Zhenhua Wu (2008), "An Empirical Study of the Antecedents and Consequences of the Ordinary Member's Commitment in Virtual Communities," *International Conference on Information Management, Innovation Management and Industrial Engineering*, 3, 329-334.

Xu, Bo, Donald R. Jones, and Bingjia Shao (2009), "Volunteers' involvement in online community based software development," *Information & Management*, 46 (3), 151-158.

Yi, Youjae (1993), "Contextual Priming Effects in Print Advertisement: The Moderating Role of Prior Knowledge," *Journal of Advertising*, 22 (March), 1-10.

Yoo, Won-Seok, Kil-Soo Suh, and Moon-Bong Lee (2002), "Exploring the Factors Enhancing Member Participation in Virtual Communities," *Journal of Global Information Management*, 10 (July-September), 55-71.

Yu, Tai-Kuei, Long-Chuan Lu, and Tsai-Feng Liu (2010), "Exploring factors that influence knowledge sharing behavior via weblogs," *Computers in Human Behavior*, 26 (1), 32-41.

Zhang, Yixiang, Yulin Fang, Kwok-Kee Wei, and Huaping Chen (2010), "Exploring the role of psychological safety in promoting the intention to continue sharing knowledge in virtual communities," *International Journal of Information Management*, 30 (October), 425-436.

Zhang, Zhongju (2010), "Feeling the Sense of Community in Social Networking Usage," *IEEE Transactions on Engineering Management*, 57 (2), 225-239.