The Pennsylvania State University

The Graduate School

ESSAYS ON STAKEHOLDER NETWORKS IN MARKETING

A Dissertation in

Business Administration

by

Franziska Schmid

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2022

The dissertation of Franziska Schmid was reviewed and approved by the following:

J. Andrew Petersen
Associate Professor of Marketing
Dissertation Advisor
Chair of Committee

Arvind Rangaswamy
University Distinguished Professor of Marketing

Stefan Wuyts
Professor of Marketing

Bruce Desmarais
DeGrandis-McCourtney Early Career Professor of Political Science

Brent Ambrose
Jason and Julie Borrelli Faculty Chair in Real Estate
Director of Smeal College of Business Ph.D. Program

ABSTRACT

This dissertation examines the importance of how stakeholder networks can generate insights for marketers. Although there has been ample research on some types of stakeholder networks in the marketing literature (e.g., customers), not all stakeholder networks have been given equal attention. In the introduction of this dissertation, I discuss the different types of stakeholder networks, including networks of customers, employees, and suppliers, as well as networks of firms' followers on social media. The following essays address gaps in the marketing literature by examining how data on two specific networks of stakeholders (salespeople and social media followers) can aid firms in identifying potentially valuable marketing strategies.

The first essay uses novel data to identify network ties between salespeople at a business-to-business (B2B) firm. Using this data on network ties, I measure salespeople's social capital within the firm. I then distinguish how the effect of social capital on B2B sales is distinct from other categories of intellectual capital (i.e., human and organizational capital). I find that categories of intellectual capital have heterogenous impacts on salespeople's success across different selling types (e.g., customer acquisitions, rebuying, and cross-selling). Among selling types, human capital is most beneficial for rebuying and social capital for cross-selling.

In the second essay, I demonstrate how to use standard network analysis methods and community detection algorithms common in computer science to identify segments of followers using a large organization's social media account. Many firms only apply a mass marketing strategy without targeting content to specific follower segments. I use four community detection methods to identify segments of social media followers with common connections or interests. My results show that among all methods tested, the algorithm that utilizes both network connections and follower interests as inputs to detect "communities" generates segments with the most specific community characteristics. These results suggest that firms can leverage their existing social network data to detect communities based on follower connections and interests, enabling more granular marketing strategies.

I conclude by discussing future directions for stakeholder network research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my advisor J. Andrew Petersen for guiding me through the Ph.D. program. I am thankful for the opportunity to learn from him and for his support every step of the way. I could not have completed this work without his valuable advice. I would also like to thank Arvind Rangaswamy, Stefan Wuyts, and Bruce Desmarais for serving on my committee, dedicating time to reviewing my work, and giving me invaluable feedback. Special thanks to my co-author Rajkumar Venkatesan for his feedback and support.

I would also like to thank all the Ph.D. students I have met at Penn State. Without their continuous support, I would not have been able to make it through my doctoral program. I want to thank Christine Mihelcic, Maxence Valentine, and Gretchen Ross for their friendship and advice over the last few years.

I also thank my parents for their unconditional love and support in anything I pursue, even though it means living thousands of miles away. I could not have accomplished this milestone without them believing in me. Finally, I am thankful for my partner, Brian Gibbons, who has provided me with support and advice throughout my time at Penn State.

Chapter 1

INTRODUCTION

1.1 Stakeholder Networks

Much of the research in marketing on social networks has focused on analyzing consumer networks. This pattern is logical because the customer is at the core of business environments, and data is readily available. Recently, however, research has shifted to focusing on other stakeholder networks (e.g., employees, suppliers) to understand better how all internal and external parties to the firm can affect marketing outcomes. The management and marketing literature has long recognized the importance of stakeholders (Freeman 2010). However, the academic analysis has lacked focus on certain stakeholder network groups (Petersen and Schmid 2021), even though social network connections among all stakeholders can provide valuable knowledge into how stakeholder groups function in relation to the firm (Borgatti et al. 2009).

In this dissertation, I analyze stakeholder networks, focusing on how firms can utilize information about these networks to improve marketing strategies. To introduce this dissertation, I discuss past research on stakeholder networks below and identify important gaps in the marketing literature. I then summarize how this dissertation addresses these gaps.

As previously stated, the marketing literature has devoted the most attention to customer networks. Research has found that analyzing customer networks provides valuable insights into the diffusion process of customer opinions (Godes and Mayzlin 2004; Goldenberg, Libai, and Muller 2001; Katona, Zubcsek, and Sarvary 2011; Van den Bulte and Lilien 2001), the adoption process of products and opinions (Hill, Provost, and Volinsky 2006; Iyengar, van den Bulte, and Lee 2015; Toker-Yildiz et al. 2017; Trusov, Rand, and Joshi 2013), as well as innovation creation (Stephen, Zubcsek, and Goldenberg 2016; Tsai 2001). These studies have demonstrated the practicality of examining networks in the marketing context, which has led others to examine how other stakeholder networks affect marketing outcomes.

Past research has also focused on employee networks. Most of the research on employee networks has been on the effect of salesperson and sales manager networks on sales performance. More specifically, research has found that employee networks are beneficial for identifying central employees who can be gatekeepers of information (Ahearne et al. 2013; Bolander et al. 2015; Gonzalez, Claro, and Palmatier 2014; Wang, Gupta, and Grewal 2017).

There has been less research in the marketing literature on the value of supplier networks. However, early research has shown how firms can use supplier networks to gather information for new products and services (Dyer 1996). Supplier networks can also aid in forming marketing alliances and contribute to firm value (Swaminathan and Moorman 2009). These studies again confirm the value of utilizing network analysis to improve marketing strategy.

Aside from some stakeholder networks receiving more attention than others, there are also still noteworthy gaps in network analysis of stakeholders in the marketing literature. These gaps primarily center around implementing different data collection methods and empirical analysis. In the marketing literature, most network research utilizes surveys to have network members self-identify connections. While survey data can provide valuable information about perceived ties in networks, it might not be feasible to frequently survey employees, suppliers, shareholders, and other communities to map all network ties in a business setting. Other disciplines, for example, political science and sociology, utilize secondary data for bipartite network analysis (Everett 2016; Everett et al. 2018; Grassi et al. 2019; Koskinen and Edling 2012). By using publicly available data or firm data, firms can analyze stakeholder networks without incurring additional costs. In the first essay, I address this gap by utilizing existing data from a business-to-business (B2B) firm's CRM system to analyze network ties between salespeople. The benefit of this type of data is that it exists at every B2B firm, and one can easily observe the network evolution over time.

Additionally, past research on customer networks has often focused on analyzing individual actors in online social networks (Borgatti et al. 2009) rather than identifying communities within networks. Shifting the focus from the individual to the group structure (e.g., network communities) can offer different insights regarding shared interests among members. Focusing on group structures also allows firms to target more specific segments of stakeholders.

Moreover, most network research in marketing analyzes the full network (i.e., all network ties within set boundaries). One approach not commonly utilized in the marketing literature is the ego network approach (Stolz and Schlereth 2021). Instead of focusing on an individual node or a full network, the ego network encompasses all the direct ties of a focal node and how the ego's direct ties are connected. This approach has the advantage that data collection is more feasible than collecting data on all possible ties. In the second essay, I analyze communities of social media followers. This type of analysis has not been the focus of many studies in marketing. More specifically, I use ego network data and focus on groups of actors by applying different community detection approaches.

1.2 Overview of Two Essays

In the first essay, I analyze an undirected B2B salesperson network (i.e., employee network). Even though past research has studied employee networks before, this essay provides new insights into what value employee networks can add by utilizing available firm data. I examine how different sources of knowledge, human, social, and organizational capital, affect B2B selling types (i.e., acquisition and retention). I find that different sources of knowledge make salespeople more likely to focus on certain selling types. Human capital (i.e., experience) is helpful for B2B rebuying, and social capital (i.e., betweenness centrality) fosters cross-selling.

In the second essay, I uncover a community structure in a firm's ego network on social media. I propose a framework to identify network communities based on social media users' network connections and their interests. The four methods in the framework consist of social network measures and community detection algorithms. With each additional method, I relax conditions and overcome the shortcomings of the previous methods. For example, forcing users into communities without allowing them to be members of multiple communities ignores weak network ties. Only relying on network connections ignores that users can share similar interests without being connected. Therefore, utilizing network connections combined with topic modeling offers a segmentation tool with new insights on how to best segment online social media followers.

Both essays highlight the importance of including stakeholder network analysis in the value generation process for firms.

Chapter 2

Essay 1: The Value of Intellectual Capital for Business-To-Business Selling

2.1 Introduction

B2B firms often sell complex products requiring in-depth technical knowledge (Schmitz, Lee, and Lilien 2014). The structure of a sales organization supporting the sale of such products is often organized by the product or service categories a firm offers (Cespedes 2014). As a result, B2B salespeople are responsible for selling products from one category to multiple customers. This selling process requires salespeople to have expert knowledge in their assigned category. Although the process differs from account-based selling (i.e., key-account management), where key-account managers sell products across categories, salespeople can still customize product solutions for their customers within their specific category.

Besides the defined product or service categories a B2B salesperson sells at a firm, a salesperson must also decide which sales opportunity seems the most valuable. More specifically, salespeople need to tradeoff their limited time between customer acquisition and customer retention, with the latter consisting of rebuying and cross-selling. This tradeoff holds different risks based on the information available to salespeople. Customer acquisition (i.e., selling to new customers who have never purchased from the firm) tends to have a lower probability of success because a salesperson must use external information to research potential buyers. Additionally, pursuing a customer acquisition can be risky because the buyer has no ties to the firm yet.

A less risky strategy for salespeople is to sell to existing customers through rebuying (i.e., buying a product that they have already purchased in the past) or cross-selling (i.e., selling a product from a category that is new to an existing customer) (Schmitz, Lee, and Lilien 2014). Rebuying tends to be a successful selling type because a customer has already bought the product from the firm before, and a salesperson has gathered customer knowledge through experience. However, the value of rebuy opportunities is usually smaller than other opportunities.

Cross-selling can be lucrative because it provides the potential for additional rebuys (Kamakura 2008). Another benefit of cross-selling is that successful cross-selling strengthens the relationship between a buyer and the firm as the buyer now purchases multiple products from the firm (Reinartz and Kumar 2003). However, cross-selling is similar to acquisition in that a customer is new to that particular category and salesperson. Even though salespeople know how to target their existing customers, they often lack the knowledge to target customers outside their category. For cross-selling, salespeople need customer knowledge about customers they have not sold to before. Thus, salespeople in other categories who have experience with these customers can provide helpful information.

So how and when do salespeople choose to pursue different selling types? To better understand this question, I examine the knowledge salespeople can access at a firm. Firms build their competitive advantage through the knowledge available at the firm. This accumulation of knowledge has been described as intellectual capital, consisting of three categories. These categories are human capital, social capital, and organizational capital (Mariadoss et al. 2014; Subramaniam and Youndt 2005; Youndt, Subramaniam, and Snell 2004).

Accessing human capital means drawing from a salesperson's knowledge acquired from experience. Human capital can aid selling because salespeople who have gathered experience with a B2B firm are more familiar with the social journey and processes at that firm (Grewal and Sridhar 2021). Accessing social capital means drawing from knowledge through connections with other salespeople. Past research has suggested that the sales force is the best resource to collect information about customers and the selling process because salespeople interact with customers daily (Hughes, Le Bon, and Rapp 2013). Therefore, a salesperson managing another product category can provide helpful account information to identify other opportunities.

The last intellectual capital, organizational capital, refers to the knowledge that salespeople can access through the firm, for example, through a customer relationship management (CRM) system. This category of intellectual capital is knowledge external to the salesperson, whereas human and social capital are internal to the salesperson.

Existing selling-related knowledge allows salespeople to pursue sales opportunities (Verbeke, Dietz, and Verwaal 2011). However, salespeople can obtain various knowledge, and different opportunities may require specific types of knowledge. Thus, I examine different sources of knowledge necessary to succeed at B2B selling by drawing from salespeople's intellectual capital. More specifically, I want to answer the following research questions about the drivers of balancing B2B selling types and sales outcomes:

1. What factors explain which selling type salespeople pursue in B2B markets?
2. What makes sales opportunities in B2B markets more likely to be successful?
3. What is the trade-off between different types of intellectual capital?

My findings show that salespeople with higher human capital are more likely to choose rebuy opportunities than acquisitions and cross-selling. Additionally, salespeople with higher social capital are more likely to focus on cross-selling than rebuys and acquisitions. Among other things, I also find that social and human capital help generate revenue for salespeople.

The rest of the essay is structured in the following way. I first discuss the background literature in B2B acquisition and retention and selling-related knowledge to arrive at the theoretical foundations. Then, I develop the conceptual model and measure the identified characteristics to build the empirical model. To test the conceptual model empirically, I use data from a B2B paper-and-packaging firm that captures salespeople's sales opportunities, including successful and unsuccessful sales. Last, I discuss implications for theory and practice as well as limitations and opportunities for future research.

2.2 Background Literature

In this essay, I draw from two streams of literature. One stream is the B2B acquisition and retention literature. The second stream of research examines past findings on salespeople's knowledge and how salespeople utilize it. The following sections discuss past findings in both areas and explain how I differentiate this work and add to the literature.

2.2.1 B2B Acquisition and Retention

In CRM, the allocation of resources is an important strategic decision around balancing customer acquisition and retention (Min et al. 2016). Researchers have examined the balance between customer acquisition and retention by examining the necessary marketing-mix investments for the highest customer profitability (Reinartz, Thomas, and Kumar 2005), the competitive environment (Musalem and Joshi 2009), and among others, the effect on innovation (Arnold, Fang, and Palmatier 2011). Salespeople who focus on customer retention focus on the selling types rebuying or cross-selling. Rebuy opportunities are usually quick wins for salespeople because they repeat a past transaction. Cross-selling, however, can provide value for firms as a CRM strategy to improve customer lifetime value (CLV) by expanding the number of categories a customer is purchasing (Reinartz and Kumar 2003). The more products a customer purchases from a firm, the higher the switching costs and the stronger the firm and customer account relationship (Kamakura et al. 2003). Cross-selling can lead to higher profitability for the firm and more efficient purchasing processes (Kamakura 2008; Kamakura et al. 2003; Malms and Schmitz 2011).

Research has often used past customer characteristics to identify selling opportunities with customers (Kamakura, Kossar, and Wedel 2004; Kumar, Venkatesan, and Reinartz 2008). For example, a manager would estimate the likelihood of a customer's purchase based on past purchasing decisions, no matter the type of selling. However, the salesperson's decision on what selling type to pursue also impacts the success of a sale. Additionally, salesperson characteristics influence the likelihood of success. Thus, other researchers have studied the following characteristics of selling success: adaptive selling behavior and customer orientation (Franke and Park 2006). Both approaches require much information about the customer and the selling process to adapt the message for customer interactions and find solutions for customer problems.

Moreover, customer orientation can be a successful strategy for salespeople to strengthen customer relationships (Homburg, Müller, and Klarmann 2011). To follow this customer focus, salespeople need the knowledge to successfully offer the right solutions for customer needs. Thus, salespeople need to know about a firm's product offering and understand customer needs (Rapp et al. 2006).

Among the B2B selling types, B2B firms have incentive structures to align salespeople's choices with the firm strategy (Bommaraju and Hohenberg 2018; Coughlan and Kissan 2012; Kishore et al. 2013). Salespeople are often incentivized to sell with commissions and bonuses based on their sales outcomes (Kishore et al. 2013; Lim, Ahearne, and Ham 2009). In this research, I solely focus on the types of intellectual capital salespeople can access to pursue sales opportunities.

More specifically, this essay examines which drivers make salespeople more likely to focus on a particular selling type. Therefore, I explore what sources of knowledge (i.e., categories of intellectual capital) are needed to succeed at customer acquisition, rebuys, and cross-selling. Past research has identified that knowledge needs to be transferred to be useful but has not examined how knowledge availability affects the balance and success of different selling types (Maurer, Bartsch, and Ebers 2011; Tsai 2001). Thus, I add to the literature by showing which categories of intellectual capital impact the choice between customer acquisition, rebuys, and cross-selling. Next, I discuss the literature on sources of knowledge because I am interested in how salespeople acquire knowledge within a firm to build their human, social, and organizational capital to balance B2B selling types.

2.2.2 Sources of Knowledge

In the second stream of literature, I draw from the literature about salespeople's sources of knowledge (see Table 2-1). Salespeople need skills to perform necessary tasks and be successful. One thing that adds to a salesperson's skill set is selling-related knowledge. Verbeke, Dietz, and Verwaal (2011) describe this knowledge as the product, technical, and customer knowledge needed to solve customer problems in sales situations and communicate with customers. In their meta-analysis, the authors find that selling-related knowledge, degree of adaptiveness, role ambiguity, cognitive aptitude, and work engagement are significant drivers of sales performance. However, selling-related knowledge was the driver with the strongest effect. This finding suggests that in B2B selling, salespeople must possess adequate technical product knowledge and understand how to approach and communicate with customers.

Additionally, salespeople need to understand the sales process and how a product can provide customer solutions to be successful at selling. Within a firm, salespeople can draw selling-related knowledge from different sources. For example, salespeople can gather knowledge with each customer interaction to learn about customer needs (Verbeke, Dietz, and Verwaal 2011). Consequently, salespeople with more experience have had more opportunities to improve their knowledge, which they can apply to future sales situations (Matsuo and Kusumi 2002). (Franke and Park 2006) show that experience helps salespeople prepare for different selling situations.

Table 2-1: Sources of Knowledge

Study	Research Questions	Knowledge Source	Data	Outcome Variables	Results
Ahearne et al. (2008)	Impact of information technology on salesperson performance	CRM system	Pharmaceutical sales team	Sales performance	Use of CRM system has a positive effect on salesperson knowledge
Ahearne et al. (2013)	Leveraging network centrality to translate intra-functional competitive intelligence (CI) into performance; most important informal network for centrality; the importance of organizational position and the difference in the effect of centrality	Intrafirm network (Social Capital)	Survey data of sales team in B2B media firm	Sales performance; CI quality	Central sales managers are more effective at problem-solving concerning CI issues; sales managers networking with peers creates more impact than salesperson networking
Gupta et al. (2019)	Effect of intra- and interfirm networks on the account profitability of key accounts	Intra- and interfirm networks	Survey data of B2B firms with more than 500 employees in various industries with KAM teams	Seller key account profitability	Account profitability is enhanced by cross-functional ties in intra-firm networks and by similar function ties in interfirm networks; centralization improves profitability if combined with one of the other network measures in interfirm networks
Mariadoss et al. (2014)	Role of salesperson competitive intelligence behavior in salesperson performance	Salesperson product knowledge; CRM system	Survey data of sales representatives at medical device companies in the biotechnology industry	Salesperson competitive intelligence behaviors; Performance	Salespeople's product knowledge has a positive effect on competitive intelligence behavior and performance; high (low) sales force automation use weakens (strengthens) the effect of product knowledge on performance
Rapp et al. (2006)	Impact of empowering leader behavior and salesperson experience and knowledge on working hard and smart	Experience	Performance and customer ratings of pharmaceutical salespeople; Archival data of salespersons' efforts and performance	Working hard; Working smart; Performance; Customer Service; Customer Satisfaction	Empowering behavior from managers helps salespeople to work smarter, but not harder and is most useful for salespeople with little experience; Salespeople with more knowledge but less experience work harder
Üstüner and Iacobucci (2012)	Effectiveness of salespeople in identifying opportunities, creating solutions, and closing the sale based on their network embeddedness	Intrafirm network	High-tech electronic components market; survey data	Salespeople's performance measured as sales manager evaluation	Social networks are a driver of salesperson performance; social network ties drive opportunity identification; work networks drive solution creation; both networks help with closing a sale
This Research	Impact of intellectual capital on selling type (acquisition and retention) effort and sales success	Experience; Intrafirm network; CRM system	Historical data from a B2B paper-and packaging firm	Acquisition; Rebuy; Cross-sell; Revenue of won opportunities	Different categories of intellectual capital have different effects on B2B selling types; human capital makes salespeople more likely to focus on rebuying, whereas social capital makes salespeople more likely to focus on cross-selling

Other research finds that higher levels of knowledge from experience result in salespeople working more efficiently (Rapp et al. 2006). Thus, through experiences, salespeople can gather knowledge that helps with sales performance in different selling situations. Additionally, experience contributes to salespeople's human capital (Youndt, Subramaniam, and Snell 2004).

Another source of knowledge that salespeople can draw from is other salespeople at the firm (Hughes, Le Bon, and Rapp 2013). To understand salespeople's potential access to information, I can examine their network connections with other salespeople. Salespeople can interact with each other and build relationships at firm-held events (e.g., sales meetings, training sessions, customer meetings), but they also communicate with each other informally. During these informal conversations, salespeople exchange information about customers and best practices. However, firms do not have insights into this informal exchange, and it often stays unreported and unmeasured even though the connections provide insights into knowledge sharing.

Within a network, which Wasserman and Faust (1994) define as a set of individuals that have relations with one another, salespeople have an environment to share different resources and leverage knowledge through their social network ties. Past B2B network studies measure several types of network centrality to investigate the effect of a central network position on sales performance (Ahearne et al. 2013; Üstüner and Iacobucci 2012). Moreover, past research has identified that social networks can help with increasing sales performance (Bolander et al. 2015; Üstüner and Iacobucci 2012).

Other research analyzes intra- and interfirm networks to find an effect on key account profitability (Gupta et al. 2019). As the intrafirm network measures, the authors use network density, cross-functional ties, and in-degree centrality. They then analyze how intrafirm network measures interact with interfirm network measures to study the effects on a key account team's performance and not a single salesperson's performance. While their measures give insights into how information might be shared across firms and departments, they do not specify how salespeople share information with other salespeople.

Üstüner and Iacobucci (2012) analyze three stages in the overall B2B sales process: opportunity identification, solution creation, and closing a sale. However, they focus their analysis on customer acquisition and do not include customer retention in their research. Overall, there is only limited research studying the effect of social networks on specific performance outcomes besides overall sales performance or profitability.

Salespeople's third source of knowledge is organizational-level knowledge stored in CRM systems (Mariadoss et al. 2014). For example, they can retrieve customer information and past sales transactions from the system to help facilitate sales. Past research has examined whether and how the use of technology impacts sales performance (Ahearne et al. 2008). For instance, Ahearne et al. (2008) find that the use of CRM data positively impacts salespeople's knowledge. Other research finds a positive relationship between salesperson knowledge, competitive intelligence behavior, and performance (Mariadoss et al. 2014). However, sales force automation or the use of technology is only valuable to a certain extent. Higher use of technology weakens the positive effect of knowledge on sales performance (Mariadoss et al. 2014). Thus, more access to firm knowledge does not always result in improved performance. All salespeople can access the same knowledge through the CRM system, whereas experience and social network connections are internal to a salesperson and can vary.

In this research, I measure all three categories of intellectual capital as salespeople's knowledge sources. More specifically, because different sources of knowledge can be helpful for different selling, I study the impact of knowledge (i.e., intellectual capital) on selling choice, successful sales outcomes, and the revenue of successful sales. Most research has investigated sources of knowledge separately, whereas I include three common sources of knowledge. I use secondary data already available to firms within their sales software for the network analysis. In contrast, past studies have used survey data for their network analysis, which can be costly, time-intensive, and inaccurate due to self-report biases (Van den Bulte 2010). Finally, this research gives insight into how salespeople can share and utilize resources for different selling types.

2.3 Conceptual Model

I developed the conceptual model by drawing from the previously discussed literature findings (Figure 2-1). I examine B2B customer acquisition and retention as two-stage processes; first, salespeople choose an opportunity in customer acquisition, rebuying, or cross-selling. Second, I analyze whether the opportunity is won or lost.

In the first stage, I identify controls such as customer exchange characteristics (e.g., past sales experience with the customer) and salesperson exchange characteristics (e.g., past sales behavior for the salesperson). Additionally, I identify a salesperson's intellectual capital (i.e., human capital and social capital) as key drivers in choosing an opportunity with a customer. All salespeople can access the same organizational capital, which salespeople cannot influence; therefore, I use it as a control but do not expect it to be a main driver of choice.

In the second stage, I include controls and intellectual capital as key drivers of closing a sales opportunity. A salesperson's human capital moderates the effect of a salesperson's social capital on the success of an opportunity. I also expect that the effort of salespeople with a customer account (i.e., marketing contact strategy) will impact the opportunity's success (i.e., won or lost).

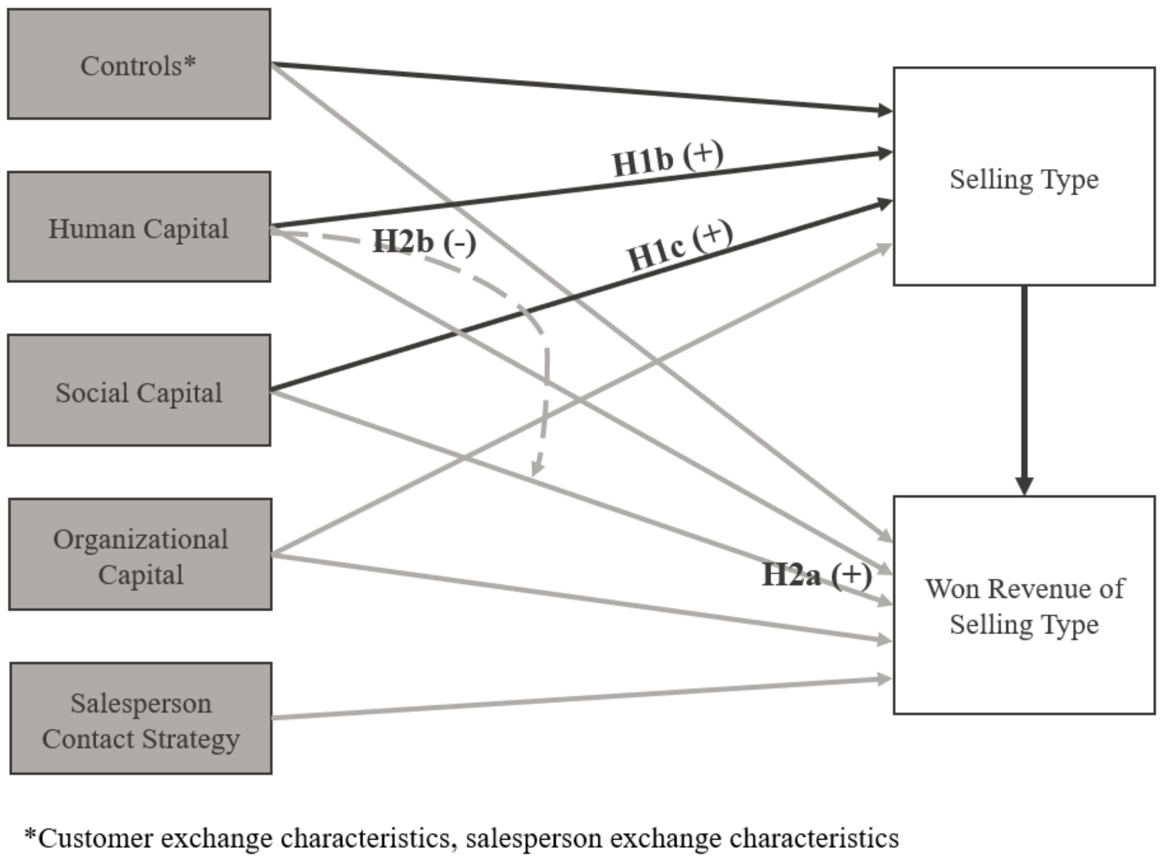


Figure 2-1: Conceptual Model

2.4 Hypotheses Development

Firms all hold knowledge, which is one main resource for conducting business. Moreover, a firm's knowledge adds to its competitive advantage and influences performance. For knowledge to be helpful, different parties in the firm must utilize it. Therefore, knowledge transfer is part of a firm's learning process and is important for its performance (Tsai 2001). Below I discuss the knowledge available through intellectual capital and build the hypotheses.

2.4.1 Intellectual Capital

Past research has described the sum of knowledge at a firm as intellectual capital (Nahapiet and Ghoshal 1998; Subramaniam and Youndt 2005; Youndt, Subramaniam, and Snell 2004). It includes individuals' knowledge, relationships, and knowledge stored in databases and processes (Youndt, Subramaniam, and Snell 2004). Because knowledge can be accumulated and utilized differently, researchers have identified three categories of intellectual capital: human capital, social capital, and organizational capital (Subramaniam and Youndt 2005).

Human capital consists of an employee's knowledge, experience, and skill set. Thus, the knowledge is internal to the individual and the individual can utilize it (Schultz 1961). Employees can strengthen their human capital with more job experience and interactions with other employees at the firm.

Social capital is the knowledge salespeople can access through network connections with other individuals. More specifically, the knowledge within a firm is a resource that employees can make available to others (Coleman 1988). Over the years, the social capital theory developed from social resource theory, and researchers have applied it in different social contexts (Borgatti, Everett, and Johnson 2018; Burt 1992; Lin 2001). Social resource theory states that a person's success partly consists of a function of social ties and the resources those ties give access to (Lin 2001). The three different social contexts of social capital are the structural, relational, and cognitive aspects (Tsai and Ghoshal 1998). The structural dimension addresses social interactions and the resulting social structures of those interactions within a network. Hence, actors can benefit from a network structure conducive to social interactions because they can access resources, for instance, knowledge in a B2B setting (Granovetter 1992; Hakansson and Snehota 1995).

Additionally, a salesperson connected to other salespeople who control resources is in an advantageous position to receive information through a possible inflow of information (Borgatti, Everett, and Johnson 2018). Inside firms, employees can share knowledge by passing down information in a formal structure from a sales manager to a subordinate or in an informal structure from peer to peer. When peers pass down information, such as peers in the same position or across functions, employees build relationships with each other, also called an informal network (Krackhardt and Hanson 1993).

The last category of intellectual capital is organizational capital, which is the knowledge within the firm's systems and processes (Youndt, Subramaniam, and Snell 2004). If the CRM data is well maintained, salespeople can access customer information and requests. However, there is evidence that salespeople often do not keep detailed records of the information they gathered (Hughes, Le Bon, and Rapp 2013; Kamakura et al. 2005). Additionally, defined processes are also part of a firm's knowledge.

In the past, research has often looked at the categories of intellectual capital independently (Madhavaram and Hunt 2017; Tsai and Ghoshal 1998; Youndt, Subramaniam, and Snell 2004). The findings can still be meaningful, but the different categories are often interdependent because they build onto each other. For instance, human capital resides within a salesperson and accumulates through experience at a firm. No one else has access to the salesperson's human capital unless the salesperson shares it with others. Thus, social capital depends on other salespeople's human capital. Moreover, salespeople with strong social capital can obtain others' knowledge (i.e., human capital). A firm can only store and expand its organizational capital when an employee diffuses human capital through social capital. In the following hypotheses, I test whether different categories of intellectual capital impact salespeople's selling type choices and their success with different selling types.

2.4.2 First Selling Stage: Selling Choice

I develop the hypotheses by drawing from the theoretical foundations of intellectual capital. Human and social capital are categories of intellectual capital unique to each salesperson, whereas organizational capital is accessible to all salespeople at the firm. Therefore, all salespeople have access to the same knowledge from organizational capital.

When salespeople engage in customer acquisition, the knowledge that will be most useful comes from external sources to the firm. If salespeople do not have high human or social capital, they are likely to target new customers. Thus, I hypothesize:

H_{1a}: *Salespeople without strong human or social capital at the firm are more likely to choose customer acquisition over other selling types.*

When salespeople engage in rebuying opportunities, they need knowledge specific to the existing customer. Because salespeople in rebuy situations have sold to the same customer before, they can draw from existing knowledge (i.e., human capital). Therefore, salespeople with high human capital are likely to choose rebuy opportunities over other selling types. Thus, I hypothesize:

H_{1b}: Salespeople with strong human capital at the firm are more likely to choose to rebuy over other selling types.

When salespeople engage in cross-selling, they need knowledge about customers who have bought from other categories. If salespeople have no experience with customers, they can rely on other salespeople who have sold to those customers before. Consequently, the intellectual capital most helpful for cross-selling is social capital because it gives access to knowledge through network ties. Thus, I hypothesize:

H_{1c}: Salespeople with strong social capital are more likely to choose to cross-sell over other selling types.

2.4.3 Second Selling Stage: Selling Performance

After a salesperson chooses a selling type, the next step is to close the sale successfully. To do so, salespeople need the right selling-related knowledge, which they can acquire from different categories of intellectual capital. Salespeople can draw from their human capital if they have had enough time at a firm to build their skill set through experience. If salespeople are new to a firm, they have not had time to build up enough human capital to close a sale without help from other salespeople. Examples of knowledge necessary to close a sale are information about the selling process or communicating with the customer's buying team. Additionally, high intellectual capital leads to stronger financial returns than low intellectual capital, regardless of the category (Youndt, Subramaniam, and Snell 2004). Thus, I hypothesize:

H_{2a}: A firm's strong human, social, or organizational capital increases sales revenue.

Rapp et al. (2006) findings suggest that salespeople with less experience but strong expertise (e.g., B2B sales requires technical product knowledge) can work harder to compensate for their lack of experience. Thus, if a salesperson does not have enough human capital through experience, relying on social capital from network ties may mitigate a lack of human capital. Based on the discussed theoretical foundations, I hypothesize the following:

H_{2b}: Strong (weak) human capital of a salesperson dampens (strengthens) the effect of social capital on sales revenue.

2.5 Empirical Application

2.5.1 Data

To empirically test how different categories of intellectual capital impact salespeople's choice of selling, I use data from a B2B paper-and-packaging firm based in the United States. The firm sells a variety of paper and packaging products (e.g., cardboard, paperboard) to customer accounts in nine different industries. These industries range from pharmaceuticals to consumer-packaged goods. For example, a customer account could be a candy manufacturer buying paper products as one category and cardboard as another. In the firm's sales department, salespeople are assigned to sell product categories in specific regions, which means that a salesperson is responsible for multiple accounts in one product category. If a customer account purchases from two categories, the customer account works with two salespeople.

In the dataset, I observe 77,251 total sales opportunities for 1,285 salespeople and 18,072 customer accounts in North America between June 2007 and February 2018. Additionally, I observe whether an opportunity was successful or not (i.e., won or lost) and the amount of each sales opportunity. I can also differentiate whether an opportunity is a customer acquisition, a cross-sell, or a rebuy. For instance, the first sale to a customer is the customer acquisition; any opportunity after the original sale could be a cross-sell or a rebuy opportunity. In addition to the selling data, I also use data from a salesperson network. Different from other B2B social network research, which mainly uses survey data, I use secondary data based on past sales opportunities as a source for measuring an informal social network among salespeople.

Table 2-2 provides descriptive statistics for the sample. 33% of customer acquisition opportunities are won, whereas 65% of rebuy opportunities are successful. These statistics show that salespeople are far more likely to convert rebuy versus acquisition opportunities into sales. However, only 28% of cross-selling opportunities are successful. This statistic confirms that most (in this data, 72%) of the cross-selling opportunities do not lead to a sale and shows that it is more difficult to win an opportunity if it is a cross-sell (Schmitz, Lee, and Lilien 2014). However, if a cross-selling opportunity is successful, the firm can generate high gains in revenue with a median value of $56,525 compared to the median rebuy value of $22,000.

Table 2-2: Winning Percentages for Sales Opportunities

(June 2007 - February 2018)

Variable	**Percentage Won**	**Median Amount Won (Lost)**
Sales Opportunities	55%	$24,650 ($50,000)
Acquisition Opportunities	33%	$35,000 ($50,000)
Rebuy Opportunities	65%	$22,000 ($50,000)
Cross-Selling Opportunities	28%	$56,525 ($100,000)

2.5.2 Variable Operationalization

Outcome variables. One of the dependent variables is the type of sales opportunity a salesperson pursues. I examine the different selling types: customer acquisition, cross-selling, and rebuying. Customer acquisition is operationalized as a binary variable based on whether a customer purchases from the firm for the first time. I operationalize a rebuy opportunity as a binary variable and measure it as a 1 when customers buy from a category they have already purchased.

Researchers have measured cross-selling in different ways, including cross-selling as a cumulative measure of product categories in B2C. Others have used B2B cross-selling performance measured by sales managers' perceptions through survey data (Kumar, George, and Pancras 2008; Malms and Schmitz 2011; Schmitz 2013; Schmitz, Lee, and Lilien 2014). Kumar, Venkatesan, and Reinartz (2008) measure a cross-buy with the number of past product types purchased by customers. I operationalize cross-selling as a binary variable indicating whether a salesperson sold from a category new to an existing customer. To do so, I identify customers' first purchases and compare subsequent purchases with previous purchase categories to accurately measure cross-selling. The other outcome variable is the won revenue amount. I measure this variable with the dollar amount of won sales revenue.

Customer exchange characteristics. Each customer account in the data set has a buying relationship with the B2B firm. Similar to previous research, I measure characteristics that reflect the exchange between a customer account and the firm due to the customer-firm relationship (Reinartz and Kumar 2003). Other research used exchange characteristics based on past purchases, purchasing frequency, and purchasing amounts (Hughes 1996; Reinartz and Kumar

2003; Rossi et al. 1996). For the exchange characteristics, I use the win proportion of salespeople with customers by dividing the number of won opportunities by the total of all opportunities. Subsequently, I include the proportion of a selling type by dividing the number of opportunities for one selling type (e.g., rebuy) by the number of all opportunities. The last customer exchange variable includes the win proportion of a selling type, which I operationalize by dividing the number of won opportunities for one selling type by all opportunities for the same selling type.

Salesperson exchange characteristics. I measure the average number of other salespeople's opportunities for a selling type per year in one category to control for the selling behavior of other salespeople. I operationalize it by taking the average of all opportunities for one selling type across other salespeople in one category per year. If 200 salespeople sell in one category, then the measure takes the average of 199 salespeople's opportunities within a year.

Human capital. I operationalize human capital with the experience of a salesperson (Youndt, Subramaniam, and Snell 2004). I measure it with the years a salesperson has worked for the firm (i.e., tenure time) because salespeople develop human capital based on experience. This measure is a proxy for human capital and indicates how much selling knowledge a salesperson has gathered over the years.

Social capital. To measure a salesperson's social capital, I take the informal network between salespeople at the firm, where I observe salesperson ties with each other, built on social interactions. In contrast, a formal network would be based on hierarchical structures in the firm. I took a two-mode data set that portrays which salesperson worked with which customer account (i.e., tie between salesperson and customer) and projected it into one-mode data (i.e., tie between salesperson and salesperson) (Borgatti, Everett, and Johnson 2018). The projection is based on whether salespeople have worked with the same customer accounts, which means they likely interacted at some point.

Additionally, I added weights to the network edges by measuring how close in time salespeople worked with the same customer accounts. By comparing the absolute minimum distance in time between salesperson assignments, I create weights for the network edges. I argue that salespeople are more likely to interact more frequently if they are responsible for a customer account closer in time. Thus, the network weights will give me a better idea of who had the potential to interact and exchange knowledge. As a result, the final network is an undirected valued social network. There are 1,285 salespeople (nodes) in the network with 17,440 edges. The average number of connections (degree) of salespeople is 13.5.

Past research has measured social capital in several ways because it gives access to various resources (Borgatti, Everett, and Johnson 2018; Maurer, Bartsch, and Ebers 2011; Tsai and Ghoshal 1998). Burt (1992) argues that structural holes give insight into actors with access to complementary information; however, his constraint measure only considers an actor's ego network, not the full network. Another measure that examines the flow of information is closeness centrality since it can be interpreted "as the minimum time until the arrival of something flowing through the network" (Borgatti, Everett, and Johnson 2018). Closeness centrality relies on redundant ties and helps conform to norms, yet it is not as helpful in spreading information across different network subgroups or isolates in the network (Centola and Macy 2007; Coleman 1988).

Betweenness centrality is one social capital operationalization that measures whether an actor is a gatekeeper of information and can allow or withhold information flow (Freeman 1977, 1979). Therefore, actors in positions that connect many other actors (i.e., high in betweenness centrality) have access to various novel information. Given that cross-selling requires novel information from teams across different product categories, salespeople with connections to multiple teams can control and access additional knowledge. Thus, I use betweenness centrality to operationalize social capital based on the informal salesperson network data.

(Freeman 1977) states that betweenness centrality measures the shortest distance (geodesic distance) between two other actors. One benefit of using betweenness centrality is that it can be calculated even when there are gaps in the network, which is impossible with other centrality measures (e.g., closeness centrality). I calculate betweenness centralization for the salesperson network with Freeman's betweenness centralization equation (Wasserman and Faust 1994):

(1) $$C_B = \frac{\sum_{i=1}^{g}[C'_B(n^*) - C'_B(n_i)]}{(g-1)}$$

where C_B is the betweenness centralization index, n refers to a node, $C'_B(n^*)$ is the maximum actor betweenness score for the set of actors, and g is the number of actors. The index can take on a value between 0 and 1. To calculate the betweennesses centrality for each salesperson at a given opportunity, I measure the weight of an edge in the following way:

(2) $$\text{Weight} = \frac{1}{|t_i - t_k|}$$

where *t* is the date when salesperson *i* and *k* worked for the same account. I used the igraph package in R for all the network calculations.

Organizational capital. As the third category of intellectual capital, I measure organizational capital as the amount of information stored in the CRM system. More specifically, I count the number of past opportunities starting with the first date in the system (2007) at the time of an opportunity to measure how much information salespeople can access during a sale.

Marketing contact strategy. To control for the number of times a salesperson contacts a customer account, I also include the salesperson contact strategy as a variable in this analysis. I measure this through the number of marketing interactions, including events, email correspondence, and phone calls a salesperson has with customer accounts for the sales cycle of a given opportunity.

Heterogeneity. Within the model, I control for the customer (firm) and salesperson heterogeneity. I also control for a customer's industry, which I observe within the data set. Additionally, I use salesperson clustered standard errors. I provide a summary of the variables and their operationalizations in Table 2-3 and descriptive statistics for the variables in Table 2-4. The correlations in Table 2-4 only represent model-free evidence. To determine the drivers of selling type choice and the success of sales opportunities, I require a more formal modeling framework to control for other factors.

Table 2-3: Variable Operationalization

Variable Name	**Operationalization**
Outcome Variables	
Acquisition	A dummy variable equal to 1 if it is the first purchase by a customer.
Cross-Sell	A dummy variable equal to 1 if an opportunity is in a new category for an existing customer account.
Rebuy	A dummy variable equal to 1 if it is a purchase by an existing customer in the same category as already purchased.
Amount Won	The revenue amount of a won sales opportunity.
Intellectual Capital	
Social Capital	Measured with betweenness centrality as the number of times an actor lies on the geodesic distance between two other actors.
Human Capital	Measured in tenure time with the number of years a salesperson has been working for the firm.
Organizational Capital	The number of opportunities stored in the CRM system at a given point in time.
Customer Exchange Characteristics	
Customer Win Proportion	The number of won opportunities with a customer account divided by the number of opportunities with that account.
Proportion of Selling Type	The number of opportunities for one type of selling (acquisition, cross-sell, rebuy) divided by the number of opportunities for all selling types.
Win Proportion of Selling Type	The number of won opportunities for one type of selling (acquisition, cross-ell, rebuy) divided by the number of all opportunities for the same selling type.
Salesperson Exchange Characteristics	
Average Selling Type Opportunities per Year and Category	The average of opportunities for one selling type across all salespeople in one category per year.
Salesperson Contact Strategy	The number of marketing interactions between the customer and the salesperson for a given opportunity.

Table 2-4: Descriptive Statistics (mean, standard deviation, correlations)

	Variable	μ	σ	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Acquisition	0.29	0.46	1													
2	Cross-sell	0.04	0.19	-0.13	1												
3	Rebuy	0.67	0.47	-0.92	-0.28	1											
4	Amount Won (log)	5.46	5.29	-0.26	-0.09	0.28	1										
5	Betweenness Centrality	0.37	0.31	-0.02	0.01	0.01	0.00	1									
6	Tenure Time	2.12	2.43	-0.13	-0.01	0.13	0.01	-0.04	1								
7	Cust. Winning Prop	0.59	0.26	-0.29	-0.12	0.33	0.45	-0.00	0.04	1							
8	Win Prop. of Selling Type	0.67	0.26	-0.43	-0.20	0.49	0.53	0.00	0.06	0.89	1						
9	Prop. of Selling Type	0.60	0.29	-0.11	-0.33	0.24	0.15	-0.01	-0.10	0.26	0.26	1					
10	Avg. Selling Type Opp. per Year and Category	0.57	0.21	-0.31	-0.47	0.49	0.21	-0.01	0.02	0.36	0.42	0.61	1				
11	Avg. Mktg. Contacts per Year per Category	0.22	0.41	0.03	0.00	-0.03	0.00	0.07	-0.02	-0.03	-0.02	-0.06	-0.15	1			
12	CRM Information	44378	25621	0.02	0.03	-0.04	-0.14	-0.04	0.36	-0.26	-0.23	-0.15	-0.17	-0.07	1		
13	Avg. Amount Won (log)	9.96	2.84	-0.22	-0.08	0.24	0.18	0.01	0.03	0.31	0.27	-0.13	0.07	0.35	0.04	1	
14	Marketing Contacts (log)	-1.00	2.52	-0.01	-0.00	0.02	0.02	-0.04	-0.02	0.06	0.04	0.09	0.13	0.22	0.02	-0.03	1

2.5.3 Methodological Development

In B2B sales salespeople can choose between three selling types. These types are customer acquisition, rebuying, and cross-selling. As a result, I need a model that can identify the drivers of different selling choices. Thus, I model the selling type choice using a multinomial probit regression, which models whether a salesperson chooses to pursue an acquisition, a rebuy, or cross-selling at a given opportunity. The dependent variable is the choice of selling type with three unordered categories (acquisition, cross-sell, rebuy). The main drivers include betweenness centrality, tenure time, and customer and salesperson exchange characteristics. In addition, I include the share per selling type of other salespeople in the same category to control for the actions of other salespeople.

Once a salesperson chooses a specific selling type (i.e., acquisition, rebuying, or cross-selling), the next step is closing the sale and generating revenue. Therefore, a salesperson's selling performance depends on the opportunity type and revenue generated from that closed opportunity. Consequently, I need a model structure that can examine the drivers that lead to a successful sale and help generate revenue. Sales revenue as the dependent variable is left truncated because generated revenue is either greater than 0 if successful or 0 if a sale is not successful. I model the success of an opportunity with a type I Tobit model to account for the left-censored data. The drivers included in the model are human capital (i.e., tenure time), social capital (i.e., betweenness centrality), customer and salesperson exchange characteristics, and the number of marketing contacts with a customer. I also include a salesperson random effect to account for unobserved salesperson heterogeneity.

In this model structure, I face the following empirical challenges. First, I need to deal with the endogeneity of the selling type choice. The decision that salespeople make when choosing a selling type depends on the other selling types and which selling types are available at the time. When using probit models, past research cautions against including predicted values of endogenous regressors (Danaher et al. 2015). To address the endogeneity of the selling type, I use "probit residuals" as covariates in the revenue model, where I use the residuals of the selling type model (Che, Chen, and Yuxin 2012; Danaher et al. 2015; Petrin and Train 2010). I save the residuals of the selling type model and then use them in the revenue model.

Second, I need to deal with the endogeneity of the marketing effort of salespeople. To correct for endogeneity, I use a control function approach (Che, Chen, and Yuxin 2012; Petrin and Train 2010). I first estimate a salesperson's marketing effort (i.e., how many times did a salesperson contact a customer account) with other previously used covariates as independent variables. I use the marketing efforts of other salespeople as an instrumental variable by measuring how many marketing touchpoints other salespeople have within a sales cycle. I use it as an instrument because the marketing efforts of other salespeople should not influence the sales performance of a salesperson. Still, they may influence the marketing actions of the salesperson. I then use the residuals of the marketing effort model in the revenue model.

2.5.4 Model Development

Selling Type Model. The selling type choice is a multinomial probit model modeled with the following equation:

(3) $$\mathrm{Selling\ Type}_{ijt} = \alpha_0 + \alpha_{HC}\mathrm{HC}_{it}^{ST} + \alpha_{SC}\mathrm{SC}_{it}^{ST} + \alpha_X \mathrm{X}_{ijt}^{ST} + \alpha_C \mathrm{Controls}_{ijt}^{ST} + \varepsilon_{ijt}^{ST}$$

where $Selling\ Type_{ijt}$ stands for what type of selling type a salesperson goes after (cross-selling or rebuy compared to acquisition), HC_{it}^{ST} is the human capital measure for salesperson *i*, SC_{it}^{ST} is the social capital measure (betweenness centrality index measure) for salesperson *i*, X_{ijt}^{ST} represents the exchange characteristics for salespeople and customer accounts, including an interaction term between human and social capital, $Controls_{ijt}^{ST}$ include the industry control variables and organizational capital, ε_{ijt}^{ST} is the error term, and *i*, *j*, and *t* represent the salesperson, customer account, and time.

Marketing Contact Model. I model the amount a salesperson contacts a customer during a sales cycle by using a regression with the following equations:

$$(4)\ \text{Marketing Contacts}_{\text{ijt}} = \alpha_0 + \alpha_{\text{HC}}\text{HC}_{\text{it}}^{\text{MC}} + \alpha_{\text{SC}}\text{SC}_{\text{it}}^{MC} + \alpha_{\text{X}}\text{X}_{\text{ijt}}^{\text{MC}} + \alpha_{\text{C}}\text{Controls}_{\text{ijt}}^{\text{MC}} + \alpha_Z\text{Z}_{\text{it}} + \varepsilon_{\text{ijt}}^{\text{MC}}$$

where $\text{Marketing Contacts}_{\text{ijt}}$ is the number of times a salesperson contacted a customer (i.e., marketing effort), $\text{HC}_{\text{it}}^{\text{MC}}$ is the human capital measure for salesperson *i*, SC_{it}^{MC} is the social capital measure (betweenness centrality index measure) for salesperson *i*, $\text{X}_{\text{ijt}}^{\text{MC}}$ represents the exchange characteristics for salespeople and customer accounts, $Controls_{ijt}^{MC}$ include the industry control variables and the organizational capital, Z_{it} is the instrumental variable, $\varepsilon_{\text{ijt}}^{\text{MC}}$ is the error term, and i, j, and t represent the salesperson, customer account, and time.

Revenue Model. I model the success of an opportunity as the amount of a won opportunity with a random-effects Tobit regression with the following equation:

$$(5)\ \text{Revenue}_{\text{ijt}}^{*} = \gamma_0 + \gamma_{\text{HC}}\text{HC}_{\text{it}}^{\text{R}} + \gamma_{\text{SC}}\text{SC}_{\text{it}}^{\text{R}} + \gamma_{\text{x}}\text{X}_{\text{ijt}}^{\text{R}} + \gamma_{\text{C}}\text{Controls}_{\text{ijt}}^{\text{R}} + \gamma_{pr}\text{pr}_{\text{ijt}} + \gamma_{mc}\text{mc}_{\text{ijt}} + \text{u}_{\text{i}}^{\text{R}} + \varepsilon_{\text{ijt}}^{\text{R}}$$

$$\text{Revenue}_{\text{ijt}} = \begin{cases} \text{Revenue}_{\text{ijt}}^{*} & \text{if Revenue}_{\text{ijt}}^{*} > 0 \\ 0 & \text{if Revenue}_{\text{ijt}}^{*} \le 0 \end{cases}$$

where $Revenue_{ijt}$ is the amount of a closed opportunity for salesperson *i* with account *j* in time *t*, $\mathrm{HC}_{\mathrm{it}}^{\mathrm{R}}$ is the human capital measure for salesperson *i*, SC_{it}^{R} is the social capital measure (betweenness centrality index measure) for salesperson i, X_{ijt}^{R} represents the exchange characteristics for salespeople and customer accounts, including an interaction term between human and social capital, $Controls_{ijt}^{R}$ include the industry control variables and the organizational capital, pr_{ijt} are the probit residuals from the selling type model, mc_{ijt} is the marketing contact model computed residual, u_{i}^{R} is the salesperson random effect, ε_{ijt}^{R} is the error term.

2.5.5 Results

Selling Type: Cross-sell. I estimate the selling type model to identify drivers of salesperson choice of the various selling types with customer acquisition as the base category. For cross-selling choice, all the variables in the model except for human capital and the interaction effect between social and human capital are significant (see Table 2-5). These results suggest that the selected drivers help identify what makes salespeople more likely to choose cross-selling. As expected, I find that a higher betweenness centrality (i.e., social capital) increases the likelihood of choosing a cross-selling opportunity over customer acquisition (β = .411; $p < .05$). Additionally, a higher human capital, measured in tenure time, is not significant, and the coefficient is negative (β = -.058; $p > .10$).

Table 2-5: Estimation Results of Selling Type Model

Coefficient (Standard Error)	**Selling Type Model: Cross-Sell**	**Selling Type Model: Rebuy**
Acquisition (base outcome)		
Intercept	3.281*** (0.451)	-2.671*** (.249)
Social Capital		
Betweenness Centrality	.411** (.169)	.236*** (.053)
Human Capital		
Tenure Time	-.058 (.049)	.093*** (.030)
Customer Exchange Characteristics		
Customer Win Proportion	-3.079*** (.431)	-5.675*** (.547)
Salesperson Exchange Characteristics		
Win Proportion of Selling Type	2.085*** (.402)	7.679*** (.558)
Proportion of Selling Type	-1.822*** (.335)	-.801*** (.299)
Average Selling Type Opportunities per Year and Category	-21.060*** (2.127)	4.160*** (.491)
Organizational Capital		
CRM Information	.00002** (.000006)	.000003 (.000002)
Interaction Effect		
Tenure Time*Betweenness	-.077 (0.048)	-.019 (.017)
Controls		
Industry Fixed Effect	Included	Included
Model Fit		
Log-likelihood	-31877.325	
BIC	64092.29	

Notes: Standard errors are clustered at the salesperson level and reported in parentheses.
Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$

A salesperson is not more likely to choose cross-selling based on their past success across all selling types. Instead, it decreases the likelihood of choosing cross-selling ($\beta = -3.079$; $p < .01$). However, the success of the specific selling type, in this case, cross-selling, makes a salesperson more likely to choose cross-selling again ($\beta = 2.085$; $p < .01$). The past cross-sell proportion for a salesperson ($\beta = -1.822$; $p < .01$) decreases the likelihood to choose cross-selling. Moreover, the number of cross-selling opportunities other salespeople go after in a category per year decreases the likelihood of choosing to cross-sell ($\beta = -21.060$; $p < .01$). The interaction effect of tenure time (i.e., human capital) and betweenness centrality (i.e., social capital) is not significant and negative (β =-.077; *p > .10*). Lastly, the effect of organizational capital is positive and significant with a small coefficient (β = .00002; *p < .01*).

Selling Type: Rebuy. For rebuy choice, I find that the variables in the model, except for the interaction effect of human and social capital, are significant (see Table 2-5). Here, I also find that a higher betweenness centrality (i.e., social capital) increases the likelihood of choosing a rebuy opportunity over customer acquisition (β = .236; *p < .01*). When comparing the effect of betweenness centrality on cross-selling and rebuying, I find that salespeople with a higher betweenness centrality are most likely to cross-sell before rebuying and acquisition. For choosing rebuys, the tenure time (i.e., human capital) is significant and increases the likelihood to rebuy (β = .093; *p < .01*).

A salesperson is less likely to choose to rebuy based on their past success across all selling types (β = -5.675; *p < .01*). Whether a salesperson has had success with rebuys makes them more likely to choose rebuys (β =7.679; *p < .01*). The more a salesperson rebuys, the lower the likelihood of choosing to rebuy again (β = -.801; *p < .01*). Opposite to cross-selling, the number of rebuy opportunities other salespeople go after in a category per year increases the likelihood of a salesperson choosing rebuying (β = 4.160; *p < .01*). The interaction effect of betweenness centrality and tenure time is not significant for rebuying (β = -.019; *p > .10*). Lastly, organizational capital is positive and significant with a small coefficient (β = .000003; *p < .01)*

Table 2-6: Estimation Results of Marketing Contact Model

Coefficient (Standard Error)	**Marketing Contacts** (without IV)	**Marketing Contacts** (with IV)
Intercept	-.006 (.008)	-.076*** (.008)
Social Capital		
Betweenness Centrality	.069*** (.006)	.053*** (.006)
Human Capital		
Tenure Time	.004*** (.001)	.004*** (.001)
Salesperson Exchange Characteristics		
ln(Avg. Amount Won)	-.001** (.0005)	-.002*** (.001)
Win Proportion of Selling Type	.15*** (.006)	0.156*** (.005)
Proportion of Selling Type	.021*** (.006)	.048 *** (.006)
Organizational Capital		
CRM Information	-.0000007*** (.00000007)	-.0000005*** (.00000006)
Interaction Effect		
Tenure Time*Betweenness	.009*** (.002)	.006*** (.015)
Instrumental Variable		
Avg. Marketing Effort per Year and Category	-	.226*** (.003)
Controls		
Salesperson Random Effect	Included	Included
Time Fixed Effect	Included	Included
Industry Fixed Effect	Included	Included
Model Fit		
Adj. R^2	0.0232	.0801

Significance: *** p < 0.01; ** p < 0.05; * p < 0.10

Marketing Contact Model. The results from the marketing contact model show that all variables are significant (see Table 2-6). The average marketing effort per year that other salespeople in a category put forth (IV) is positive and significant and shows it is a good fitting instrument (β = .226; $p < .01$). Comparing the regression results without the instrumental variable, I see that the F-value increases from 142.05 to 481.17 when adding the instrumental variable. I also find that the two categories of intellectual capital internal to a salesperson, social capital (β = .053; $p < .01$) and human capital (β = .004; $p < .01$), are positive and significant. Additionally, the size of a successful opportunity is negative and significant (β = -.002; $p < .01$). How many opportunities were of the same selling type (i.e., the proportion of selling type) is positive and significant (β = .058; $p < .01$). The interaction effect of social and human capital is positive and significant (β = .006; $p < .01$). Lastly, the organizational capital is positive and significant (β = .000; $p < .01$).

Table 2-7: Estimation Results of Revenue Model

Coefficient (Standard Error)	**Revenue Model**
Intercept	-8.65*** (.333)
Social Capital	
Betweenness Centrality	.203* (.12)
Human Capital	
Tenure Time	.646*** (.039)
Salesperson Exchange Characteristics	
Cross-selling	.947*** (.222
Rebuying	.289*** (.089)
Proportion of Selling Type	.865*** (.194)
Proportion of Won Selling Type	20.482*** (.199)
ln(Average Amount Won)	.066*** (.016)
Firm-Initiated Marketing Efforts	
ln(Marketing Contacts)	-3.031*** (.389)
Organizational Capital	
CRM Information	-.0001*** (.000003)
Interaction Effect	
Tenure Time*Betweenness	-.028 (.046)
Instrumental Variables	
Marketing Contact Residuals	4.090*** (.89)
Cross-Sell Model Errors	.152 (.141)
Rebuy Model Errors	-4.982*** (.889)
Control Variables	
Industry Fixed Effect	Included
Model Fit	
Log-likelihood	-162724.54
BIC	325696.7

Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$

Revenue Model. I estimate the revenue model and find that all the variables in the model, except for the interaction effect of human and social capital, are significant (see Table 2-7). I find that social capital positively affects the likelihood of winning an opportunity and generating a higher revenue amount (β = .203; $p < .10$). Human capital also increases the likelihood to be successful at B2B selling (β = .646; $p < .01$). Although both social and human capital are drivers of success, the interaction effect of tenure time and betweenness centrality is not significant (β = -.028; $p > .10$). The more organizational capital a salesperson has, the less likely the success of a sale (β = -.0001; $p < .01$). If an opportunity is a cross-sell, the likelihood to generate positive revenue increases (β = .947; $p < .01$). If an opportunity is a rebuy opportunity, the likelihood of winning and generating revenue increases as well (β = .289; $p < .01$). The proportion of opportunities for the same selling type is positive and significant (β = .865; $p < .01$). Past wins with the same selling type as the opportunity at hand increase the likelihood of generating revenue (β = 20.482; $p < .01$). The number of times a salesperson contacts a customer per sales cycle decreases the likelihood of succeeding at generating revenue (β = -3.031; $p < .01$).

2.5.6 Discussion

The positive results of the intellectual capital categories internal to salespeople provide empirical support for the hypotheses (see Table 2-8). When comparing all three selling types, salespeople with a higher social capital are more likely to cross-sell over rebuys but more likely to put effort into rebuys over spending time on acquisitions (H1a). The order in which salespeople go after selling types changes for salespeople with higher human capital. With increased experience, salespeople are more likely to pursue rebuying than acquisitions and cross-selling because rebuying requires knowledge about a salesperson's existing customers (H1b). In addition, a salesperson with high social capital (i.e., betweenness centrality) is in an advantageous position to receive information from salespeople in other categories. Therefore, social capital is most helpful in choosing cross-selling opportunities and provides the ability to access knowledge about customers new to the salesperson to make a sale (H1c).

Table 2-8: Supported Hypotheses

Hypotheses		Supported
H_{1a}	*Salespeople without strong human or social capital at the firm are more likely to choose customer acquisition over other selling types.*	✓
H_{1b}	*Salespeople with strong human capital at the firm are more likely to choose to rebuy over other selling types.*	✓
H_{1c}	*Salespeople with strong social capital are more likely to choose to cross-sell over other selling types.*	✓
H_{2a}	*Strong human, social, or organizational capital at a firm increases sales revenue.*	✓
H_{2b}	*Strong (weak) human capital of a salesperson dampens (strengthens) the effect of social capital on sales revenue.*	

However, the more successful a salesperson is overall, the more the likelihood to cross-sell goes down compared to acquisitions; and the likelihood to pursue rebuys goes down as well. An explanation could be that salespeople with a high win proportion have already depleted their existing options to sell and, thus, try to acquire new customers. Suppose salespeople have been successful with a specific selling type in the past. In that case, they are more likely to pursue that same selling type because they have generated knowledge on how to succeed at the selling task. Salespeople need to succeed with cross-selling to pursue cross-selling opportunities in the future.

The more other salespeople cross-sell in a particular category, the less likely a salesperson is to do it in that category. It seems possible that there are not as many opportunities left if a customer account already buys multiple products from the B2B firm. Therefore, the sales team has exhausted the cross-selling options. If other salespeople do more rebuying, then a salesperson is still more likely to focus on rebuying over acquisitions because every salesperson sells to their existing customers. Thus, opportunities will not be exhausted because customers need to replenish products over time.

When examining the number of times a salesperson contacts a customer during a sales cycle, I find that people with higher social capital reach out to customers less, suggesting that building and maintaining a network can save time and costs. Additionally, salespeople with higher human capital also reach out to customers less because they likely already have strong knowledge about their customers. When other salespeople in the same category increase their marketing contacts, a salesperson in that same category will contact their customers less.

Regarding revenue success, salespeople with higher social capital are more likely to close a sale successfully, as well as salespeople with higher human capital. Both effects show that salespeople with higher internal intellectual capital know what works to close a sale because they have acquired enough experience or can rely on their network connections to access necessary information for sales. Acquiring more organizational capital does not increase the likelihood of winning a sale and generating revenue. This result aligns with past research that has found that if salespeople start relying on stored CRM data, they rely less on their other skills and intuition (Mariadoss et al. 2014). Therefore, not all intellectual capital automatically helps with generating revenue (H2a). To succeed at selling, salespeople must build their own human and social capital. The interaction effect of social and human capital would indicate that they might have a substitutable relationship; however, the effect is not significant (H2b).

2.5.7 Network Formation

I define the network connections based on whether a salesperson worked on the same account as another. Based on my knowledge of the B2B firm, salespeople are assigned to customer accounts by the region the salesperson works in and the product category a salesperson sells. Therefore, there should not be any salesperson tie selection based on the performance, selling effort, or other characteristics of a salesperson.

To test for endogenous network formation in the network data, I run a Stochastic Actor-Oriented Model (SAOM) with the SIENA (Simulation Investigation for Empirical Network Analysis) package in R. This model can model the changes in network ties in combination with behavior changes. Thus, it analyzes the co-evolution of the network and actor behavior simultaneously by applying method-of-moments estimations based on repeated computer simulations of networks, which then produces parameter effects with standard errors (Ripley et al. 2021; Snijders, Steglich, and Schweinberger 2007; Steglich, Snijders, and Pearson 2010).

The model compares different time periods (i.e., waves) of the salespeople's network connections behaviors. The first wave of observations is used as starting parameters in the network tie and behavior simulation. As outcome variables for the model, the changes in the network and the behavior are both defined. To model the frequency of opportunities where actors can decide about a tie change, the SAOM models rate functions for each wave. These opportunities for change are micro steps. Table 2-9 gives an overview of the components of the SAOM model as described.

Table 2-9: Schematic Overview of Model Components

	Occurrence	**Rule of Change**
Network changes	Network rate function	Network objective function
Behavioral changes	Behavioral rate function	Behavioral objective function

(Steglich, Snijders, and Pearson 2010)

Several effects could impact changes in network ties. One effect could be structural network endogeneity based on reciprocity, transitivity, or structural holes (Snijders, Steglich, and Schweinberger 2007). The other effect could come from behaviors or characteristics of network actors, which can also be determinants of network change. Sharing characteristics or behaviors can be signs of homophily, preference for similarity, or a node's attractiveness for a relationship.

As the dependent variables, I use the network of salespeople and the selling behavior of salespeople in the form of rebuying and cross-selling. The data set is unique because it is an undirected network, and the network dependent variable only increases. For this specific case, when a dependent variable only increases, some of the basic effects of the SAOM are not identified. These effects are the outdegree (i.e., the basic tendency toward tie formation) for the network variable and the linear shape effect for the behavior variable. Some effects considering in- and out-degrees are not a good fit for this model because the network is undirected.

For this undirected network, I employ model type 2, called the *forcing model,* assuming that if one of two actors initiates a tie, a tie is automatically created (Ripley et al. 2021). I measured three time periods of B2B selling at the firm, from 2011 to 2013. Each wave in time is one year long because B2B sales cycles tend to be longer; thus, shorter waves would not present enough opportunities for salespeople to create a tie.

Table 2-10 shows the results of the SIENA model. In general, positive effect sizes show a tendency for salespeople to create new network ties. The rate function models the number of opportunities actors have to change their network ties (i.e., form a tie). The results show the network function (i.e., network dynamics) and the behavior functions for each behavior (i.e., behavior dynamics).

Table 2-10: SAOM Estimation Results

	Log Odds	**SE**
Network Dynamics		
Rate Effect Period 1	0.233***	0.019
Rate Effect Period 2	1.854***	4.89
Actor pairs at distance-2	0.106***	0.026
Rebuy ego x Rebuy alter	0.028	0.101
Cross-sell ego x Cross-sell alter	0.344***	0.092
Behavior Dynamics Rebuy		
Rate Effect Period 1	2.630***	0.725
Rate Effect Period 2	16.802	8.659
Behavior Rebuy linear shape	0.756***	1.237
Behavior Rebuy quadratic shape	0.372***	0.094
Behavior Dynamics Cross-sell		
Rate Effect Period 1	32.956***	1.182
Rate Effect Period 2	233.222***	58.398
Behavior Cross-sell linear shape	6.060***	1.377
Behavior Cross-sell quadratic shape	1.241***	0.293

Significance: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$

The structural network effect of *actor pairs at distance-2* tests the tendency of salespeople to keep other salespeople at a social distance of two. Therefore, it measures whether network closure is a mechanism that plays a role in network formation and is a common endogenous network covariate (Manger and Pickup 2016). I expect to find a positive effect because salespeople should not be concerned about forming relationships with other salespeople's ties.

The effect of rebuy on the network tests whether salespeople with more rebuy effort have a stronger tendency to form network ties. Similarly, the effect of cross-selling on the network, tests whether salespeople with more cross-selling effort are more likely to form ties with other salespeople. The *ego X alter* effect tests whether the tendency to form a tie depends on the levels of selling effort for two salespeople considering a tie.

Looking at the results, I find that the effect for endogenous network formation, *actor pairs at a distance-2*, is positive and significant. As expected, salespeople do not tend to be connected to other salespeople's ties just to be connected. The *Rebuy ego X rebuy alter* effect is not significant, meaning that salespeople do not create ties with other salespeople based on their rebuy levels. The *Cross-sell ego X Cross-sell alter* effect is significant and shows a positive tendency for salespeople to create ties with other salespeople with higher cross-selling levels. This effect makes sense because salespeople cross-sell to customers who have not purchased from their category. Therefore, other salespeople are more involved in the cross-selling process to provide insights about the customer new to the cross-selling salesperson, or salespeople recommend a customer for cross-selling. The effect is particular to cross-selling because of how cross-selling works and is defined in this research.

The behavior function's linear and quadratic shape effect shows the rebuying distribution. The quadratic shape effect for rebuy is significant, which means the tendency to reach different rebuy levels is non-linear. The quadratic shape effect for cross-selling is also significant and shows that the tendency to reach different cross-sell levels is non-linear. Both effects are positive, which can be interpreted as positive feedback with a positive effect on a salesperson's behavior. This result means that the more a salesperson is cross-selling, the more they will cross-sell in the future. Other work has described a result like this as a signal for a lower chance of influence through others.

Overall, the results of the SIENA model show what I expected to find. The structural endogenous network formation effect provides evidence that salespeople do not solely tend to create more ties to achieve network closure. I also find that more selling does not automatically increase the likelihood of forming ties with other salespeople. Cross-selling is a selling type that increases the likelihood of forming a tie. This result is logical, as cross-selling occurs when salespeople sell to a customer that has never purchased from them. Therefore, it will create a new tie for a salesperson because they have now sold to a new customer to which salespeople in other categories have already sold. Consequently, the cross-selling salesperson forms new ties with salespeople from the other categories. Finally, the results provide evidence that there is no selection regarding tie creation because the assignment of salespeople to customer accounts is based on geography and product category.

2.6 Implications

In this essay, I examine how categories of intellectual capital impact salespeople's choices in B2B acquisition and retention. I first examine salespeople's decisions about which selling type to pursue (acquisition, rebuying, or cross-selling). Subsequently, I identify drivers of success for selling opportunities of a specific selling type. The results from the empirical models show that different categories of intellectual capital foster different types of B2B selling, and not all categories of intellectual capital positively impact selling success.

Human capital, a category of intellectual capital internal to salespeople, positively impacts rebuying because salespeople engage with customers they have sold to in the past. Social capital, which is also intellectual capital internal to salespeople, positively impacts cross-selling as salespeople need knowledge about customers who have only bought from other categories before. Organizational capital, which I control for in the model, has a slightly negative effect on the likelihood of success at closing a sale. Until now, research has only shown that overall sales performance increases because of a firm's network connections, but not which type of selling benefits most from networking. This finding has important implications for managers to understand how salespeople can become more successful at rebuying and cross-selling.

2.6.1 Theoretical Implications

Past research in the B2B acquisition and retention literature has not tied specific categories of intellectual capital to either acquisition or retention types. I examine these categories as drivers of B2B retention and show that intellectual capital internal to salespeople increases the likelihood of engaging in rebuying and cross-selling opportunities. By measuring all categories of intellectual capital (human, social, and organizational capital), I show that human and social capital are drivers of rebuying and cross-selling, respectively.

Specifically, I show that salespeople with higher human capital are more likely to go after rebuying opportunities. This category of intellectual capital is internal to a salesperson and comes with selling experience. Salespeople with higher social capital, which I measure with betweenness centrality, are more likely to choose cross-selling opportunities. Prior research has identified that complex knowledge is necessary for cross-selling but has not given much detail about how salespeople can access this knowledge (Schmitz, Lee, and Lilien 2014). My results suggest that being in a network position that gives access to various customer knowledge helps with cross-selling and succeeding at closing sales. Unlike other intrafirm network research, I specifically focus on the informal interaction between salespeople, which commonly occurs in the field among salespeople, but is difficult to measure.

With this unique data set, I can measure cross-selling success based on historical data and do not need to rely on sales managers' perceptions like in past studies. Contrary to other social network studies, I do not use survey data but historical data to investigate the salesperson network over time. Moreover, I also measure the strength of network ties by adding weights to the network ties based on the time salespeople worked together.

2.6.2 Managerial Implications

My findings show that knowledge from various sources can be helpful in different selling situations. To improve the internal knowledge of salespeople, managers can help by providing learning opportunities for salespeople to gain experience.

One internal knowledge source is social capital. To improve social capital, sales managers can help with relationship development. For instance, managers can address the importance of working together across categories to cross-sell successfully. I also find that cross-selling becomes more likely the more successful salespeople are at cross-selling. Therefore, it is crucial to set up salespeople for cross-selling success to not deter them from pursuing further opportunities. Assuring relationship development between salespeople across categories will help salespeople with accessing customer information and provide the right conditions for salespeople to succeed.

Additionally, as experienced by many over the last year, more work can be done remotely. Thus, often firms now also have a remote onboarding experience. When employees do not start their jobs in person, they miss out on important relationship building that could set them up for success in the future. (Wiseman et al. 2022) found that salespeople who received a more individualized onboarding experience performed better compared to a less individualized experience. Even though many jobs can be done remotely, individual social interactions within a firm should not be ignored. Thus, fostering network building is especially important when salespeople do not get many opportunities to meet other colleagues.

Another way for managers to support network building is to set up sales networking events (in-person or virtual) to meet each other, create new connections, and rekindle relationships. If salespeople have worked on the same customer accounts in the past, they share a network tie but could have lost touch if they now work on different customer accounts. Hence, bringing salespeople together can facilitate communication within and across customer accounts.

2.7 Limitations and Future Research

One limitation of my work is that I only have access to an intrafirm network of one B2B firm, where I can only observe the effect of salesperson networks in one industry. However, many B2B firms are structured similarly, and the informal communication patterns among B2B salespeople should not differ based on the industry. Future research should test whether my findings hold with data from multiple B2B firms in other industries. In addition to analyzing other industries, it would also provide more insights to analyze B2B data that includes information about firms' events and the frequency of opportunities where salespeople can interact with each other.

By using secondary data for the network analysis, I avoid the self-reporting bias of network connections. One limitation of the data is that I do not have survey data to compare the differences between the data collection methods. Future research should test secondary network data and survey-based network data at the same firm to provide more insights into the usefulness of secondary network data.

Chapter 3

Essay 2: Identifying Community Structures Among Social Media Followers: Evidence From an Online Ego Network Perspective

3.1 Introduction

No matter the type of organization (e.g., for-profit or non-profit), organizations miss out on exchanging information with consumers and other businesses without having a social media presence. Some organizations' social media accounts have even attracted a cult following by communication with witty social media posts. For instance, Wendy's Twitter account currently has 3.8 million followers.[1] Even though all these people follow Wendy's, they probably do not follow the account for the same reasons and may have different expectations from the organization. Similarly, American Express's Twitter account with 857,000 followers also attracts users for various reasons.[2] Some users may follow to get information about promotions, and others may be interested in travel tips through American Express travel. The account can also serve B2C or B2B customers and inform them about the company's events and sponsorships. Users could also follow American Express for more than one reason and cannot be categorized into just one type of interest.

Through social media posts, organizations can connect with many different followers. Social media posts are also a great way to listen to consumers, but most social media users do not engage with organizations' social media accounts. On average, an organization's post only has

---

[1] https://twitter.com/Wendys
[2] https://twitter.com/AmericanExpress

an engagement rate of 0.045%.[3] Besides these standard social media metrics (e.g., engagement rate, number of followers), organizations analyze social media posts for feedback from followers and customers (McKay 2017). They can also observe and analyze keywords, hashtags, demographic information, and geo-locations to understand their audience (Newberry 2020). However, the accuracy of the information on social media platforms depends on what users decide to reveal on their profiles. Thus, interests identified with analytics tools based on profile information might not always be correct.

One part of the social media data that organizations typically do not leverage is their followers' online social network connections. Network connections are available with all social media accounts and can give insights into how followers connect over similar interests. People often create connections based on homophily. Thus, understanding how followers are connected helps organizations understand their followers better.

Within the follower network, followers follow an account for different reasons, and organizations must consider communicating different messages to different follower segments. Knowing whom you are talking to is essential to engaging users. Therefore, a social media account's content needs to interest the audience. Organizations can proactively communicate the right content if they understand their followers' interests.[4]

In this essay, I propose a multi-step research framework to segment and target social media followers by incorporating online social media networks of followers and text data from online social media posts to identify various follower segments. In academic research, online social networks have been the focus of various studies in the past, often in the context of product

[3] https://www.statista.com/statistics/798414/daily-twitter-brand-audience-mentions-by-vertical/
[4] https://blog.hootsuite.com/social-media-engagement/

adoption, diffusion of information, and social influence (Ebbes, Huang, and Rangaswamy 2016; Hinz et al. 2011; Katona, Zubcsek, and Sarvary 2011). However, not many studies have used the group structures in a network to identify follower segments on social media.

Additionally, most studies have used the full network in their analysis. Full networks include all existing ties within a defined set of actors (Stolz and Schlereth 2021). In practice, it is not feasible to analyze a full network if a user does not have a direct connection to an account. In this essay, I use the ego network approach. It alleviates this data collection problem because an ego network only contains first-degree ties (i.e., direct connections) and how these ties are connected (Stolz and Schlereth 2021). This approach is a novel contribution to both the marketing literature on social media and to industry. Organizations can implement this approach to collect their ego network by observing first-degree ties to their social media accounts.

To understand the group structure of followers on social media and how to extract follower characteristics from segments, I answer the following research questions with a multi-step research framework.

What are the follower insights generated from segments when using:

1) standard network measures (i.e., centrality measures)?
2) discrete segmentation within a firm's ego network?
3) overlapping segmentation within a firm's ego network?
4) overlapping segmentation detected from ego network connections and social media posts?

In my analysis, I begin with standard social network measures, which give information about individuals in the network and the overall network structure. Next, I apply a non-overlapping community detection method (i.e., users can only be members of one community) to segment followers based on their network connections. These results give follower segments but do not allow followers to be part of multiple segments, which is an unrealistic assumption. Therefore, I then relax the condition of unique community membership and allow individuals to have multiple community memberships with an overlapping community detection method. These findings give a more realistic identification of follower segments. However, the method does not give any information about followers' interests. Therefore, I go beyond assigning communities based solely on network ties by adding text data on topics from user posts. Here, I use an overlapping community detection method with node attributes, assigning characteristics to each follower.

The overlapping community detection with node attributes exposes community structures based on network connections and common topics of interest among users. In contrast, other methods ignore some connections or ignore user similarities. By utilizing connections and post topics, I find that this method is the only one that differentiates between broad and niche topics. Understanding the interest in the different topics could help organizations target specific users with niche topics and choose more broad topics for all followers.

In the following sections, I go over the background literature, the research framework, the methods of analysis, and my results. Additionally, I discuss implications for theory and practice, limitations, and potential for future research.

3.2 Background Literature

I draw from the social networks and customer engagement literature to develop my research framework. The following section discusses past findings and how my research closes gaps in the existing literature.

3.2.1 Online Engagement

Even though measuring the direct impact of a social media presence can be challenging, organizations can use the engagement metrics of their social media accounts as a tool to help achieve strategic goals. These goals include creating brand awareness, retaining customers, and increasing revenue. Consequently, online social media provides many opportunities for organizations to communicate with customers.

Customer engagement is how organizations initiate a firm-to-customer interaction to create a bond between the firm and the customer (Kumar and Pansari 2016; Meire et al. 2019). Organizations must understand what motivates customers to engage on social media to create a strong emotional bond. The organization's post content must be relevant to customer interests. Thus, many organizations create content beyond promoting products to determine what customers care about (Venkatesan 2017). For example, Meire et al. (2019) examine what type of information leads to different customer sentiments. Additionally, more posted content from organizational social media accounts can increase customer engagement with the organization on social media (Gill, Sridhar, and Grewal 2017). The result of more interactions between customers and an organization is higher trust and loyalty, which means that organizations must be active in creating content on social media (Hajli et al. 2017).

Organizations can improve customer engagement by catering to multiple customer segments because they can reach more followers (Venkatesan 2017). However, most communication on social media is not targeted to individual customer segments and looks more like mass marketing. Organizations can improve engagement and successful diffusion by understanding which type of information resonates best with individual follower groups (Vermeer et al. 2019). Thus, organizations should align their online social media content with customer interests.

In this research, I utilize users' frequent social media post topics in combination with the social network connections of the organization's social media account to identify different groups of followers and their common interests. To the best of my knowledge, there has not been any research in the marketing literature that uses online social network connections and post topics to segment and target online social media followers.

3.2.2 Social Networks

Katona, Zubcsek, and Sarvary (2011) argue that marketing strategies can be implemented more effectively with the help of social media networks because these networks give insight into customer communication patterns. Thus, past marketing literature has measured network centrality to identify influential individuals in social networks (Katona, Zubcsek, and Sarvary 2011; Trusov, Rand, and Joshi 2013). Identifying the most influential nodes in a network can provide helpful insights, for instance, for seeding information in an online social network (Hinz et al. 2011). Additionally, knowing whom to target based on the network position is better than not utilizing any information about the ties in a social network. Past academic studies have also identified influential actors based on different traits, particularly the most influential positions within networks (e.g., Bakshy et al. 2011; Van den Bulte and Joshi 2007; Trusov, Bodapati, and Bucklin 2010; Yoganarasimhan 2012).

However, researchers have also found that individuals get a sense of identity from a cohesive network structure (Susarla et al. 2012; Yoganarasimhan 2012). In a cohesive network, people share similar viewpoints and relate to each other's values and interests. These findings suggest that it is important to identify cohesive sub-groups in social networks instead of only focusing on one influential individual. Subgroups are also called communities and are defined by the frequency of members' interactions. Thus, members inside the community have more in-group interactions than out-group interactions (Tang and Liu 2010). For example, group members have common interests, which gives them reasons to interact frequently.

Social network analysis (SNA) helps identify influential individuals. However, we also know that followers can group over similar interests and targeting multiple follower communities is most effective for engagement (Venkatesan 2017). I add to this literature by introducing a novel approach to segmenting social media followers with community detection based on network connections and node attributes (e.g., follower interests) and utilizing ego networks. Moreover, this approach helps uncover community structures, enabling organizations to rely less on broadcasting messages to everyone and instead target different communities based on their interests. Organizations can use these methods in practice to collect and process network data.

3.3 Research Framework

I propose a multi-step research framework to provide a novel approach for the segmentation and targeting of social media followers (see Figure 3-1). I start by explaining some of the standard methods of analysis, which range from individual network measures to community detection methods. With each additional method, I alleviate the challenges of previous methods. In the subsequent section, I describe each method in order of complexity, going from the least to the most complex.

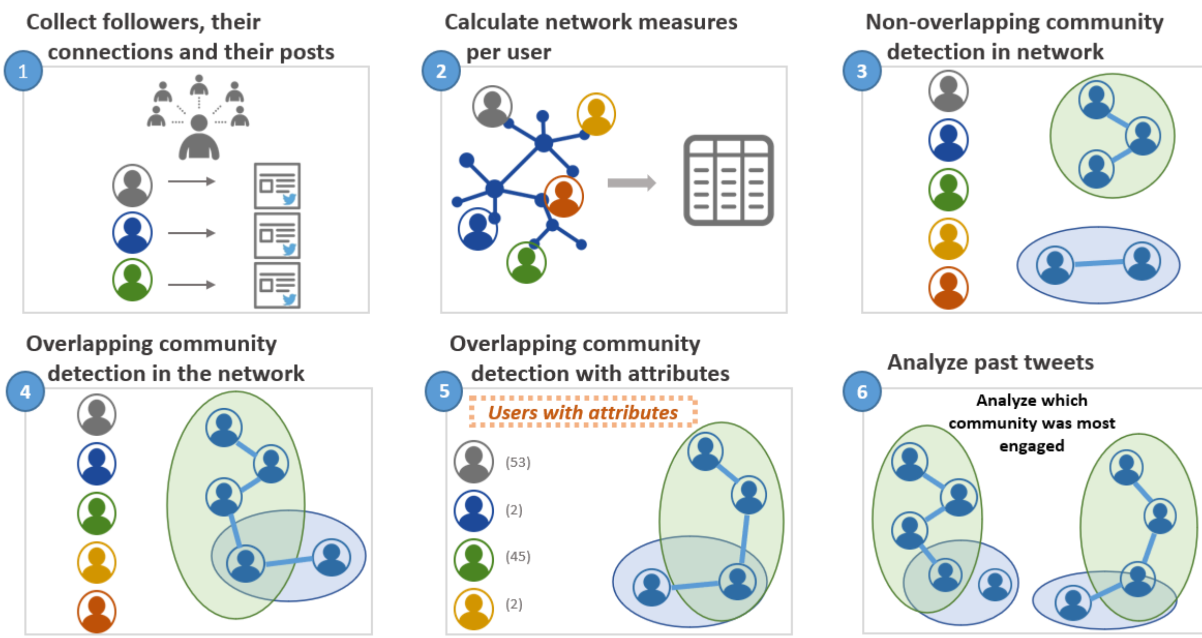


Figure 3-1: Research Framework

3.3.1 First Method: Social Network Analysis

The most commonly used measures in SNA are centrality measures. These measures give insight into connections between followers, identify positions within a network, and estimate how information can spread through the network.

Measures. The most used network centrality measures are degree centrality and closeness centrality. Degree centrality counts the number of nodes adjacent to a node, providing a proxy to measure the activity of an actor in a network (Wasserman and Faust 1994). For example, if a user has connections to three other users in the network, the number of degrees is also three. However, past research has shown that the number of connections does not automatically translate into having influence and may not be as effective of a measure (Katona, Zubcsek, and Sarvary 2011). Moreover, this measure only gives information about either the full network or individual nodes but not on subgroups within the network.

Closeness centrality counts how many individuals one must go through to reach all nodes in the network and measures how closely connected individuals are with others. It is the inverse of the sum of distances from one node to all other nodes (Freeman 1979). This measure causes problems if there are individuals with no connections. If nodes are undefined (i.e., disconnected), no path exists to them (Borgatti, Everett, and Johnson 2018). Therefore, it is not an ideal measure for all networks.

The clustering coefficient (i.e., transitivity) also measures the interconnectedness of a network (Watts and Strogatz 1998). It divides the number of existing edges by the total number of neighbors of the actor. One can calculate the clustering coefficient for the full network or each node. For this case, I use the individual measure because I first want to get information about individual social media followers.

Constraint. While standard network centrality measures can give insights into a node's position and connectivity, they do not tell us anything about group structures or behavior. In addition, these network measures do not provide information about follower interests and what topics might resonate with followers. Therefore, as a next step, I examine community detection methods to group followers based on their network ties in an ego network.

3.3.2 Second Method: Non-Overlapping Community Detection

SNA offers several ways to detect communities to identify heterogeneous groups of followers within a network. A community is a subgroup in a network, where nodes within a community interact more frequently than with nodes outside the community (Tang and Liu 2010). Past research has used community detection in the marketing literature to identify competitive market structures and brand-associative networks (Netzer et al. 2012; Ringel and Skiera 2016). However, very little research has used community detection algorithms to segment online social media followers.

There are four distinct categories of community detection methods: node-centric, group-centric, network-centric, and hierarchy-centric (see Table 3-1). Node-centric community detection defines communities by identifying all nodes that contain the same properties. For instance, a community could be a group of nodes with the same number of ties (degrees). Another more specific method is to measure cliques or k-cliques. A clique is the strictest approach because all nodes in the subgraph (community) must be fully connected. Whereas for k-cliques the condition can be relaxed because each node must have *k* number of connections within the clique. This method is time complex as it takes a long time to identify cliques within large networks.

Table 3-1: Community Detection Categories

Community Detection Category	Definition	Method Examples
Node-Centric	All nodes must satisfy the same properties.	cliques, k-cliques, k-clan, k-plex
Group-Centric	The group as a whole has to satisfy the same properties.	group density
Network-Centric	Nodes show structural similarities to partition the network into disjoint sets.	vertex similarity, latent space models, block models, spectral clustering, modularity maximization
Hierarchy-Centric	Use of network topology to create hierarchical community structures.	edge betweenness, agglomerative hierarchical clustering

Tang and Liu (2010)

Group-centric community detection measures the properties of each community. For example, each subgraph must have a specific density defined by a threshold. Network-centric community detection means that nodes must have similarities in network structure. For instance, one method is vertex similarity, which looks at whether two nodes are connected to the same nodes and thus, are structurally equivalent. In practice, structural equivalence is too restrictive to detect communities.

Lastly, hierarchy-centric community detection uses network topology to create communities. Divisive hierarchical clustering shows a hierarchical structure of groups by partitioning the network into disconnected sets, and each set gets divided further until the sets contain a small number of nodes. One way to apply divisive hierarchical clustering is by using edge betweenness, which means that edges with high betweenness get removed progressively (Newman and Girvan 2004; Tang and Liu 2010). The disconnection of the networks results in a hierarchical structure. One of the limitations is that it has a high computational cost, so it is challenging to apply it to large networks (Tang and Liu 2010).

The other type of network-centric community detection is agglomerative hierarchical clustering. The approach measures the deviation of the network interaction from the expected random connections (Tang and Liu 2010), known as modularity. At first, the algorithm assigns one community per node. It then considers the neighbors for each node and compares the modularity from moving the node out of its community and placing it into the neighboring community. The process places the node in the community with the highest modularity gain. As a second step, the algorithm then regards every community as a node, and the optimization process gets repeated (Blondel et al. 2008; Newman 2006).

Measure. As a non-overlapping community detection method for online social media, an applicable community detection algorithm must be able to handle large amounts of data. Following Yang, Algesheimer, and Tessone (2016), I utilize the Multilevel algorithm (Louvain method) from Blondel et al. (2008), which uses agglomerative hierarchical clustering, the preferred method for large networks.[5] The Louvain method has the fastest computing time and best accuracy.[6] The results of the Louvain method are stable for different network sizes, which is desirable when dealing with social media data.

---

[5] The algorithms they compare to find the best method for large networks (>1000 nodes) are all part of the igraph package in R (Edge betweenness, Fastgreedy, Infomap, Label propagation, Leading eigenvector, Multilevel, Spinglass, Walktrap) and are built on the different community-detection categories. The algorithm that outperforms others is the Multilevel algorithm from Blondel et al. (2018), also known as the Louvain method, which uses agglomerative hierarchical clustering. The authors used the LFR benchmark graphs (for Lancichinetti, Fortunato & Radicchi) to test the different algorithms. The parameters used for the LFR benchmark graphs are the number of nodes, maximum degree, maximum community size, average degree, degree distribution exponent, community size distribution exponent, and a mixing parameter, which divides the external degrees of a node (i.e., outside of the community) by the total degrees of that node (Lancichinetti, Fortunato, and Radicchi 2008).

[6] Yang, Algesheimer, and Tessone (2016) measure accuracy with two methods. The first one is the measure of normalized mutual information (NMI), which compares the number of estimated communities to the real ones, thus if NMI equals one, the real and estimated communities are the same. The other way to measure accuracy is with the ratio of the number of detected communities and the number of communities generated by the LFR model.

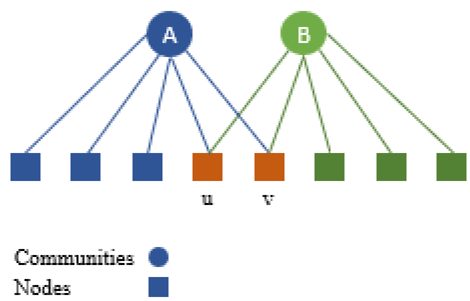
Constraint. The Louvain method helps identify communities in large networks based on network connections. However, it creates distinct communities without allowing for multiple community memberships. Followers are forced into one community, while the algorithm ignores network ties with members from other communities. Realistically, online social media users do not only have one interest and one circle of users with whom they interact. Instead, users are interested in a variety of different topics and, thus, belong to multiple groups that share similar interests. Therefore, I apply overlapping community detection below to relax this restrictive constraint.

3.3.3 Third Method: Overlapping Community Detection

In real-world online networks, nodes can belong to more than one community because individuals are usually members of more than one social group (Yang and Leskovec 2012a). Social groups form around friendships or topics of interest. Thus, one criterion should be to allow community members to be part of multiple communities when applying community detection. Even though node-centric community detection methods let nodes be part of multiple communities, many existing models show overlapping community parts as sparsely connected. However, Yang and Leskovec (2012a) find that the overlapping parts of communities have more dense connections than non-overlapping parts. By identifying densely connected overlaps, the authors also extend the concept of homophily. In the past, homophily has been identified through one dimension of interest, but the users' dense connections across overlapping communities show that users have several similar interests.

Measure. The Community-Affiliation Graph Model (AGM) is a probability model where the probability of being a member of a community is based on shared connections. It measures the most likely community affiliation of a node in a graph. By having more interests in common, users are also more likely to connect (Yang and Leskovec 2012a). Yang and Leskovec (2012a) test the model with ground truth data (i.e., data about real-world community membership), which is not always available, and verify the accuracy of this community detection method.

Figure 3-2 shows two communities (A, B) and two nodes *u* and *v* that belong to both communities. With the two nodes being part of two communities, they share an overlap and are more likely to share a connection (i.e., edge).



Adapted from (Yang and Leskovec 2012a)

Figure 3-2: Community Affiliation Network

The probability $p(u, v)$ of the AGM creating an edge E between the nodes $(u, v) \in E$ is measured with the following equation:

$$p(u, v) = 1 - \prod_{k \in C_{uv}} (1 - p_k) \tag{1}$$

Where $C_{uv} \subset C$ is the set of communities that *u* and *v* have in common, and *k* is the common community membership (Yang and Leskovec 2012a). The AGM has the advantage that it was tested against ground-truth when the authors developed the model, which means that they could reliably test different methods with ground-truth data to compare whether the detected community resembles communities in the real world. Overall, it is more accurate and scalable for large networks than other community detection methods.

Constraint. The AGM shows realistic assignments for multiple communities based on network ties, and it suggests that users connect over multiple similar interests (i.e., pluralistic homophily). However, the results from the AGM do not tell us what interests make users connect. The extension below adds attributes to nodes to get additional information for the community detection process.

3.3.4 Fourth Method: Overlapping Community Detection with Node Attributes

Adding node attributes to community detection means detecting groups that form over similar interests (Susarla et al. 2012). Because communities form over various interests, information about users (i.e., nodes) can be used as a foundation for community detection. For example, organizations can collect user information from their social media profiles for online social media networks. Due to privacy concerns, users do not always promote their personal information on social media. Thus, organizations might only get sparse data.

One part of social media data that organizations can access is followers' posts. Collecting text from posts is generally a good listening tool, but organizations can also extract topics from posts. These topics give insight into what followers talk about and their interests. Therefore, I can use topics of interest as node attributes (i.e., follower attributes). In addition to the topics, I still use social network connections between followers to identify communities.

Most existing clustering methods focus on attributes or network connections (Yang, McAuley, and Leskovec, 2013). Additionally, existing methods that consider both types of information sources either do not allow for overlaps or do not allow strong membership to multiple communities. I use a method that combines overlapping community detection based on connections and node attributes for the research framework.

Measure. The method I use is called ‘Community from Edge Structure and Node Attributes’ (CESNA). The CESNA algorithm calculates two weights for each node, one based on connections and the other on having an attribute (Yang, McAuley, and Leskovec 2013). The combined weight of a node must pass a threshold to determine community assignment. It uses the probabilistic AGM process by maximizing the likelihood of the community membership weight and the logistic weight of community membership to a node attribute (Martínez-Seis 2017; Yang, McAuley, and Leskovec 2013).

For the estimation of the network weight, the algorithm assumes that two nodes (u, v) of a community c have the probability P_{uv} and are connected through community membership F (see Equation 2).

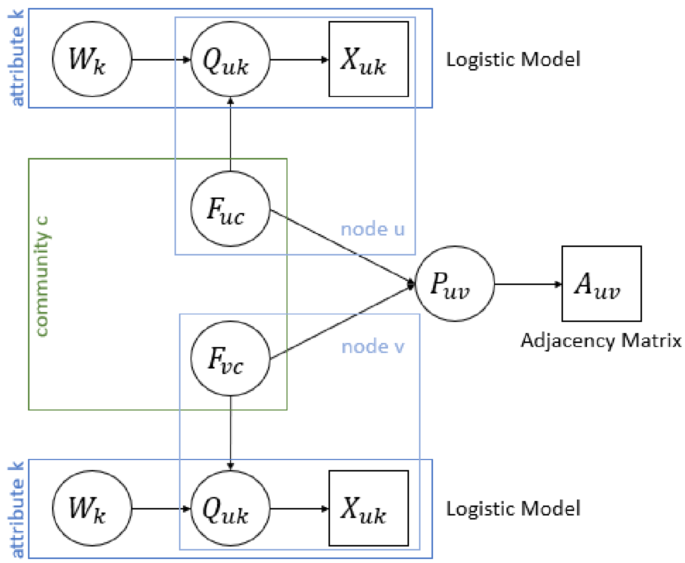
$$P_{uv}(c) = 1 - \exp(-F_{uc} \cdot F_{vc}) \tag{2}$$

If none of the nodes u or v are part of c, then $P_{uv}(c) = 0$ because they are not connected. The node attributes are modeled with a logistic regression model. The assumption for the k-th attribute of a node u X_{uk} is that attributes are binary-valued. The input features are group memberships of a node $F_{u1}, \ldots, F_{uc}$ and the associated logistic weight factor W_{kc}. The probability that X_{uk}=1 is Q_{uk} modeled with the following equation:

$$Q_{uk} = \frac{1}{1 + \exp\left(-\sum_c W_{kc} \cdot F_{uc}\right)} \tag{3}$$

Figure 3-3 shows how the models for each weight work together in the CESNA model. To determine community membership for a node u, the respective F_u must be above threshold δ (see Equation 4).

$$\delta = \sqrt{-\log\left(1 - \frac{1}{N}\right)} \tag{4}$$



Adapted from (Yang, McAuley, and Leskovec 2013)

Figure 3-3: CESNA Model

Constraint. The CESNA model can assign social media users (e.g., followers) to communities based on their connections and node attributes. One constraint of the model is that it automatically assigns nodes into communities based on the general threshold δ, which performs well compared to other selected values. Thus, some users are not assigned to any communities because they do not meet the cut-off. To overcome this constraint, I manually calculate a network and attribute weight to assign followers to communities.

As a first step, I calculate the network weight for each unassigned follower. Following Yang and Leskovec (2012b), I calculate a conductance score for each follower. The conductance measures “the fraction of total edge volume that points outside the cluster” with the following equation:

$$f(S) = \frac{c_S}{2m_S + c_S} \tag{7}$$

Where S is the set of nodes, c_S is the number of edges on the boundary of S, and m_S is the number of edges in S. With the conductance score, I get a follower weight for the likelihood of being a member of the communities. As the attribute weight, I use the logistic weights per community calculated by the CESNA model. After we have both weights, I add them together to get a combined community weight per community for each follower. Subsequently, I assign followers to a community with the highest combined weight.

3.4 Empirical Application

In the empirical application, I apply the research framework to online social media data collected for one organization. Figure 3-4 depicts the steps to go through the various community detection methods. As a first step, I collected online social media data. Next, I analyzed standard social network measures to get an overview of the network structure and insights into individual followers. Subsequently, I apply the Louvain method (i.e., non-overlapping community detection) to identify distinct communities. Next, I allow for overlapping communities by using the AGM (i.e., overlapping community detection) to investigate how users cluster together when allowed multiple group membership. Lastly, I add another layer of follower information and use the CESNA model (i.e., non-overlapping community detection with node attributes) to identify follower communities based on their connections and common topics.

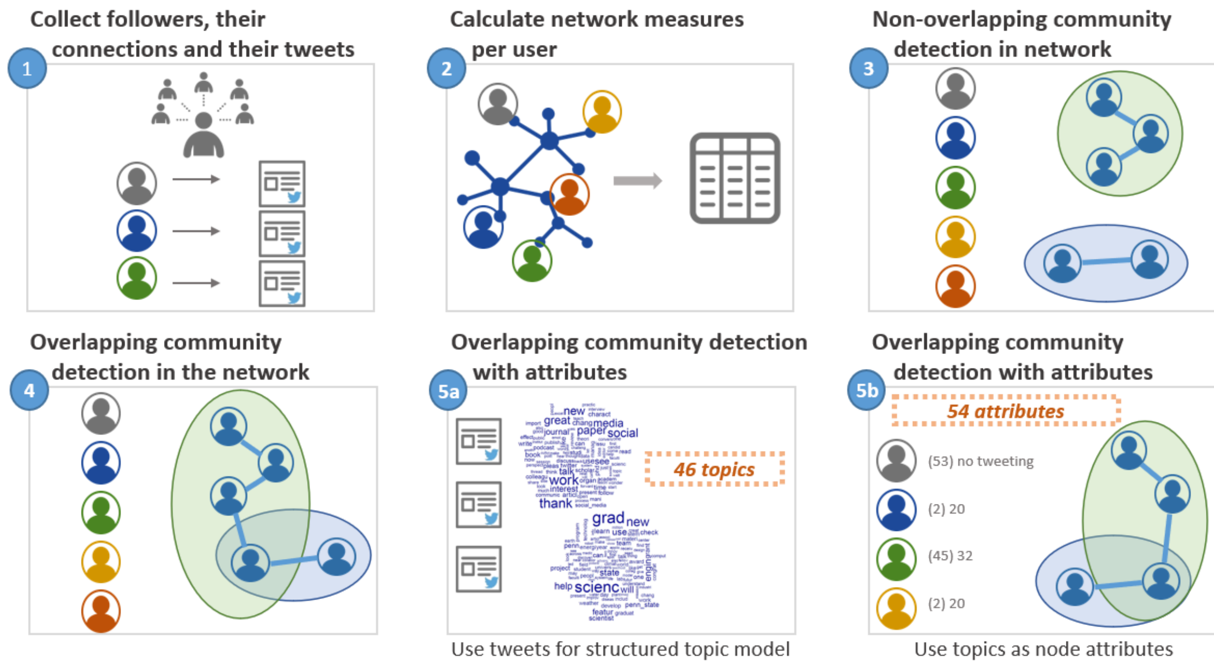


Figure 3-4: Empirical Application

3.4.1 Step 1: Data Collection

The data I use comes from a U.S. business school's Twitter account. Over the timeframe of November 2018 until July 2020, I scraped tweets that mentioned the business school's account and retweeted the school's tweets. To collect the tweets, I used NodeXL, similar to (Sterner and Felmlee 2017), where I entered a search term to collect all tweets with that search term, including all followers of Twitter users who tweeted. The search term was the account's Twitter handle to compile all communication with the account. The program NodeXL creates a network tie if a user retweets, "mentions" (if an account's name appears in the tweets), or follows someone else.

As a second data collection step, I also scraped all followers of all Twitter users collected in the previous step. Since I am interested in the followers of the business school and it is not feasible to collect the full network with all second-degree connections, I use the business school as the network's focal node (ego). The final ego network data of the account has 5,917 nodes and 55,241 connections. Because I collected the follower data of an ego, every user in the network has at least one connection from being connected to the ego (focal Twitter account). The total number of tweets is 92,784 from 718 active (i.e., engaged) followers.

Figure 3-5 depicts the ego network of the business school. The nodes that have many degrees are large circles. From the network graph, I can see that some followers cluster together by sharing some of the same connections, which is model-free evidence that subgroups of followers exist in the network.

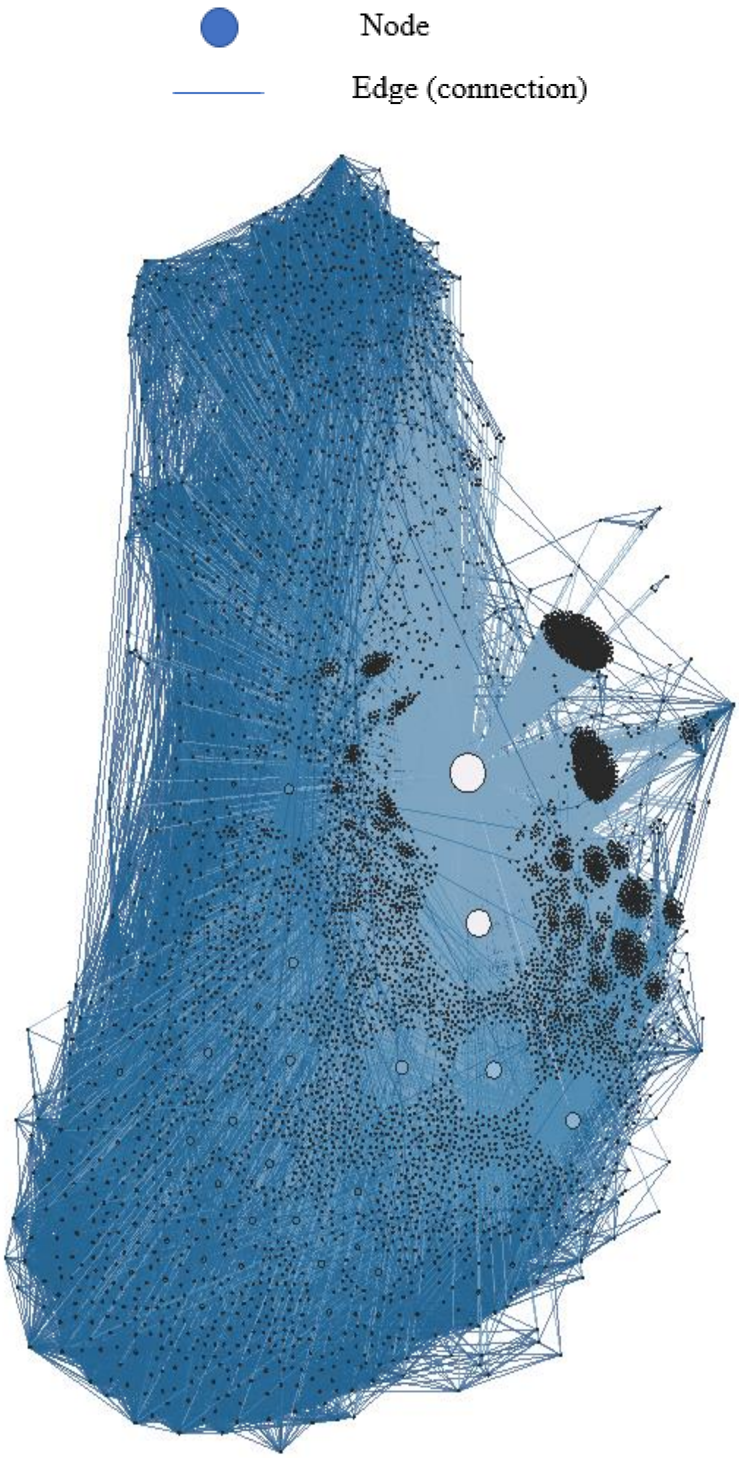


Figure 3-5: Network Graph of Business School Ego Network

3.4.2 Step 2: Social Network Analysis

As a first step, I analyze the ego network of the business school. The results give us insights into the network structure and the connections between followers. For an overview of the ego network, I measure the total number of degrees, transitivity, closeness, and followers' engagement (see Table 3-2). Because the network data is undirected, I measure degree without direction (i.e., in- and out-degree). In the network data, the mean degree is 18.67 with a standard deviation of 118.5, which means that, on average, a user has almost 19 connections. However, there is a high variation in the number of degrees within the network.

Table 3-2: Standard Social Network Measure Results

Network measure	Mean	Standard Deviation
Degree	18.67	118.50
Transitivity	0.86	0.33
Closeness	0.00008	0.000001
Engagement	1.59	10.51

The network transitivity measures the ratio of whether adjacent nodes of a node also share degrees with another. For instance, if a follower has two degrees, the transitivity measure gives insight into whether the follower's two connections are also connected. The mean transitivity in the data set is 0.86 with an S.D. of 0.33. This result shows that, on average, users' ties are strongly connected.

Another measure I use to identify the network structure is closeness centrality. The closeness centrality for the Twitter network is 0.00008, which is the inverse of the sum of distances from one node to all other nodes. Since the network is a rather large network with 5,917 nodes (large network ≥ 1000 nodes) (Yang and Leskovec 2012a), it seems reasonable that the closeness centrality is low because the network nodes spread far.

Lastly, I also measure the level of engagement among the users in the network. I define engagement as the sum of activities on Twitter, including tweeting, mentioning other accounts, and replying to tweets. The mean engagement is 1.59 with an S.D. of 10.51. For reference, the average engagement of a post on Twitter is 0.07.[7] Thus, the followers in the ego network are relatively active.

These findings show that the network has a few followers with many degrees (i.e., connections) and a few who actively tweet. While these summary network measures give us an idea about the network's structure, they do not tell us how and why users connect. Additionally, it is not feasible to target every user individually on Twitter with personalized tweets. Therefore, it may be helpful to investigate how followers can group together and whether understanding these heterogeneous groups can help better position tweets among the right target population.

3.4.3 Step 3: Non-overlapping Community Detection

After running the Louvain algorithm in the igraph package in R, I get a result of five non-overlapping communities (see Table 3-3). Each Twitter user is assigned to one distinct community. The smallest community has five members with ten connections, meaning that every member is connected. The largest community has 3,925 members and is also the community to which the ego is assigned (community 3). From this analysis step, I get distinct groups of users that may be used for targeting. However, it is challenging to infer similar interests just from homophily.

[7] https://www.socialinsider.io/blog/social-media-industry-benchmarks/

Additionally, this method does not allow users to be part of multiple communities; thus, users are forced into one community based on the strength of their network ties and other network ties are ignored. To relax the condition of distinct community membership, the following method of analysis I apply is overlapping community detection, which allows users to be part of multiple communities.

Table 3-3: Results of Non-Overlapping Community Detection

Community	Members	Edges	Min Degree	Max Degree	Mean Degree
1	704	6,515	1	318	19
2	1,021	19,776	3	763	39
3	3,925	13,782	1	3,924	7
4	5	10	4	4	4
5	262	870	1	93	7

3.4.4 Step 4: Overlapping Community Detection

In real-world networks, people belong to more than one community based on different interests. Thus, I apply the AGM to the data to allow for multiple community memberships. After running the AGM, I get a result of six communities with overlapping membership of users. All six communities are closer in size than the non-overlapping communities from the previous step, ranging from 367 to 2,763 members, and members seem more evenly distributed across communities (see Table 3-4). The ego of this network is a member of all six communities, which validates the algorithm's robustness because the ego has connections to all other users.

Table 3-4: Results of Overlapping Community Detection

Community	Members	Edges	Min Degree	Max Degree	Mean Degree
1	367	19,154	22	564	104
2	531	14,410	8	639	54
3	564	10,890	6	652	39
4	1,057	16,768	4	1,142	32
5	1,654	12,975	1	1,760	16
6	2,763	19,718	1	2,923	14

Table 3-5: Summary of Users in Overlapping Communities

Community Memberships	Number of Users	Mean Degree	Mean Closeness	Mean Transitivity	Mean Engagement
1	367	19,154	22	564	104
2	531	14,410	8	639	54
3	564	10,890	6	652	39
4	1,057	16,768	4	1,142	32
5	1,654	12,975	1	1,760	16
6	2,763	19,718	1	2,923	14

In total, 28 users are members of all six communities (see Table 3-5). These members have an average engagement of 96.89, which shows that the most active users are also the most connected and may connect over various interests. However, most users are only part of one community, which still does not reveal enough information to segment users and target the respective groups. Thus, as a next step, I apply the overlapping community detection with attributes by segmenting the network and using followers' interests based on conversations through tweets.

3.4.5 Step 5: Overlapping Community Detection with Node Attributes

To understand how followers cluster together around network connections and similar interests, I apply the CESNA algorithm that detects overlapping communities based on network connections and node attributes. I expect that users group around specific topics, which can help identify different follower segments. As a first step, I identify and assign the node attributes before running the CESNA algorithm.

Topics as Attributes. The attributes assigned to a node can be any available characteristic of a node. For this analysis, I utilize the tweet text I have collected, and I apply topic modeling to the tweet text to analyze the tweets and extract user interests. With the topics prevalent in the tweets, I can then assign user attributes based on the most likely tweet topics.

*Topic Modeling***.** I use structural topic modeling (STM) on the data and use the 'stm' package in R. STM is built to uncover topics in open-ended text and allows the use of metadata to estimate the relationship between the topics and the metadata (Roberts, Stewart, and Tingley 2019).

Before applying STM to the data, I clean and pre-process the tweet text. The total number of tweets in the data is 92,784 from 718 users. As a first step, I combine the tweets per user into 718 text documents. Next, I go through the following data cleaning steps to reduce noise in the data:

1. Remove punctuation – This step removes any non-alphabetical characters (e.g., period, comma, parenthesis) used in the tweets.
2. Remove any symbols and numbers – Because I use social media posts as data, most posts contain the @ or # symbol, which I remove before applying the STM. Additionally, I also remove any numbers from the tweet text.

3. Remove urls – Social media posts often contain urls to link to another website. This information is not relevant for identifying topics, so I remove them from the text.
4. Remove profane words – I also remove any derogatory language[8] from the text because I am not looking to identify the sentiment of the posts. The profane terms should not provide any insights into topics.
5. Remove common words (i.e., stop words) – Commonly used words do not give us any specific information about context or topics. Some of these words are the, is, and at. Thus, I remove the top 100 words in the English language.

As a next step, I tokenize (i.e., break text into units) each word and create uni- and two-grams (Berger et al. 2020). N-grams are n words that are grouped together. For example, two-grams are two words grouped into one term based on their frequency of appearance. I also stem the words in the text data because the word family will have the same meaning, whether a verb or a noun. The stemming process only keeps the root form of words.

As the last step of preparing the text data, I use the 'prepDocuments' function to get the data into the correct format and remove words that appear less frequently in the text (Roberts, Stewart, and Tingley 2019). The R package allows to either manually choose the number of topics or let the model automatically select the number of topics through spectral initialization. The model solves for the convex hull of the word co-occurrences to find the number of topics. I can also examine topic diagnostics for a manual topic number selection. The number of topics generated by the model is 46 topics. This number aligns with the topics diagnostics of exclusivity and semantic coherence, which should be high while still achieving low residuals. These results

[8] https://www.cs.cmu.edu/~biglou/resources/bad-words.txt

from the topic modeling show what followers tweeted by topic. For example, topic 28 shows words that all seem related to sustainability (see Figure 3-6).

Figure 3-6: Word Cloud of Topic 28

*Topics as Attributes***.** The STM results show that followers are interested in diverse topics. Based on the topic proportion assigned to each follower, I find that some followers have a wide dispersion of topics in their tweets. To get insight into how dispersed topics are per follower, I calculated the entropy score with the following equation (Borda 2011):

$$\mathrm{H}(X) = -\sum_{i=1}^{n} P(x_i)\, logP(x_i) \tag{6}$$

Where H is the entropy, *X* is the information content, *P* refers to the proportion. The minimum entropy score is 0.28, and the maximum is 3.48. The mean score is 2.23, which shows that most followers talk about many different topics. The lower the entropy score, the more concentrated the topics. To factor in a topic's relevance, I looked at the proportions per topic and only selected topics that passed a threshold of 0.2. Because I want to use topics as follower attributes to depict their interests, I am only interested in high prevalence topics.

CESNA uses the overlaps in attributes to assign individuals to communities. Therefore, I need to have several attributes that enough followers share, but at the same time, the attributes need to be different enough to show variety. The topic proportion threshold with the highest average followers per attribute was 0.2, with an average of 9.94 followers per attribute. Thus, I used a topic proportion as a threshold of 0.2 and the top two topics based on proportions per follower. Some followers never tweeted, and some did not pass the topic proportion threshold. For those followers, I assign an attribute called 'no tweeting' or 'diverse tweeting' (i.e., did not pass the threshold) respectively. Table 3-6 shows the final list of attribute numbers with the associated topics with 54 attributes.

Table 3-6: Attribute Overview – List of Topics Assigned as Attributes

Attributes	Topics	Attributes	Topics
0	1	27	31;16
1	10	28	31;26
2	11	29	31;4
3	12	30	32
4	13	31	32;11
5	15	32	32;31
6	16	33	33
7	17	34	34
8	18	35	35
9	2	36	36;33
10	20	37	37
11	21	38	38
12	22	39	39
13	22;11	40	39;31
14	23	41	4
15	24	42	40
16	24;16	43	41
17	25	44	43
18	25;12	45	44
19	26	46	45
20	26;4	47	46
21	27	48	6
22	28	49	7
23	29	50	8
24	3	51	9
25	30	52	diverse tweeting
26	31	53	no tweeting

CESNA Model. I ran the CESNA model in C++ with the attribute list and network data. As a result, I get three communities with overlaps. The business school account (i.e., ego) is a member of all three communities, which provides some validation for the community assignments from the model. The communities' memberships range in size from 219 (community 2) to 453 (community 3) and 601 (community 1) members (see Table 3-7). These communities are even closer in size than the communities resulting from AGM. However, based on these numbers, only 1,273 followers out of 5,917 were assigned to communities, which is one method constraint. The algorithm drops any user with less than 10 degrees from the community assignment, and any user whose weight does not pass threshold δ does not get assigned. This threshold cuts off people from community assignments. Thus, I manually calculate the network and attribute weight to assign followers to the three communities.

Table 3-7: Results of Overlapping Community Detection with Attributes

Community	Members	Edges	Min Degree	Max Degree	Mean Degree
1	601	25,636	18	823	85
2	219	4,122	4	279	38
3	453	8,616	4	523	38

Manual Community Assignment. After calculating the network weight from the conductance score and the attribute weight for each non-assigned user, I can create a community membership for each user. Overall, I manually assigned 160 users to community 1, increasing community members' total community membership for community 1 to 761. I assigned 4,403 users to community 2, making community 2 the largest community with 4,622 members. Lastly, I added 181 users to community 3 for a total membership of 634.

3.4.6 Discussion

The purpose of this framework is to propose a method that helps to identify social media follower segments (i.e., communities). Each community detection method resulted in a different number of communities. Non-overlapping community detection resulted in five communities, overlapping community detection resulted in six communities, and overlapping community detection with node attributes resulted in three. Each method of detection removed some restrictions. At first, I removed the restriction of being forced into only one community and then the restriction of only using network connections to detect communities.

Even though some of the applied community detection methods do not use topics, I investigate whether users already cluster around similar topics just by being connected and whether thematic groups are apparent. Therefore, I examine the topics assigned to each user for each community. To further investigate the topics, I focus on topics relevant to the university, leaving 22 topics.

When examining the five non-overlapping communities, I observe a range in how many people discuss a particular topic (see Figures 3-7 to 3-11). Some users' interests might be ignored by forcing users into communities because they are only allowed one community membership.

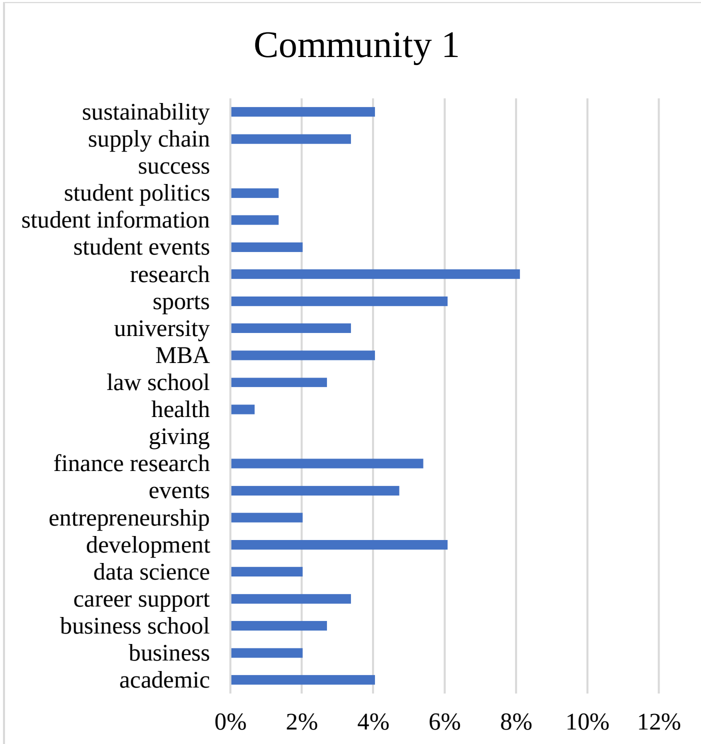


Figure 3-7: Topic Distribution by Non-Overlapping Community (Community 1)

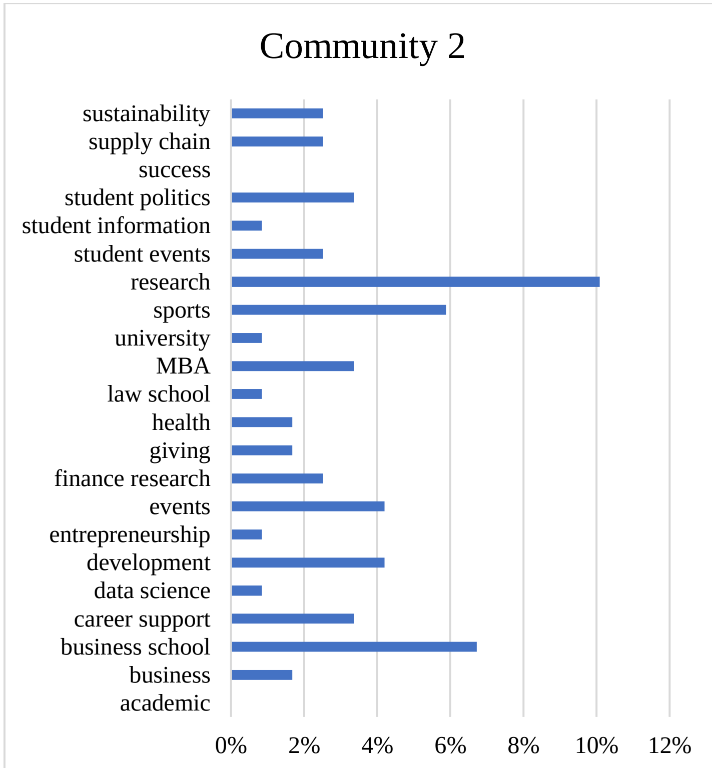


Figure 3-8: Topic Distribution by Non-Overlapping Community (Community 2)

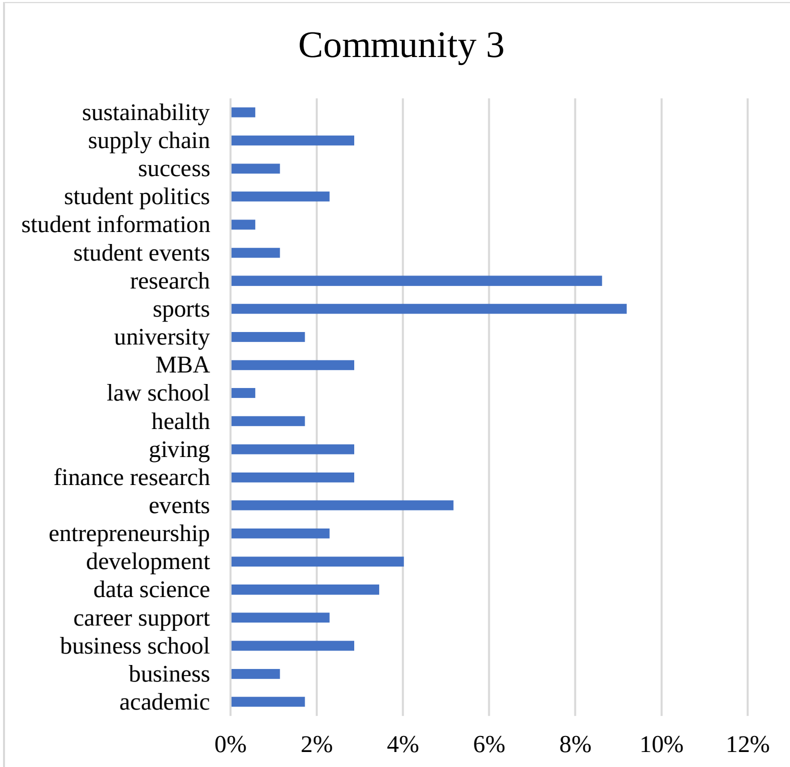


Figure 3-9: Topic Distribution by Non-Overlapping Community (Community 3)

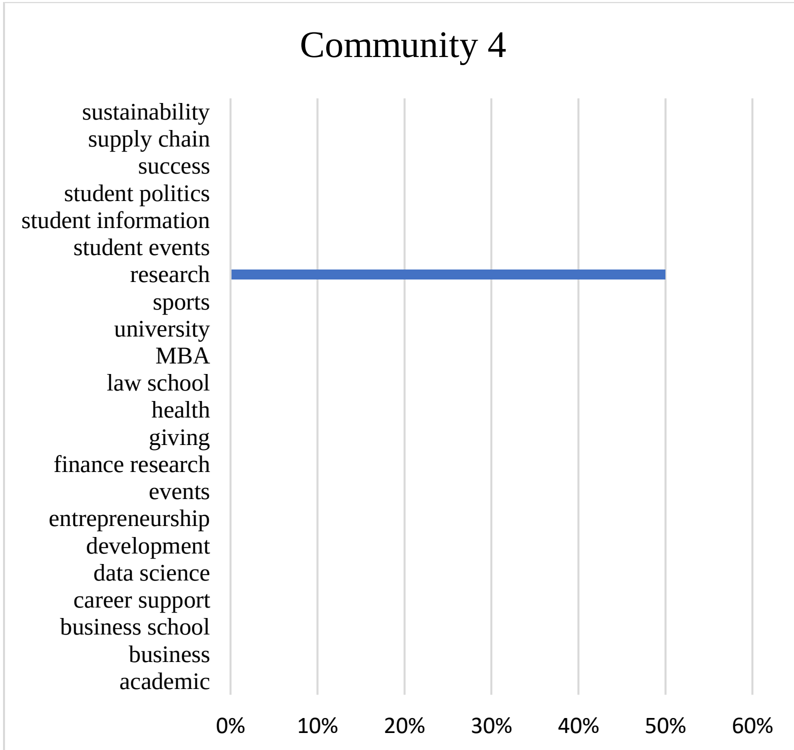


Figure 3-10: Topic Distribution by Non-Overlapping Community (Community 4)

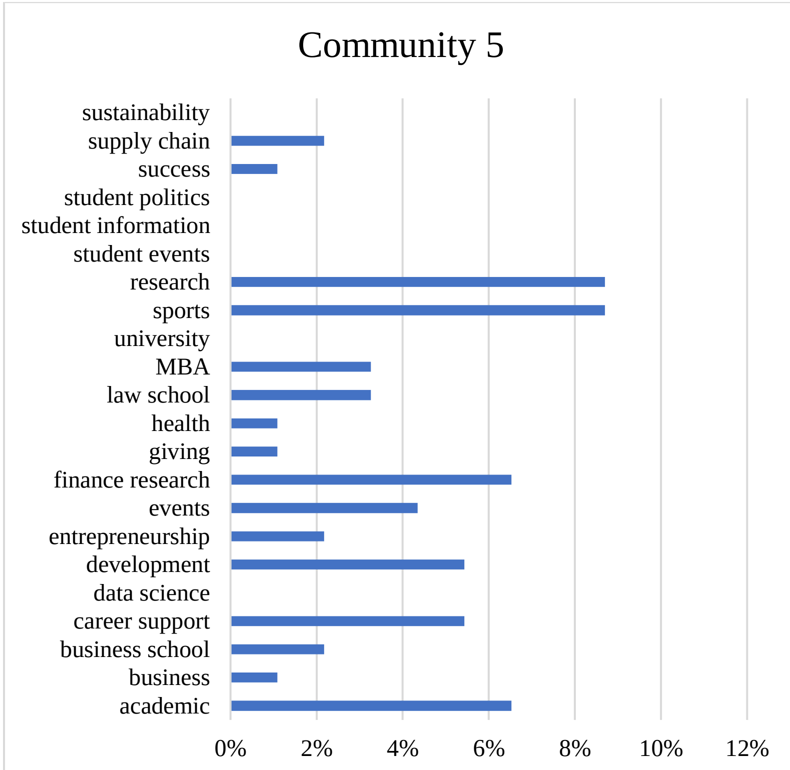


Figure 3-11: Topic Distribution by Non-Overlapping Community (Community 5)

When examining the six overlapping communities, I see that the variation in business school topics decreases compared to the non-overlapping community detection because users are now allowed to be members of multiple communities (see Table 3-8). Therefore, all overlapping communities have a higher alignment on topics they seem to care about because users can be part of multiple communities (see Figures 3-12 to 3-17).

Table 3-8: Topic Variance Change for each Community Detection Method

Community Type	Number of Topics with Variance Decrease	Number of Topics with Variance Increase
Distinct → Overlapping	22	-
Distinct → Overlapping with Attributes	19	3
Overlapping → Overlapping with Attributes	10	12

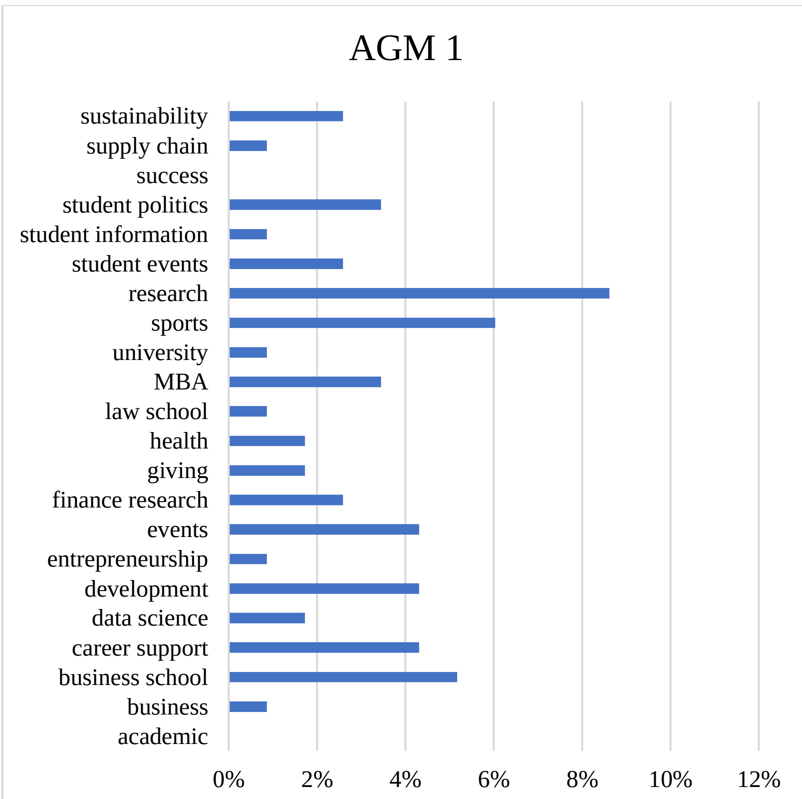


Figure 3-12: Topic Distribution by Overlapping Community (AGM Community 1)

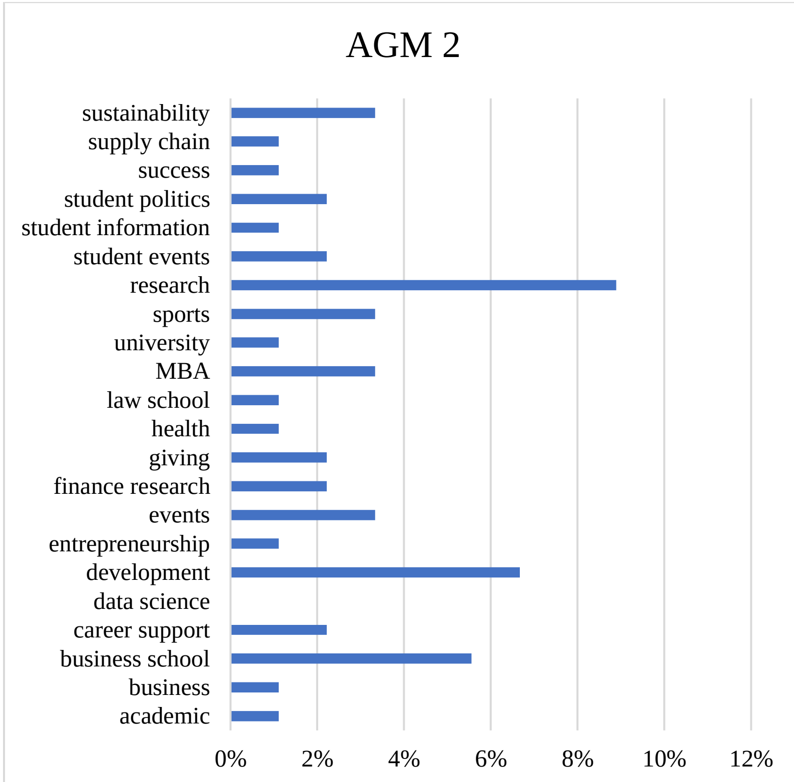


Figure 3-13: Topic Distribution by Overlapping Community (AGM Community 2)

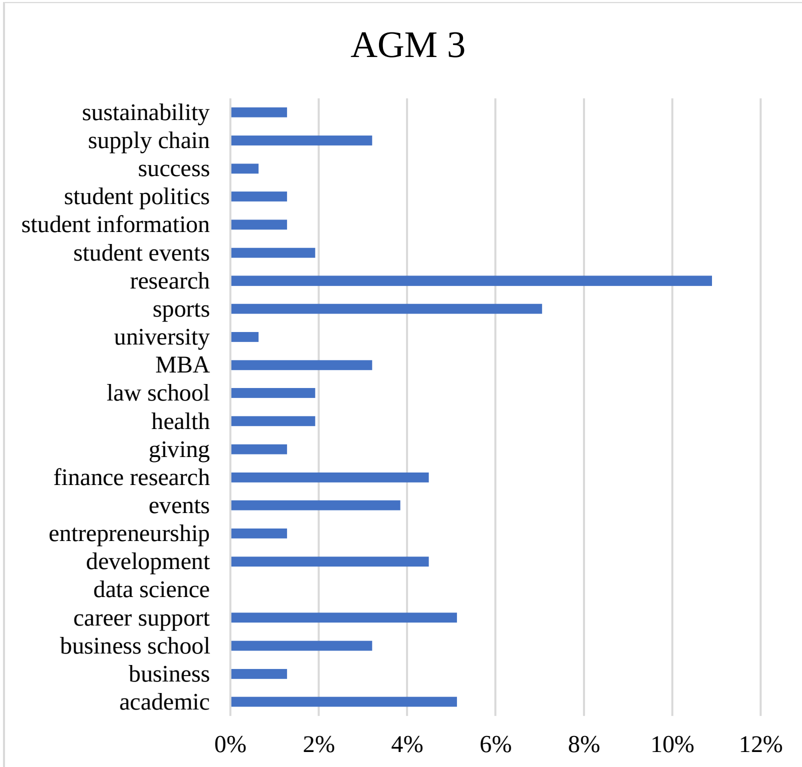


Figure 3-14: Topic Distribution by Overlapping Community (AGM Community 3)

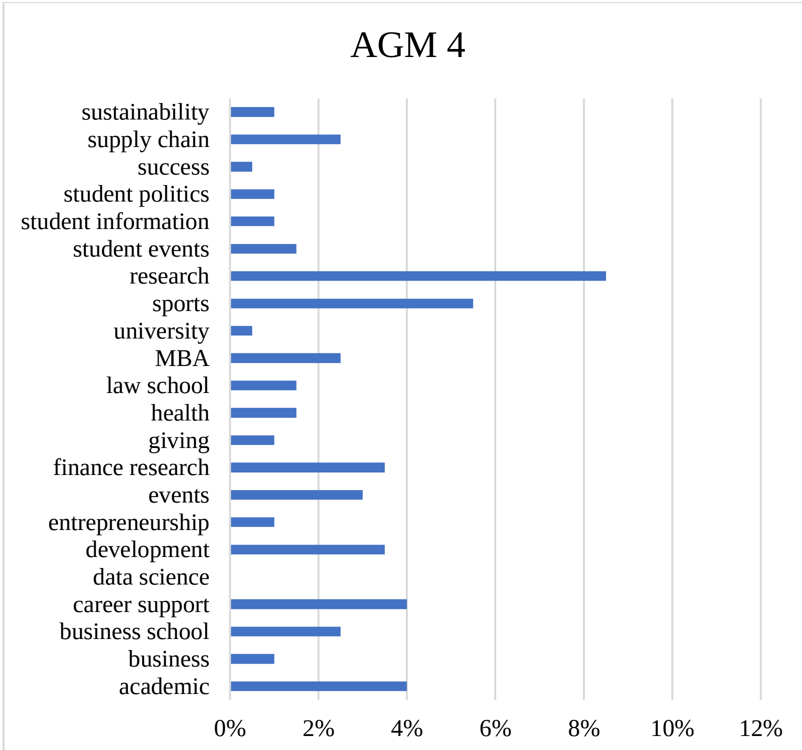


Figure 3-15: Topic Distribution by Overlapping Community (AGM Community 4)

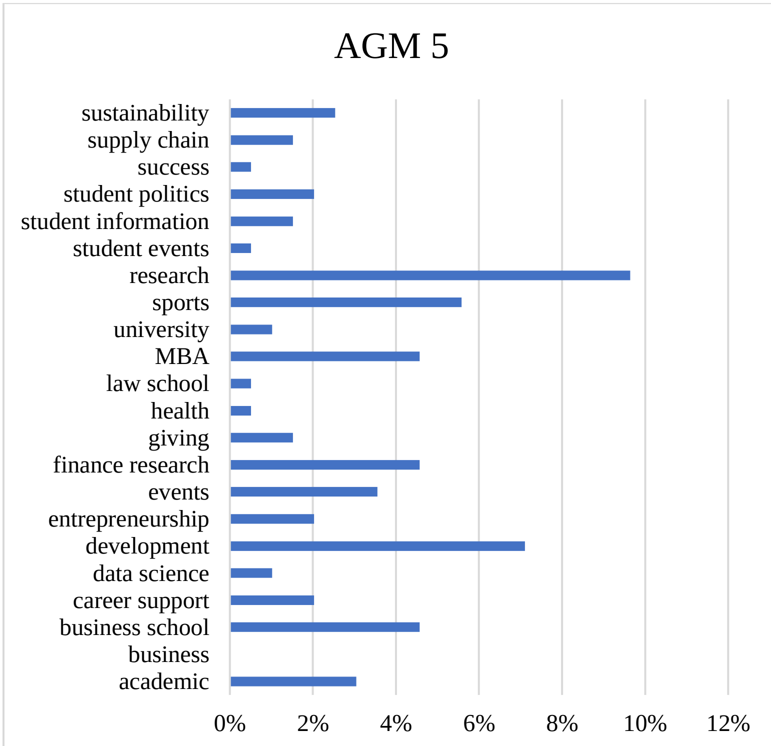


Figure 3-16: Topic Distribution by Overlapping Community (AGM Community 5)

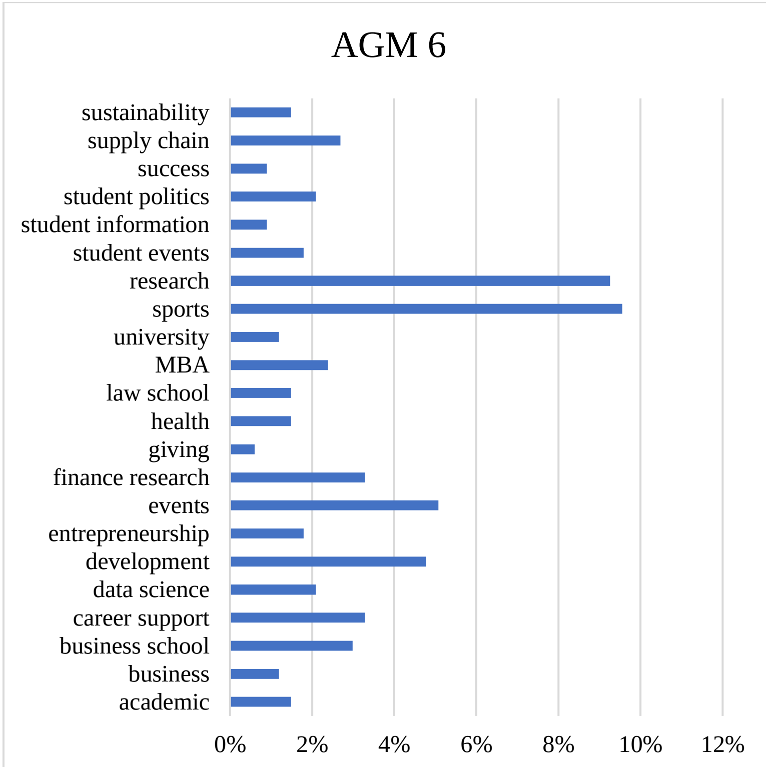


Figure 3-17: Topic Distribution by Overlapping Community (AGM Community 6)

Table 3-9: Topics by Community

	academic	business	business school	career support	data science	develop-ment	entrepre-neurship	events	finance research	giving	health
Non-overlapping	7%	2%	7%	5%	3%	6%	2%	5%	2%	5%	7%
AGM	5%	1%	3%	3%	2%	4%	1%	2%	1%	2%	2%
CESNA	4%	1%	1%	5%	2%	3%	2%	3%	2%	3%	1%
Non-AGM	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Non-CESNA	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
AGM-CESNA	100%	100%	100%	0%	100%	100%	0%	0%	0%	0%	100%

	law school	MBA	university	sports	research	student events	student information	student politics	success	supply chain	sustainability
Non-overlapping	3%	2%	2%	3%	4%	3%	3%	9%	42%	3%	1%
AGM	2%	1%	1%	1%	2%	1%	1%	6%	2%	2%	1%
CESNA	2%	2%	2%	3%	1%	1%	1%	5%	2%	1%	1%
Non-AGM	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Non-CESNA	100%	0%	0%	100%	100%	100%	100%	100%	100%	100%	0%
AGM-CESNA	0%	0%	0%	0%	100%	0%	0%	100%	100%	100%	0%

When I examine the three communities from the overlapping community detection with node attributes, I find that the topic variation decreases for topics that all communities across all methods care about (i.e., broad topics), as shown in Table 3-9. Thus, users in these communities show even fewer differences among broad topics. However, the variation for topics that are either very specific (i.e., niche topics) to the business school or not a typical business school topic increases (see Figures 3-18 to 3-20). I define a topic as a niche topic if it is a topic that is specific or novel to the business school, whereas broad topics can be topics still relevant to the business school but also apply to the whole university. For example, the topics business, business school, supply chain, and research are topics that align very well with a business school. These are niche topics because they are specific to the business school. Topics that are also niche topics in the business school context are topics that are not as expected of a business school (e.g., politics). These differentiations are important to understand. After all, a broadcasting approach might not work well for all niche topics because not everyone values them the same.

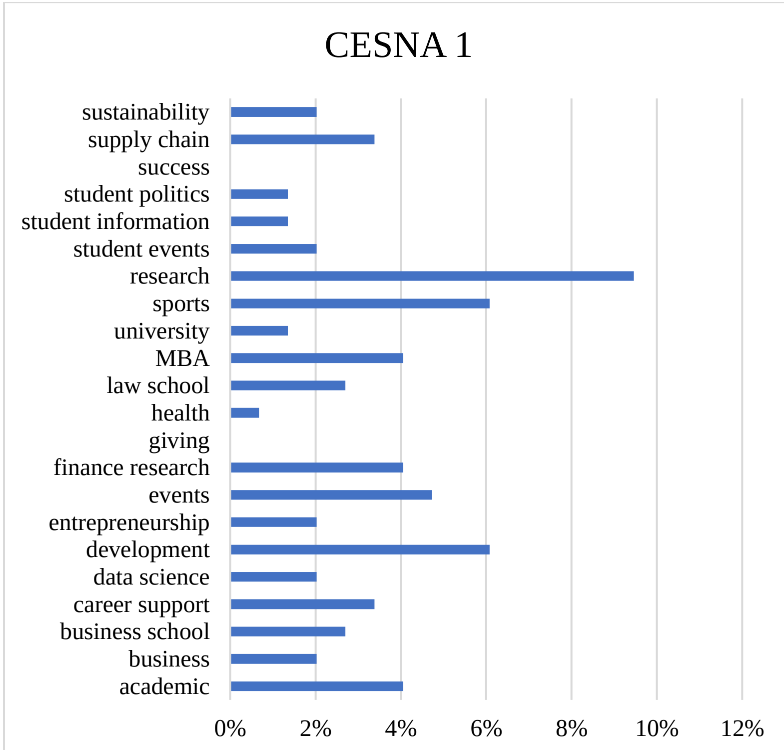


Figure 3-18: Topic Distribution by Overlapping Community w/ Attributes (Community 1)

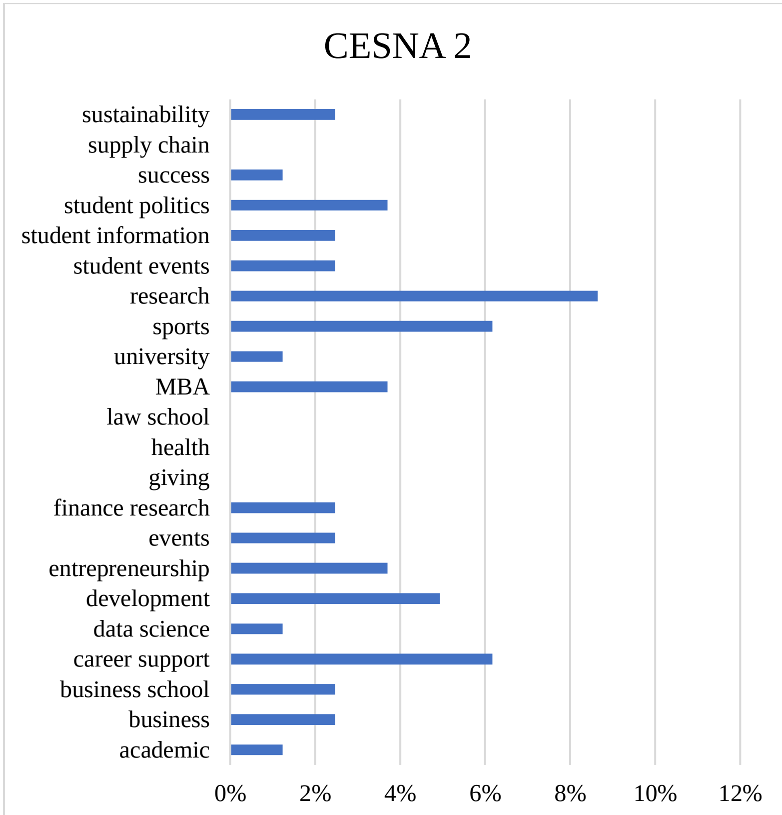


Figure 3-19: Topic Distribution by Overlapping Community w/ Attributes (Community 2)

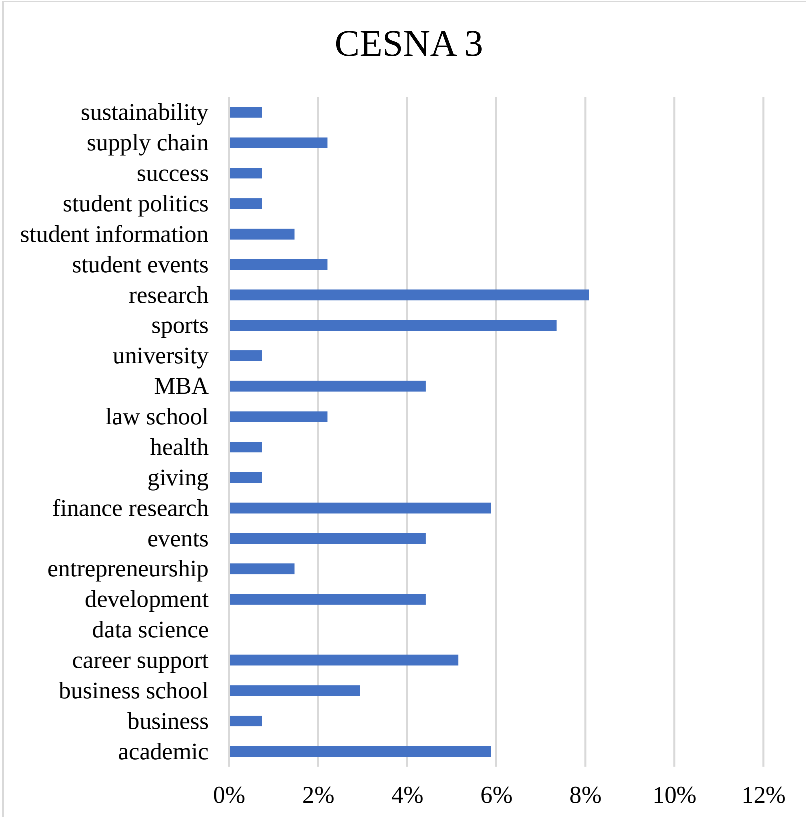


Figure 3-20: Topic Distribution by Overlapping Community w/ Attributes (Community 3)

Based on these observations, users seem to cluster around specific topics of interest. By only using the network connections to detect these communities, you can identify broad or higher-level topics and can use these topics to target many users by broadcasting the same message to a big audience. Only using topic modeling to examine topics that users post about would leave out the existing community structure among social media followers and followers that do not engage frequently. The analysis of network connections should be combined with additional attributes to detect communities and understand which niche topics users cluster around, as depicted in Figure 3-21. Different detection methods provide different insights and can be helpful for different social media strategies of organizations.

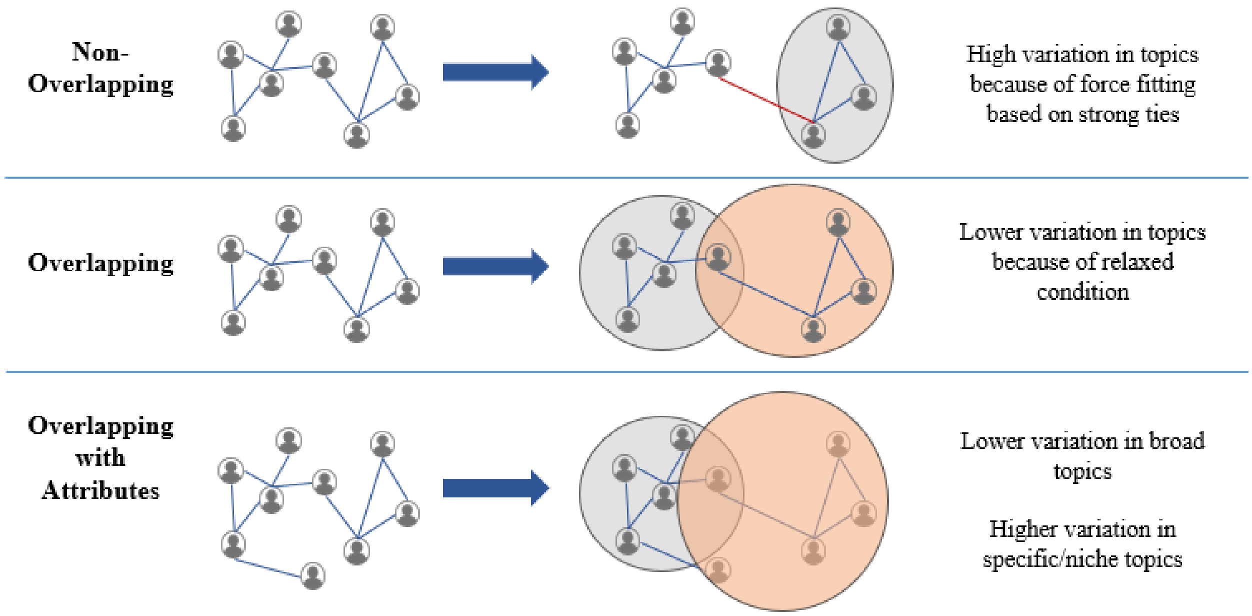


Figure 3-21: Community Development

3.5 Implications

My analysis shows that taking an organization's ego network and detecting follower communities can give insights into the various follower segments. I show the benefits and shortcomings of various methods by showing different steps of using SNA to detect follower communities. The community detection method that best meets the needs of identifying follower communities with a common thematic is the CESNA model. Knowing what segments followers belong to can help organizations target their follower base through improved social media communication.

This essay introduces a framework for organizations to better segment and target their online social media followers. I describe an evolution of four different methods that provide insight into followers. Past social network research has mainly focused on identifying influential individuals for targeting. This research framework incorporates social network connections and user interests based on topic modeling to segment and target groups of followers.

3.5.1 Theoretical Implications

The framework provides a different perspective for SNA in marketing by highlighting and applying the benefits of using an ego network approach. With many privacy regulations on social media, obtaining full network data is challenging. Additionally, full network data might not always be as meaningful because direct connections are the ones that are likely to engage with the account.

Within this framework, I overcome the constraints of several SNA methods in a step-by-step approach. Not allowing for overlapping communities does not realistically depict communities because followers are likely to follow an account for more than one reason. When relaxing the condition of distinct community membership, the topic variance smooths out because users are allowed to be part of multiple communities. Thus, community membership should account for multiple topics of interest. When detecting communities solely based on social network ties, similarities among followers who are not yet connected are ignored.

Only applying topic modeling for follower segmentation gives an overview of what followers are interested in, but it leaves out users that do not frequently engage on social media. In this data, only 12% of followers engaged within the observed time frame. Thus, most followers would be ignored even though they show an interest in the organization by following it. Adding network ties can help identify similar interest groups because people often cluster around similar interests (i.e., homophily) (McPherson, Smith-Lovin, and Cook 2001).

The combination of network analysis and node attributes uncovers communities based on network connections and similar interests. More specifically, overlapping community detection with node attributes helps to identify broad topics that all followers are interested in and differentiate between broad and niche topics.

3.5.2 Managerial Implications

With 3.8 billion social media users worldwide, organizations should leverage the vast online social media data available. Follower connections can give many insights into how they communicate and their preferences. In practice, most organization do not utilize their social network data and do not analyze social network structures.

Organizations can detect follower segments based on network connections and attributes by assigning attributes for each user. Attributes can aid with community detection if some users are not connected. Additionally, while it becomes more difficult for organizations to reach their followers because of frequently changing algorithms on social media, understanding the follower segments becomes even more important. My framework helps organizations better understand their follower segments, how they cluster together, and what specific interests users have. Subsequently, organizations can craft specific content for their defined follower segments. Based on the communities' preferred content, organizations can understand what segment to target and what content to post according to the segment's interest. Aligning content with the follower segments is one thing organizations have control over when posting in contrast to a platform's algorithmic feed creation. Instead of broadcasting messages to everyone without regard to user preferences, organizations can use thematic community structures to post community-specific content.

Overall, the discussed community detection methods give firms a novel way to analyze data and understand diverse groups of potential customers. This research framework provides a toolkit for all types of organizations across industries to better segment and target their social media followers. Based on the data available to organizations, they can choose the most applicable method for their data. However, ideally, organizations should gather additional data besides the network to employ node attributes to segment followers based on network connections and attributes.

3.6 Limitations and Future Research

One limitation of my proposed research framework is that the CESNA model requires a manual community assignment for followers without many connections. Another limitation is that I only had access to one type of attribute (i.e., topics from STM). Including an additional attribute would provide even more data on followers and help with assigning followers to communities. Future research should investigate additional attributes that can be used for community detection and segmenting of social media followers.

Furthermore, there are limitations to evaluating and validating community detection algorithms. Ideally, community detection results should be compared with ground truth data to evaluate whether users value similar things and follow an account for similar reasons (Tang and Liu 2010). Another limitation is that I only have data from one organization with few engagements. Therefore, other research should test the community detection methods on other organizations' online social media networks and link the identified communities and their preferences to follower engagement. It would also be interesting to see whether for-profit organizations show the same community structures in their online ego networks or if the community structure I find is specific to organizations like universities.

More specifically, future research should design an experiment to test the validity of the community detection methods by comparing engagements from a broadcasting approach and a targeted approach based on identified communities from the network and attributes. A critical requirement of engagement measures is that an account must have sufficient engagement to test and validate different content alignments.

Chapter 4

CONCLUSION

Every firm has different stakeholders in its business environment, and the focuses on stakeholder groups vary. One important thing about stakeholders is that they all provide value for firms and understanding how these stakeholders are connected within their group generates valuable insights. In the two essays, I highlight stakeholder networks' impact on firm performance and the information organizations can extract.

The first essay analyzes an employee network of B2B salespeople to examine the effect that different categories of intellectual capital have on selling types and performance. I differentiate this network analysis from previous network research in marketing by utilizing secondary data (i.e., bipartite network) to understand network ties between salespeople. My results show that different sources of knowledge (i.e., intellectual capital) impact different B2B selling types. Human capital is more beneficial for rebuying, social capital is more helpful for cross-selling, and having low levels of human and social capital makes salespeople more likely to spend effort on acquisitions. Additionally, I perform a robustness check on whether the network formation among salespeople is endogenous by applying a SAOM. My findings provide evidence that B2B salespeople do not form network ties endogenously.

The second essay analyzes a social media ego network of followers. In this research, I provide an example and a shift of focus from a full network to an ego network, which is a more feasible approach for firms. I examine community networks, whether there is value in segmenting followers, and which method is suitable for this type of data. More specifically, I apply different social network analysis methods and relax the constraints of each additional method's previous methods. My findings suggest that a combination of network analysis and topic modeling can give the most specific understanding of follower communities. Focusing on network connections alone disregards users that share similar interests but have not had a chance to connect. Topic modeling alone ignores users that do not frequently engage on social media, which is the majority of users.

This dissertation highlights the value of two specific stakeholder networks for firms and addresses some shortcomings of previous social networks research. In the following section, I discuss other potential research directions.

4.1 Discussion of Future Stakeholder Network Research

As a quote from Mark (Newman 2018) states: “The ultimate goal of studying networks is to better understand the behaviors of the systems they represent.” Within a firm’s business environment, stakeholders are part of its system; therefore, understanding stakeholder groups' behaviors will let firms adapt their strategies and better meet stakeholder needs.

Because stakeholder groups are often also connected, there is potential to research the network within one stakeholder group and across stakeholder groups. Some marketing research is already going in this direction. Gupta et al. (2019) analyze inter- and intra-firm network ties and the effect these relationships have on key account management profitability. The authors examine within-seller ties by utilizing network connections at the seller firm and seller-buyer connections by utilizing connections the seller team has with the buying team at a firm. Future research should continue analyzing network connections across stakeholders.

Additionally, the pandemic has accelerated many trends in business, and one of these trends is virtual selling in B2B. More and more customers prefer a virtual sales interaction over an in-person interaction, and this selling practice requires more insights into what type of relationships benefit from less personal interaction.[9] Future research should utilize networks of salespeople and customers to study the interactions that occur virtually and support them appropriately.

Moreover, marketing research should analyze other stakeholder networks that have not been given as much attention. For example, with shifting business goals toward a more sustainable future, supplier relationships with each other and other firms can give insights into how quickly change can take place. How do new business goals change supplier and distribution networks and the relationship between firms and suppliers? One specific example is that Patagonia is in the process of changing its suppliers based on environmental impact. What does that mean for existing relationships?[10] Will other suppliers assimilate?

It would also be interesting to understand whether suppliers in a network would change their manufacturing practices based on what competitors are doing, what important customers are doing, or what other downstream suppliers are doing. Having a better grasp of network connections among suppliers can help anticipate reactions or disruptions in the supply chain.

These are just a few examples of future directions for stakeholder network research. As business environments become more connected, researchers may also find new ways to collect and integrate network data into their analyses.

[9] https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/these-eight-charts-show-how-covid-19-has-changed-b2b-sales-forever

[10] https://www.nytimes.com/2021/12/10/business/ryan-gellert-patagonia-corner-office.html

BIBLIOGRAPHY

Ahearne, Michael, Eli Jones, Adam Rapp, and John Mathieu (2008), "High touch through high tech: The impact of salesperson technology usage on sales performance via mediating mechanisms," *Management Science*, 54 (4), 671–85.

Ahearne, Michael, Son K. Lam, Babak Hayati, and Florian Kraus (2013), "Intrafunctional Competitive Intelligence and Sales Performance: A Social Network Perspective," *Journal of Marketing*, 77 (5), 37–56.

Arnold, Todd J., Eric (Er) Fang, and Robert W. Palmatier (2011), "The effects of customer acquisition and retention orientations on a firm's radical and incremental innovation performance," *Journal of the Academy of Marketing Science*, 39 (2), 234–51.

Bakshy, Eytan, Jake Hofman, Winter Mason, and Duncan Watts (2011), "Everyone's an Influencer: Quantifying Influence on Twitter," *Proceedings of the fourth ACM international conference on Web search and data mining SE - WSDM '11*, 65–74.

Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), "Uniting the Tribes: Using Text for Marketing Insight," *Journal of Marketing*, 84 (1), 1–25.

Blondel, Vincent D., Jean Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre (2008), "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, 2008 (10), 1–12.

Bolander, Willy, Cinthia B. Satornino, Douglas E. Hughes, and Gerald R. Ferris (2015), "Social Networks within Sales Organizations: Their Development and Importance for Salesperson Performance," *Journal of Marketing*, 79 (6), 1–16.

Bommaraju, Raghu and Sebastian Hohenberg (2018), "Self-selected sales incentives: Evidence of their effectiveness, persistence, durability, and underlying mechanisms," *Journal of Marketing*, 82 (5), 106–24.

Borda, Monica (2011), *Fundamentals in Information Theory and Coding*, Springer Science & Business Media.

Borgatti, Stephen P., Martin G. Everett, and Jeffrey C. Johnson (2018), *Analyzing Social Networks*, Sage Publications.

Borgatti, Stephen P, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca (2009), "Corrected 24 April 2009; see last page," *Science*, 323 (April), 892–96.

Burt, Ron S. (1992), *Structural Holes: The Social Structure of Competition*, Cambridge, MA: Harvard University Press.

Centola, Damon and Michael Macy (2007), "Complex Contagions and the Weakness of Long Ties," *American Journal of Sociology*, 113 (3), 702–34.

Cespedes, Frank V. (2014), "Sales Organization: Choose Your Strengths and Limitations," in *Aligning Strategy and Sales: The Choices, Systems, and Behaviors That Drive Effective Selling*, Boston, MA: Harvard Business Review Press, 152–53.

Che, Hai, Xinlei Chen, and Chen Yuxin (2012), "Investigating Effects of Out-of-Stock on Consumer Stockkeeping Unit Choice," *Journal of Marketing Research*, 49 (4), 502–13.

Coleman, James S (1988), "Social Capital in the Creation of Human Capital," *American Journal of Sociology*, 94, S95–120.

Coughlan, Anne T. and Joseph Kissan (2012), "Sales Force Compensation: Research Insights and Research Potential," in *Handbook of Business-to-Business Marketing*, G. L. Lilien and R. Grewal, eds., Northampton, MA: Edward Elgar Publishing, 473–95.

Danaher, Peter J., Michael S. Smith, Kulan Ranasinghe, and Tracey S. Danaher (2015), "Where, When, and How Long: Factors that Influence the Redemption of Mobile Phone Coupons," *Journal of Marketing Research*, 52 (5), 710–25.

Dyer, Jeffrey H. (1996), "Specialized supplier networks as a source of competitive advantage: Evidence from the auto industry," *Strategic Management Journal*, 17 (4), 271–91.

Ebbes, Peter, Zan Huang, and Arvind Rangaswamy (2016), "Sampling designs for recovering local and global characteristics of social networks," *International Journal of Research in Marketing*, 33 (3), 578–99.

Everett, Martin G. (2016), "Centrality and the dual-projection approach for two-mode social network data," *Methodological Innovations*, 9, 1–8.

Everett, Martin G., Chiara Broccatelli, Stephen P. Borgatti, and Johan Koskinen (2018), "Measuring knowledge and experience in two mode temporal networks," *Social Networks*, 55, 63–73.

Franke, George R. and Jeong Eun Park (2006), "Salesperson adaptive selling behavior and customer orientation: A meta-analysis," *Journal of Marketing Research*, 43 (4), 693–702.

Freeman, Linton C. (1977), "A Set of Measures of Centrality Based on Betweenness," *Sociometry*, 40 (1), 35–41.

Freeman, Linton C. (1979), "Centrality in social networks: Conceptual clarification," *Social Networks*, 1 (3), 215–39.

Freeman, R. Edward (2010), *Strategic Management: A Stakeholder Approach*, Cambridge: Cambridge University Press.

Gill, Manpreet, Shrihari Sridhar, and Rajdeep Grewal (2017), "Return on engagement initiatives: A study of a business-to-business mobile app," *Journal of Marketing,* 81 (4), 45–66.

Godes, David and Dina Mayzlin (2004), "Using online conversations to study word-of-mouth communication," *Marketing Science,* 23 (4).

Goldenberg, Jacob, Barak Libai, and Eitan Muller (2001), "Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth," *Marketing Letters,* 12 (3), 211–23.

Gonzalez, Gabriel R., Danny P. Claro, and Robert W. Palmatier (2014), "Synergistic effects of relationship managers' social networks on sales performance," *Journal of Marketing,* 78 (1), 76–94.

Granovetter, Mark (1992), "Problems of Explanation in Economic Sociology," in *Networks and Organizations: Structure, Form, and Action,* N. Nohria and R. G. Eccles, eds., Boston: Harvard Business School Press, 25–56.

Grassi, R., F. Calderoni, M. Bianchi, and A. Torriero (2019), "Betweenness to assess leaders in criminal networks: New evidence using the dual projection approach," *Social Networks,* 56, 23–32.

Grewal, Rajdeep and Shrihari Sridhar (2021), "Commentary: Toward Formalizing Social Influence Structures in Business-to-Business Customer Journeys," *Journal of Marketing*, 85 (1), 98–102.

Gupta, Aditya, Alok Kumar, Rajdeep Grewal, and Gary L. Lilien (2019), "Within-Seller and Buyer–Seller Network Structures and Key Account Profitability," *Journal of Marketing*, 83 (1), 108–32.

Hajli, Nick, Mohana Shanmugam, Savvas Papagiannidis, Debra Zahay, and Marie Odile Richard (2017), "Branding co-creation with members of online brand communities," *Journal of Business Research*, 70, 136–44.

Hakansson, Hakan and Ivan Snehota (1995), *Developing Relationships in Business Networks*, London: Routledge.

Hill, Shawndra, Foster Provost, and Chris Volinsky (2006), "Network-based marketing: Identifying likely adopters via consumer networks," in *Statistical Science,* 256–76.

Hinz, Oliver, Bernd Skiera, Christian Barrot, and Jan U. Becker (2011), "Seeding strategies for viral marketing: An empirical comparison," *Journal of Marketing*, 75 (6), 55–71.

Homburg, Christian, Michael Müller, and Martin Klarmann (2011), "When Should the Customer Really be King? On the Optimum Level of Salesperson Customer Orientation in Sales Encounters," *Journal of Marketing*, 75 (2), 55–74.

Hughes, Arthur M. (1996), *The Complete Database Marketer*, New York: McGraw-Hill.

Hughes, Douglas E., Joël Le Bon, and Adam Rapp (2013), "Gaining and leveraging customer-based competitive intelligence: The pivotal role of social capital and salesperson adaptive selling skills," *Journal of the Academy of Marketing Science*, 41 (1), 91–110.

Iyengar, Raghuram, Christophe van den Bulte, and Jae Young Lee (2015), "Social contagion in new product trial and repeat," *Marketing Science,* 34 (3), 408–29.

Kamakura, Wagner A. (2008), "Cross-Selling: Offering the Right Product to the Right Customer at the Right Time," *Journal of Relationship Marketing*, 6 (4), 41–58.

Kamakura, Wagner A., Bruce S. Kossar, and Michel Wedel (2004), "Identifying innovators for the cross-selling of new products," *Management Science*, 50 (8), 1120–33.

Kamakura, Wagner A., Michel Wedel, Fernando de Rosa, and Jose Afonso Mazzon (2003), "Cross-selling through database marketing: A mixed data factor analyzer for data augmentation and prediction," *International Journal of Research in Marketing,* 20 (1), 45–65.

Kamakura, Wagner, Carl F. Mela, Asim Ansari, Anand Bodapati, Pete Fader, Raghuram Iyengar, Prasad Naik, Scott Neslin, Baohong Sun, Peter C. Verhoef, Michel Wedel, and Ron Wilcox (2005), "Choice models and customer relationship management," *Marketing Letters,* 16 (3–4), 279–91.

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 (3), 425–43.

Kishore, Sunil, Raghunath Singh Rao, Om Narasimhan, and George John (2013), "Bonuses versus commissions: A field study," *Journal of Marketing Research*, 50 (3), 317–33.

Koskinen, Johan and Christofer Edling (2012), "Modelling the evolution of a bipartite network—Peer referral in interlocking directorates," *Social Networks*, 34 (3), 309–22.

Krackhardt, David and Jeffrey R. Hanson (1993), "Informal Networks: The Company Behind the Chart," *Harvard Business Review*.

Kumar, V., Morris George, and Joseph Pancras (2008), "Cross-buying in retailing: Drivers and consequences," *Journal of Retailing*, 84 (1), 15–27.

Kumar, V. and Anita Pansari (2016), "Competitive advantage through engagement," *Journal of Marketing Research*, 53 (4), 497–514.

Kumar, V., Rajkumar Venkatesan, and Werner Reinartz (2008), "Performance implications of adopting a customer-focused sales campaign," *Journal of Marketing*, 72 (5), 50–68.

Lim, Noah, Michael J. Ahearne, and Sung H. Ham (2009), "Designing sales contests: Does the prize structure matter," *Journal of Marketing Research*, 46 (3), 356–71.

Lin, Nan (2001), *Social Capital: A Theory of Social Structure and Action*, Cambridge, UK; New York: Cambridge University Press.

Madhavaram, Sreedhar and Shelby D. Hunt (2017), "Customizing business-to-business (B2B) professional services: The role of intellectual capital and internal social capital," *Journal of Business Research*, 74, 38–46.

Malms, Oliver and Christian Schmitz (2011), "Cross-divisional orientation: Antecedents and effects on cross-selling success," *Journal of Business-to-Business Marketing*, 18 (3), 253–75.

Manger, Mark S. and Mark A. Pickup (2016), "The Coevolution of Trade Agreement Networks and Democracy," *Journal of Conflict Resolution*, 60 (1), 164–91.

Mariadoss, Babu John, Chad Milewicz, Sangwon Lee, and Arvin Sahaym (2014), "Salesperson competitive intelligence and performance: The role of product knowledge and sales force automation usage," *Industrial Marketing Management*, 43 (1), 136–45.

Martínez-Seis, Bella (2017), "RELNA: Ranking attributes in social networks to detect overlapping communities efficiently," in *Proceedings - International Conference on Data Engineering*, IEEE Computer Society, 1431–35.

Matsuo, Makoto and Takashi Kusumi (2002), "Salesperson's procedural knowledge, experience and performance: An empirical study in Japan," *European Journal of Marketing*, 36 (7/8), 840–54.

Maurer, Indre, Vera Bartsch, and Mark Ebers (2011), "The value of intra-organizational social capital: How it fosters knowledge transfer, innovation performance, and growth," *Organization Studies*, 32 (2), 157–85.

McKay, Bob (2017), "The Six Most Effective Social Media Metrics To Understand Your Campaign's Success," (accessed May 2, 2018), [available at https://www.forbes.com/sites/forbesagencycouncil/2017/07/14/the-six-most-effective-social-media-metrics-to-understand-your-campaigns-success/2/#108b9a1a7ac8].

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, 27, 415–44.

Meire, Matthijs, Kelly Hewett, Michel Ballings, V. Kumar, and Dirk Van den Poel (2019), "The Role of Marketer-Generated Content in Customer Engagement Marketing," *Journal of Marketing*, 83 (6), 21–42.

Min, Sungwook, Xubing Zhang, Namwoon Kim, and Rajendra K. Srivastava (2016), “Customer acquisition and retention spending: An analytical model and empirical investigation in wireless telecommunications markets,” *Journal of Marketing Research*, 53 (5), 728–44.

Musalem, Andrés and Yogesh V. Joshi (2009), “How much should you invest in each customer relationship? A competitive strategic approach,” *Marketing Science*, 28 (3), 555–65.

Nahapiet, Janine and Sumantra Ghoshal (1998), “Social Capital, Intellectual Capital, and the Organizational Advantage,” *The Academy of Management Review*, 23 (2), 242–66.

Netzer, Oded, Ronen Feldman, Jacob Goldenberg, and Moshe Fresko (2012), “Mine your own business: Market-structure surveillance through text mining,” *Marketing Science*, 31 (3), 521–43.

Newberry, Christina (2020), “How to Find and Target Your Social Media Audience,” *Blog Hootsuite*, [available at https://blog.hootsuite.com/target-market/].

Newman, M. E. J. (2006), “Modularity and community structure in networks,” *Proceedings of the National Academy of Sciences*, 103 (23), 8577–82.

Newman, M. E. J. (2018), *Networks*, United Kindom: Oxford University Press.

Newman, M. E.J. and M. Girvan (2004), “Finding and evaluating community structure in networks,” *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69 (2 2).

Petersen, J. Andrew and Franziska Schmid (2021), “Leveraging stakeholder networks with outside-in marketing,” *Industrial Marketing Management*, 92, 72–75.

Petrin, Amil and Kenneth Train (2010), “A Control Function Approach to Endogeneity in Consumer Choice Models,” *Journal of Marketing Research*, 47 (1), 3–13.

Rapp, Adam, Michael Ahearne, John Mathieu, and Niels Schillewaert (2006), “The impact of knowledge and empowerment on working smart and working hard: The moderating role of experience,” *International Journal of Research in Marketing*, 23 (3), 279–93.

Reinartz, Werner J. and V. Kumar (2003), “The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration,” *Journal of Marketing*, 67 (1), 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.

Ringel, Daniel M. and Bernd Skiera (2016), “Visualizing asymmetric competition among more than 1,000 products using big search data,” *Marketing Science*, 35 (3), 511–34.

Ripley, Ruth M., Tom A.B. Snijders, Zsofa Boda, Andras Voros, and Paulina Preciado (2021), “Manual for SIENA version 4.0 (version May 11, 2021),” Oxford: University of Oxford, Department of Statistics; Nuffield College.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley (2019), "Stm: An R Package for Structural Topic Models," *Journal of Statistical Software*, 91 (2).

Rossi, Peter E, Robert E Mcculloch, Greg M Allenby, Peter E Rossi, Robert E Mcculloch, and Greg M Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15 (4), 321–40.

Schmitz, Christian (2013), "Group influences of selling teams on industrial salespeople's cross-selling behavior," *Journal of the Academy of Marketing Science*, 41 (1), 55–72.

Schmitz, Christian, You-Cheong Lee, and Gary L. Lilien (2014), "Cross-Selling Performance in Complex Selling Contexts: An Examination of Supervisory- and Compensation-Based Controls," *Journal of Marketing*, 78 (3), 1–19.

Schultz, Theodore W. (1961), "Investment in Human Capital," *The American Economic Review*, 51 (1), 1–17.

Snijders, Tom, Christian Steglich, and Michael Schweinberger (2007), "Modeling the Coevolution of Networks and Behavior," in *Longitudinal Models in the Behavioral and Related Sciences*, K. In van Monfort, H. Oud, and A. Satorra, eds., Mahwah, NJ: Lawrence Erlbaum, 41–71.

Steglich, Christian, Tom A.B. Snijders, and Michael Pearson (2010), "Dynamic Networks And Behavior: Separating Selection From Influence," *Sociological Methodology*, 40 (1), 329–93.

Stephen, Andrew T., Peter Pal Zubcsek, and Jacob Goldenberg (2016), "Lower connectivity is better: The effects of network structure on redundancy of ideas and customer innovativeness in interdependent ideation tasks," *Journal of Marketing Research*, 53 (2), 263–79.

Sterner, Glenn and Diane Felmlee (2017), "The Social Networks of Cyberbullying on Twitter," *International Journal of Technoethics*, 8 (2), 1–15.

Stolz, Simon and Christian Schlereth (2021), "Predicting Tie Strength with Ego Network Structures," *Journal of Interactive Marketing*, 54, 40–52.

Subramaniam, Mohan and Mark A. Youndt (2005), "The Influence of Intellectual Capital on the Types of Innovative Capabilities," *The Academy of Management Journal*, 48 (3), 450–63.

Susarla, Anjana, Jeong-ha Oh, Yong Tan, and Yong Tan (2012), "Social Networks and the Diffusion of User-Generated Content: Evidence from YouTube," *Information Systems Research*, 23 (1), 23–42.

Swaminathan, Vanitha and Christine Moorman (2009), "Marketing alliances, firm networks, and firm value creation," *Journal of Marketing*, 73 (5), 52–69.

Tang, Lei and Huan Liu (2010), *Community Detection and Mining in Social Media*, Morgan & Claypool.

Toker-Yildiz, Kamer, Minakshi Trivedi, Jeonghye Choi, and Sue Ryung Chang (2017), "Social interactions and monetary incentives in driving consumer repeat behavior," *Journal of Marketing Research,* 54 (3), 364–80.

Trusov, Michael, Anand V Bodapati, and Randolph E Bucklin (2010), "Determining Influential Users in Internet Social Networks," *Journal of Marketing Research,* 47 (4), 643–58.

Trusov, Michael, William Rand, and Yogesh V. Joshi (2013), "Improving prelaunch diffusion forecasts: Using synthetic networks as simulated priors," *Journal of Marketing Research,* 50 (6), 675–90.

Tsai, Wenpin (2001), "Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance," *Academic of Management Journal,* 44 (5), 996–1004.

Tsai, Wenpin and Sumantra Ghoshal (1998), "Social Capital and Value Creation: The Role of Intrafirm Networks," *The Academy of Management Journal,* 41 (4), 464–76.

Üstüner, Tuba and Dawn Iacobucci (2012), "Does Intraorganizational Network Embeddedness Improve Salespeople's Effectiveness? A Task Contingency Perspective," *Journal of Personal Selling & Sales Management,* 32 (2), 187–205.

Van den Bulte, Christophe (2010), "Opportunities and Challenges in Studying Customer Networks," in *The Connected Customer: The Changing Nature of Consumer and Business,* S. H. K. Wuyts, M. G. Dekimpe, E. Gijsbrechts, and F. G. M. (Rik) Pieters, eds., London: Routledge, 7–35.

Van den Bulte, Christophe and Yogesh v. Joshi (2007), "New Product Diffusion with Influentials and Imitators," *Marketing Science,* 26 (3), 400–421.

Van den Bulte, Christophe and Gary L. Lilien (2001), "Medical Innovation Revisited: Social Contagion versus Marketing Effort," *American Journal of Sociology,* 106 (5), 1409–35.

Venkatesan, Rajkumar (2017), "Executing on a customer engagement strategy," *Journal of the Academy of Marketing Science,* 45 (3), 289–93.

Verbeke, Willem, Bart Dietz, and Ernst Verwaal (2011), "Drivers of sales performance: A contemporary meta-analysis. Have salespeople become knowledge brokers?," *Journal of the Academy of Marketing Science,* 39 (3), 407–28.

Vermeer, Susan A.M., Theo Araujo, Stefan F. Bernritter, and Guda van Noort (2019), "Seeing the wood for the trees: How machine learning can help firms in identifying relevant electronic word-of-mouth in social media," *International Journal of Research in Marketing,* 36 (3), 492–508.

Wang, Rui, Aditya Gupta, and Rajdeep Grewal (2017), "Mobility of top marketing and sales executives in business-To-business markets: A social network perspective," *Journal of Marketing Research*, 54 (4), 650–70.

Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge, UK: Cambridge University Press.

Watts, Duncan J. and Steven H. Strogatz (1998), "Collective dynamics of 'small-world' networks," *Nature*, 393 (6684), 440–42.

Wiseman, Phillip, Michael Ahearne, Zachary Hall, and Seshadri Tirunillai (2022), "Onboarding Salespeople: Socialization Approaches," *Journal of Marketing*, 002224292210764.

Yang, Jaewon and Jure Leskovec (2012a), "Structure and Overlaps of Communities in Networks," *arXiv*.

Yang, Jaewon and Jure Leskovec (2012b), "Defining and Evaluating Network Communities Based on Ground-Truth in Online Social Networks," in *Proceedings of the 2012 IEEE International Conference on Data Mining*.

Yang, Jaewon, Julian McAuley, and Jure Leskovec (2013), "Community detection in networks with node attributes," *Proceedings - IEEE International Conference on Data Mining, ICDM*, 1151–56.

Yang, Zhao, René Algesheimer, and Claudio J. Tessone (2016), "A comparative analysis of community detection algorithms on artificial networks," *Scientific Reports*, 6 (August).

Yoganarasimhan, Hema (2012), "Impact of social network structure on content propagation: A study using YouTube data," *Quantitative Marketing and Economics*, 10 (1), 111–50.

Youndt, Mark A., Mohan Subramaniam, and Scott A. Snell (2004), "Intellectual Capital Profiles: An Examination of Investments and Returns," *Journal of Management Studies*, 41 (2), 335–61.

VITA

FRANZISKA SCHMID

EDUCATION

Penn State Smeal College of Business **University Park, PA**
Ph.D. in Business Administration (Emphasis Marketing) *2017 – 2022*
Minor: Social Data Analytics

Penn State Smeal College of Business **University Park, PA**
Master of Business Administration *2015 - 2017*
Concentrations: Marketing & Strategic Leadership

Dickinson State University **Dickinson, ND**
Bachelor of Science in International Business, Summa cum laude *2008 - 2011*

RESEARCH INTERESTS

Social Networks, CRM, B2B-Sales, Social Media, and Social Influence

PH.D. DISSERTATION AND COMMITTEE

Dissertation Title: "Essays on Stakeholder Networks in Marketing"
Committee Members: J. Andrew Petersen (Chair), Arvind Rangaswamy, Stefan Wuyts, and Bruce Desmarais (Political Science)

Essay 1: "The value of intellectual capital for business-to-business selling"

Essay 2: "Identifying community structures among social media followers: Evidence from an online ego network perspective"