ESSAYS ON NETWORK EFFECTS, SERVICE PRICING PLANS

AND ONLINE MESSAGE VIRALITY

by

Fereshteh Zihagh

APPROVED BY SUPERVISORY COMMITTEE:

____
Brian Ratchford, Chair

____
Xiaolin Li, Co-Chair

____
B.P.S. Murthi

____
Sanjay Jain

To my husband Masoud, my sisters Asal & Neda, and my parents.

ESSAYS ON NETWORK EFFECTS, SERVICE PRICING PLANS

AND ONLINE MESSAGE VIRALITY

by

FERESHTEH ZIHAGH, BS, MBA

DISSERTATION

Presented to the Faculty of

The University of Texas at Dallas

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY IN

MANAGEMENT SCIENCE

THE UNIVERSITY OF TEXAS AT DALLAS

August 2019

ACKNLOWLEDGEMENTS

It would not have been possible to write this dissertation without the help, valuable advice and support of my kind supervisor, Professor Brian Ratchford. Besides my advisor, I would also like to thank my committee members Drs. Xiaolin Li, B.P.S. Murthi, and Sanjay Jain for their encouragement and insightful comments on this dissertation and my other academic activities.

May 2019

ESSAYS ON NETWORK EFFECTS, SERVICE PRICING PLANS AND ONLINE MESSAGE VIRALITY

Fereshteh Zihagh, PhD
The University of Texas at Dallas, 2019

Supervising Professors: Brian Ratchford, Chair
Xiaolin Li, Co-Chair

The first part of my dissertation investigates information technology and examines the matching between buyers and sellers in IT outsourcing markets, and how the social network between firms explains the observed matches. Specifically, I examine how interfirm connections and the information and resources accruing to firms from such connections-i.e., firm social capital - affect firms' decision to sign a contract with a specific firm. The idea is that information spillovers from these connections constrain economic activities and that firms' competitiveness changes by changing their network position. Using a unique panel data of 49,072 contracts signed between buyers and sellers in 1989-2013, I construct the network such that the nodes are the firms which are connected because they collaborate on shared projects. I use the network measures and client-specific characteristics in a two-sided matching model to quantify the change in the created joint value in the match when the firms' network position changes. Results suggest that a network of vendors and clients is more valuable as the size of their network grows (i.e., higher joined social capital is more valued) and that clients with higher social capital derive disproportionally more value from a vendor's network. Moreover, although synergies created from the similarity in

network positions are valuable in most cases, there are exceptions to this which are explained by transaction costs specific to such instances. The findings help firms design their competitive strategy by predicting the effect of repositioning in the network.

In my second chapter, I use a randomized field experiment to investigate customers' reaction to pricing and tariff design in the internet service industry. Specifically, I examine the effect of promotions and tariff structure (3-part tariff) on internet subscriber revenue and churn using a field experiment. Researchers have reported cases in which the impact of pricing decisions on profits is beyond what economic interpretations justify. However, behavioral effects of pricing and tariff structure on post-purchase outcomes are ignored in empirical studies even though they provide useful insight into consumer behavior and can have significant policy implications. I propose a parsimonious model that allows pricing, tariff structure, and new service introduction to impact two relevant behaviors in contractual settings: the level of transactions after plan choice and the decision to churn. My model also accounts for customer heterogeneity by including "level of price-sensitivity" and comparing the behavior of high and low price-sensitive segments. I use data from an Internet Service Provider (ISP) in a natural field experiment setting and find that promotions attract new customers, but they do so at the cost of increasing customers' price sensitivity and lowering their inertia. Moreover, I find that customers who are not exposed to promotions spend more on plan and credit purchases even after I control for possible self-selection of more price-sensitive customers into the group that is exposed to promotions. I also find that high price-sensitive customers respond differently to constant exposure to promotions compared with low price-sensitive segments. I conclude that doing less relevant and targeted promotions wins price-sensitive customers, but it also encourages showrooming and comparison behavior. In other words, a marketing campaign that attracts new customers may also hurt customer lifetime value. I recommend that firms need to consider this trade-off when designing pricing policies.

In the last chapter, I develop an empirical model to study how users' social capital, mediated by image motives (i.e., driven by the perception of others) and intrinsic motives (i.e., driven by personal satisfaction rather than posting consequences), influences the propensity to post/retweet positive or negative contents online. My findings show that the identity of users can explain their motivations to post on online platforms and the receivers' engagement with the posted contents. Results show that there is an inverted U-shaped relationship between the number of followers and motives. The breakdown of motives reveals that both image and intrinsic motives are highest for users with a medium number of followers. Moreover, I find that as the number of followers increases, users are more likely to post due to intrinsic motives than image-related motives. Even though users with a higher number of followers believe negative content can lead to more virality, they do not post negative content mainly because they do not care as much about the reciprocity from followers as they care about the intrinsic satisfaction that they derive from posting contents. The results of my model show that intrinsic motives (in order of importance: validate thoughts, have fun, be a listener, amplify news, entertain/inform, save tweets) are more conducive to posting of positive content. On the other hand, image motives (in order of importance: look clever/expert, identify with a group, gain followers) are associated with the negative content but have no significant association with the positive content. The findings have implications for designing social media content strategy and fostering reader engagement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

STRUCTURAL EMBEDDEDNESS AND BUSINESS PARTNER SELECTION: A NETWORK PERSPECTIVE

1.1 Introduction

As technology becomes key to differentiation, more complex and core business activities are outsourced[1] and outsourcing business processes (e.g., finance, HR) are among the fastest growing trends of outsourcing. However, challenges remain for firms to capture value from their complex outsourced activities with many firms reporting failed matches (Rouse and Corbitt 2006). Vendors' ability to comply with firms' predetermined processes, data security, and capability to offer flexible solutions with timely upgrades are among the top client concerns (CIO.com). How do firms approach these concerns ex-ante and make buying decisions in business-to-business (B2B) contexts? Research on outsourcing performance has largely been dominated by Transaction Cost Economics (Williamson 1985.) TCE explains how the choice of the governance mechanism (e.g., the choice of a partnering firm) and the contract design (e.g., fixed-price/time & materials, flexibility/contingency provisions, termination clauses, monitoring) are dependent on the transaction-specific characteristics and known exchange hazards, namely, asset specificity, transaction uncertainty, and transaction complexity.

---

[1] Global expenditure in IT service outsourcing was $3.3 trillion in 2015, and it is expected to exceed $3.8 trillion starting from 2020 (Gartner, 2016). HR BPO spending in cloud HCM applications will likely reach $4.7B in 2021, with a 19% CAGR from 2016.

TCE is based on two behavioral assumptions of “bounded rationality” and “bounded reliability.” The objective is to align incentives and safeguard against the potential foreseeable contingencies to guarantee the functional or technical aspects of the IT outsourcing development process such that it works as planned. In doing so, TCE acknowledges that the ability to foresee future contingencies reduces transaction costs and improves contracting efficiencies (Williamson 1979). However, since the level of analysis is transactions, it assumes away production costs and horizontal market effects such as knowledge/capability differences between firms. Information differences matter since firms operate in competitive markets and their competitiveness impacts their potential to overcome reliability and uncertainty concerns and will be reflected in the final matching outcome.

A related issue is whereas TCE assumes zero spillover effect−the fact that firms can redeploy the information and assets to their next projects. In reality, firms continuously build up their stock of information through learning by doing and investments, and they are likely to consider spillover effects when deciding to work on projects. Since spillover effects impact governance, they are related to the scope of TCE. Network embeddedness (Granovetter 1985) provides a proper framework to study these information effects. The idea is that local knowledge does not freely circulate in space, and interfirm connections play an important role in this process (Breschi and Lissoni 2001). Vendors and clients are related to each other because the existing information and resources embedded in dyads− social capital−carry over to other dyads that share a client or vendor, through learning by doing or tacit knowledge stocked in organizations (Williamson 1971). Thus, this information will be reflected in the firms’ matching decisions and contract design. Embeddedness implies that business firms are socially and historically constructed; therefore, they are constrained by the structure of the inter-firm relationships that emerge and evolve through continuous interactive processes of working on projects.

Extant TCE research is oblivious to horizontal market factors, such as the presence of rival agents, and structural constraints that firms face, for instance, the knowledge, information, and resources accumulated in a firm with a specific location. The firm governance choice impacts its social capital by adding value or devaluing it and, thus, potentially changing the firm's competitive position. In this chapter, I aim to quantify the firms' preference for matching with certain firms that occupy specific positions in the business network and thus are endowed with different types of information and levels of capabilities.

In sociology, "embeddedness" refers to the centrality and prominence of actors. It is assumed that embeddedness constrains economic activities (Gulati and Gargiulo 1999). However, marketing scholars have generally analyzed business markets at the dyad level as opposed to considering the body of dyads in a social network (Lilien et al. 2010). Srivastava et al. (2001) report that relational assets such as social networks and intellectual assets such as know-how embedded in individuals and processes are market-based resources that are relevant to the field of marketing and can be leveraged to create customer value. Recent marketing scholars have adopted the embeddedness concept and the social capital ensuing from it to study project success in open source systems (Grewal et al. 2006), innovation project development (Moon 2011), film success (Packard 2016), innovation and firm performance (Muller and Peres 2017), the success of key account management (KAM) (Gupta et al. 2017), and managerial social capital of top marketing and sales executives and firm performance (Wang et al. 2017).

In studying which buyers and sellers form matching partners in sourcing markets, my model takes into account the impact of knowledge interdependencies between a client and a vendor in their mutual decision to work together. The argument is that since IT system development is a collaborative process, and firms with better resources are perceived and indeed do better in terms of conceiving and implementing better strategies, client information and experience in IT outsourcing also predict project success. The knowledge-based view (KBV) (Grant 1996, Kogut and Zander 1992) is commonly used to study the impact of firm knowledge on outsourcing decisions. KBV complements TCE by including the requirement that the client business knowledge and vendor technical knowledge must be exchanged and integrated to perform a collaborative IT process (Dibbern et al. 2008, Grant 1996, Tiwana and Bush 2007). This complementary aspect provided by KBV is important in my context since TCE does not address the risks created by the degree to which the knowledge about the client's requirements can be communicated to the vendor, which is a critical consideration since firms equipped with this knowledge are more likely to outsource than vertically integrate.

My approach has notable advantages relative to earlier work on incentives that make certain types of agents sign contract with certain types of principals. Empirical work on contracts typically regresses contract choice and matching outcomes on observed principal and agent characteristics. To date, the literature primarily focuses on vertical characteristics of firms−traits of the focal dyad of analysis−in answering how contracts are awarded and mostly ignores horizontal features of the market, including the presence of rival agents and the competitiveness of them determined by their market level factors such as social capital. I add additional insights in this chapter of my dissertation. First, my method is easily applied to evaluate potential matches in markets in which either side has information opacity regarding the other party. For instance, data on younger firms or geographically distant firms may not be publicly available and ratings for such firms are not readily available, but data on who contracts with whom is not proprietary (Pakes 2010). Second, my structural estimation accounts explicitly for the endogenous matching process by which vendors match with clients, which is an important source of endogeneity in determining which characteristics matter for outsourcing value creation. Ni and Srinivasan (2015) find a non-negligible impact of endogenous matching on estimates of retailer and manufacturer matching preferences in apparel sourcing markets and provide a comparison of reduced-form results and formal matching parametric estimations.

To study firms' behavior, I consider a model in which the "with whom" decision of the outsourcing and the pre-entry competition are addressed simultaneously by applying the interfirm network model embedded in a two-sided matching model. This analysis provides additional insights that are important but missing in the extant research. First, because equilibrium outcomes result from mutual choices and not choice sets, identification cannot be based on the analysis of single-agent demand models (Fox 2010). Choice models such as logit are also not entirely appropriate due to biased preference estimates arising from the omission of horizontal market effects, such as the presence of outside options and competition among economic actors. Second, recent growing research on how knowledge and capabilities complement the traditional focus on transaction costs in managing vertical relationships (Argyres and Zenger 2012) shows that the consideration of knowledge spillovers beyond the focal transaction may significantly affect the governance of the focal transaction. Knowledge spillover is a determinant factor in choosing to use own employees rather than subcontractors (Mayer 2006), and in suppliers' incentives to make partner-specific investments (Kang et al. 2009). As long as firms consider such spillover effects when deciding to work on projects, an empirical model without capturing them may suffer from concerns about endogeneity. Following the framework proposed by Mayer (2006), I integrate transaction cost theory and the knowledge-based view of the firm to understand better the competitiveness of both sides of contracting, and a firm's potential costs and benefits that extend beyond an individual transaction, which is called spillover effects. Interfirm network embeddedness provides a proper framework to study these information effects. Panel data also improves generalizability regarding these spillover effects and mitigates the limitations of single-year, cross-sectional data that exist in the extant literature.

If firm embeddedness explains who contracts with whom, this finding will have important implications for vendors' *resource allocation* and *targeting* strategies that impact firms' long-term outsourcing growth strategies and help them win larger contracts. It will also help clients evaluate the competitiveness of sellers and reduce the costs of soliciting bids from less promising suppliers. Thus, a model based on a set of relationships that have been stable over some time is informative for both buyers and sellers. If social capital does not impact the matching outcome or if it does not do so for all types of clients and vendors, it is important to understand why and when this is the case. For example, do all dimensions of social capital predict the matching outcomes? These dimensions include "Structural Embeddedness" (SE), a measure of historical relationships around an actor, "positional embeddedness" (PE), a measure of being connected to other firms who are themselves central, and "junctional embeddedness" (JE) locating a position at the edge of two groups and building relations between dissimilar segments. SE denotes firms' network quantity, whereas PE and JE represent firms' network quality. In this chapter, I investigate the importance of the seller's network quantity versus network quality in matching outcomes. In particular, which dimensions of social capital have the most impact on the decisions of clients and vendors to match? Do client size and revenue impact the weight firms put on vendors' SC? In other words, do small, medium, large, and enterprise firms put more value on different aspects of vendors' social capital? I provide systematic empirical evidence for these questions.

My data comes from a comprehensive dataset of IT outsourcing contracts signed in years 1989-2013. I construct an interfirm network of clients and vendors that transact in IT service outsourcing markets and provide a novel explanation for the posed questions. In this panel dataset, on the client side, I observe client size measured by the number of employees, revenue, location, and the history of outsourcing contracts signed with vendors. I also observe vendors' history of contracts signed with clients. Deal characteristics include data, deal size in dollars, duration, tasks included (e.g., IT, business process, app development, HR), contract type (new, renegotiated, extension, expansion) and award mechanism (auction or negotiation). To empirically test my hypotheses, I use a two-sided matching model that is estimated with constructed network measures and other firm-specific characteristics such as revenue and size. I use a nonparametric two-sided matching model, Fox maximum score estimator (Fox 2002), to explain the vendor and the client's mutual choices that characterize contract formation. The purpose of the model is to quantify the impact of firm characteristics (size, revenue, social capital) on the likelihood of observing certain matches in the market for each task.

1.1.1 Contributions

My study contributes to the existing literature on contracts in three ways. First, it contributes to the literature on TCE by examining the impact of horizontal market effects and competition on contracting outcomes. The results show that in the presence of outside options, the market is non-monotonically sorted with respect to client size, positively sorted with client revenue and vendor social capital, and is positively sorted with respect to client network and vendor network.

The second contribution is that the model includes information spillover effects and network constraints in the TCE analysis and measures them using three concepts: SE, JE, and PE. Third, since the business network evolves as the social capital stock changes, the panel data structure allows us to capture the effects and improves the generalizability of the recommendations. While commonly used aggregate methods in answering information asymmetry concerns are effective for forecasting purposes (Muller et al. 2017), many managerial questions are normative in that the managers need to choose between paths of actions and need to know which of them will be optimal. In such cases, the aggregate approach is limited, so the structure of the underlying social network is important and should be incorporated into decision making. Therefore, my results have important managerial implications for both selling managers and outsourcing executives who want to allocate their resources to multiple projects.

1.2 Conceptual Background

1.2.1 Transaction Costs Economics

Transaction cost economics (TCE) (Williamson 1985) has focused on how transaction attributes, specifically contractual hazards arising from asset specificity (human, physical, and site), determine how a transaction should be governed. TCE focuses on how information asymmetry between agents, and uncertainty about the action of parties, demand uncertainty, and technological uncertainty influence ex-post contracting costs (Williamson 1985). When parties are bilaterally dependent, and their joint activities are interrelated, firms attempt to craft complex contracts that define remedies for foreseeable contingencies or set out processes to resolve unforeseeable outcomes to decrease lock-in concerns (Williamson 1985). When such contracts are too costly to craft and enforce, managers may choose to rely on other available information to manage the uncertainties. For example, the presence of cooperative attitudes is an alternative solution to an otherwise intractable economic problem. (Gulati 1995)

Because TCE focuses on individual transactions, it does not adequately consider the power of social context (Granovetter 1985) and the possibility that firms may be constrained in their decisions by their past choices (Argyres and Zenger 2012). However, information and knowledge that circulate between firms while working on projects have a crucial impact on firms' competitive position by adding (e.g., informational/knowledge effects) or devaluing (e.g., reputational effects) resources, and they help build or maintain competitiveness. Such knowledge spillover is studied in interfirm networks (Tsai 2001), the formation of buyer-supplier relationships (Wuyts and Geyskens 2005), R&D consortiums (Inkpen and Tsang 2005), and open source project managers (Grewal et al. 2006). I incorporate the effect of historical constraints by adopting business network theory and the interform embeddedness concept since a network establishes reliability and imposes cooperative behavior (Gulati and Gargiulo 1999) and is the source of information flow (Burt 2000) while maintaining theories' underlying behavioral assumptions. In a similar vein, when deciding to work on projects, firms consider the potential benefits and costs that go beyond the transaction. For instance, firms may decide to extensively monitor suppliers to ensure quality (Mayer 2006) when information spillover is more likely. These effects that are beyond dyadic are not addressed in the extant TCE research, but the effects are important due to firms' changing competitive position in B2B markets.

1.2.2 Knowledge-based View

TCE explains firms' outsourcing choices by analyzing transaction costs of external markets, such as search, monitoring, and enforcement costs, and not incorporating internal coordination costs such as communication. However, since IT system development is a collaborative process and firms' activities are interrelated, the successful exchange of information and seamless integration of vendors' technical knowledge with the client's business knowledge is required for effective project performance (Tiwana and Bush 2007). Anecdotal evidence also highlights the importance of client knowledge and competence in managing the outsourced project since a client's idiosyncrasies may create a new environment even for an experienced vendor. The knowledge-based view of firms (Grant 1996) deals with the impact of such knowledge consideration on firm decisions, and it predicts that higher client knowledge leads to outsourcing due to a decrease in perceived exchange hazards. KBV also examines how firms increase their stock of knowledge and consequently their competitiveness. KBV complements TCE by addressing challenges posed by requirement uncertainty, the degree to which client requirements can be communicated to a vendor (Benaroch et al. 2015).

I argue that since new projects entail inherently uncertain knowledge exchanges, both the vendor and client need to have a common understating of what needs to be done and how it will be done to increase the likelihood of project success. A client with little experience could make decisions that lead to lock-in. A firm with a higher ability to foresee and delineate the requirements is more likely to be selected as a business partner.

1.2.3 Firm Embeddedness and Social Capital

Organizational sociologists extend the efforts made by Williamson (1985) to build on bounded rationality and imperfect information concepts and argue that organizational routines, processes, and structures are embedded in the broader social context (Smelser and Swedberg 1994). Uzzi (1997) argues that the embeddedness of exchanges within social structures circumvents and thus economizes on time otherwise spent in costly contract renegotiations, leading to savings on ex-post transaction costs. The extent to which firms rely on this information depends on the uncertainties, ex-ante information asymmetry, and ex-post transaction costs, which vary by project characteristics such as size (Gopal 2013). I view firm social capital as the relationship among firms, including vendors and clients, and contracted projects that provide access to firms about information and embedded resources (Portes 1998). In the collaborative process of software development, clients form ties with vendors and exchange information and knowledge, and, in turn, gain access to complementary resources and knowledge that would not otherwise be accessible. Consequently, it should be easier for firms with higher social capital to put together the know-how with the requisite skill sets (i.e., human capital) to develop the IT systems.

The role of human capital in consumer price and search activities in packaged goods is discussed in Ratchford (2001). According to the author, human capital includes "knowledge, skills, or expertise embodied in people and acquired through investments in training or learning by doing" and explains much of consumers' loyalty. Also, the research of Luo et al. (2013) document that capital built up from past experiences with a brand can affect current consumption of the brand through changing consumers' time allocation to activities they know well. The analysis in Ratchford (2001) is at the individual level, but network theory (network effects: Ruef et al. 2003) provides a framework to study differential information capital including knowledge, skills, expertise, and reputation that firms have acquired through investments in projects and collaborating with other firms. The network theory has been applied to studying organizational activities such as receiving financing (Uzzi 1999) and hiring top managers (Granovetter 1995).

Viewed in this sense, interfirm networks function as a conduit for three distinct types of information. First, embeddedness impacts information acquired through learning by doing. This type of information increases firms' functional capabilities such as collaborative process management capabilities by helping them anticipate and communicate the requirements and better respond to uncertainties. This information also increases technical capabilities due to accumulated knowledge. Moreover, higher embeddedness has been associated with information reach (i.e., volume and speed of information access) (Wang et al. 2017). Second, social networks establish reliability (Gulati and Gargiulo 1999) since embeddedness promotes collaboration. Third, vendors with favorable locations that connect otherwise disconnected segments have access to unique, rich, nonredundant tacit information (i.e., information flow control advantage) (Burt 2005) and are perceived and indeed are more agile, which is a source of entrepreneurial efforts and product flexibility. This factor impacts the scope of the vendor's business, product mix, depth, and breadth of solutions and the firm's ability to offer tailored solutions.

1.2.4 Forms of Embeddedness and Firm Selection Decisions

How firm social capital manifests itself is of interest to both academics and practitioners. Research in network effects investigates the importance of vendor and client locations: How central is a location (Portes 1998)? How strong are the ties that the location provides (Granovetter 1973)? And How strategic is the location in terms of controlling information flow? Burt (2000) identifies two types of advantages that accrue from firms' social capital: informational advantages induced by firm embeddedness and control advantages created by bridging. The former is measured by "structural embeddedness" (SE- a measure of historical relationships around an actor) and "positional embeddedness" (PE- a measure of being connected to other firms that are themselves central). The latter is reflected in firm "junctional embeddedness," JE or brokerage- locating a position at the edge of two groups and building relations between dissimilar segments.

Network theory (Borgatti and Halgin 2011) associates structural and positional embeddedness with information reach, expertise, status, junctional embeddedness with information richness, and the likelihood of receiving more nonredundant information about other firms at any given time, which, in turn, can provide firms with the capability of offering customized solutions, performing better and faster, or being perceived as the source of new ideas. High embeddedness is closely linked to "search embeddedness" (DiMaggio and Louch, 1998). Search embeddedness indicates many direct and indirect ties with other firms, which affects information reach, including the volume and speed of information that can be accessed (Freeman 1979). Higher search embeddedness gives firms more exposure to business opportunities for two reasons: (1) Quantity derives sales. When firms want to increase their overall sales, the more clients they have, the more results they will get; and (2) Quantity helps exposure. Thus, higher sales impact the rate at which the vendor is contacted. More firms will learn about a buyer through their familiar relations if the seller increases the number of sales. Equivalently, quantity increases referrals. Therefore, firms' attempts to increase social capital by selectively entering into an exchange relationship might benefit them with search and information advantages, resulting in a reputational, functional or technical value that helps them match with business partners in the future. In this chapter of my dissertation, I aim to test whether these forms of competitive advantage impact matching outcomes. I also test for boundary conditions when these resources and capabilities matter the most by including client size.

Marketing researchers have used "information reach" to refer to the quantity of information that firms get or can be exposed to due to higher centrality (measured by firm degree), and "information richness" to refer to the quality of information (non-redundant and unique) accrued to the firm due to its brokerage potential (Wang et al. 2017). Thus, I see a trade-off between quantity and quality of information/expertise as a result of a buyer's decision to locate specific positions in the industry. High centrality with a relative lower junctional embeddedness (i.e., brokerage) implies that the firm has intentionally sold generic versions of the solutions to more buyers, since, by definition of constructs, in this case, the firm has focused on more homogenous segments of buyers. On the other hand, higher JE implies that the seller has sold more customized/differentiated solutions or has covered more niche segments. Firms decide whether to invest in projects with generic information requirements or projects that need more customization. These strategies can have a different impact on buyers' perceptions of seller's competencies. In a recent study on how this decision impacts sellers' trustworthiness, Ruth et al. (2017) adopt an analytical model to study if seller investments before signing a contract with a specific buyer can signal their trustworthiness and find a separating equilibrium in which a selfish seller chooses general investments that improve the seller's outside options, whereas a trustworthy seller chooses specific investments, leading to an increase in seller-quality generating efforts. Moreover, when knowledge spillover effects are present (i.e., when vendors can potentially reuse the learned knowledge in their future projects), they may be incentivized to sacrifice quality and customization efforts and offer a less generic solution to increase the chances of reuse (Levina and Ross 2003) and serve the competitors of the client (Bönte and Wiethaus 2007) or diffuse the knowledge to other vendors, say through staff mobility (Pack and Saggi 2001).

Table 1.1: Outsourcing Risks and the Role of Embeddedness

Risk Factor	Risk Determinants	Hazard	Existing Solutions	Role of Embeddedness
Transaction Uncertainty (TU)	-Technological Uncertainty (TeU) -Technical Complexity (TC) -Requirement Uncertainty (RU) (Gopal et al. 2003)	-Unexpected Transitions and Management Costs (Gopal et al. 2003) -Lock-in (Susarla 2012)	-IT Process, Infrastructure and Quality Certifications (ISO9001, CMM, BCP) (Paulk 1993, Banerjee and Duflo 2000) -Warranties -Pilot Projects (Snir & Hitt 2004) -Flexibility Provisions in Contracts (Susrala 2012) -Overall track record, feedback rating systems (Yoganarasimhan 2013, Lin et al. 2016, Kokkodis et al. 2015)	-High embeddedness gives access to better information and resources and increases contracting, benchmarking, and communication skills by better anticipation of requirements and contingencies. (KBV+TCE) -High embeddedness is perceived as a source of higher coordination skills, better process management, and vendor consultancy abilities, which are helpful in the presence of demand ambiguity and TU. (KBV+TCE) -Favorable location (betweenness) increases unique tacit knowledge and positively impacts perceptions of development agility in offering tailored and innovative solutions, which is also reflected in the breadth and depth of products and solutions. (TCE) -Search embeddedness. (KBV+ TCE) High embeddedness gives firms exposure and impacts the likelihood of finding matching firms that possess the required technical and functional capabilities; increasing the possibility of finding a better match.
Behavior Uncertainty (BU)	Information Asymmetry: -Geographical and cultural boundaries -unfamiliar or less experienced service providers (Lin et al. 2016)	-Opportunism and bounded reliability (Maladaptation hazard, Underinvestment) (Verbeke and Graidanus, 2009) -Ex-post transaction costs of bargaining/disputes/overruns (Bajari et al. 2007, Williamson 2010)	-Comprehensive contracts (safeguards) -Monitoring -Trust mechanisms -External expertise -Overall track records, feedback rating systems -Auction (Banerjee and Tadelis 2001) - suppliers pre-contract investments (Beer, Ann, Leider 2018) -Vertical Integration (Williamson 1971)	-Firms with high embeddedness are perceived to have *historically* behaved well, and less likely to be involved in opportunism. (TCE) -High embeddedness restrains firms from acting opportunistically in the future. (TCE) -Search embeddedness. (KBV+ TCE) High embeddedness gives firms greater information access and impacts the likelihood of finding reliable matching partners.

CMM refers to the capability maturity model (a methodology used to develop and refine an organization's software development process).
BCP refers to business continuity plans to ensure uninterrupted support/services.
Technological Uncertainty refers to shifts in customer preferences due to technological changes or end customer demand ambiguity.
Requirement Uncertainty refers to the degree to which knowledge about the client's requirements can be communicated to a vendor.
Technological Complexity refers to the breadth of tasks or type of task/engagement

I argue that social capital plays a critical role in determining the matching of partners, but the role depends on the boundary condition of the client size. Firms have different initial endowments of resources, and their subsequent matching decisions also impact this heterogeneity. The argument is based on social capital function in reducing the risks created by ex-ante information asymmetry and ex-post opportunism, leading to savings on ex-ante costly contracts and ex-post transaction costs. Table 1.1 summarizes the uncertainties and risk factors involved in the pre-selection stage in IT outsourcing and the role of firm social capital in addressing these factors.

1.2.5 Theories of Matching

The outsourcing market is conceptualized as a two-sided matching model with transfers. In a two-sided matching model, it is assumed that both sides have preferences for the agents on the other side. These models have been applied to study the professional sports market (Yang et al. 2009), auctions of licenses used in the radio spectrum for mobile phone service (Fox and Bajari 2009), online dating (Hitsch et al. 2010), and celebrity endorsements of brands (Zamudio 2016). In matching with transfers, firms must pay a potential partner to match, and the required payment involves the characteristics of rival agents (Shapley and Shubik 1972, Becker 1973). Fox (2009b) studies identification in matching games with transfers and determines the order of matches (i.e., which matches provide the highest value). The relative importance of complementarities in payoffs from different agent traits, say firm size or experience, is identified using data on matches but not the equilibrium transfers. Thus, even though transfers are observable in my context, I do not use them for estimation of preferences because Fox (2009b) proves that preference rankings are identifiable from observed matches and not necessarily from transfers.

Ackerberg and Botticini (2002) supplement the contract choice probit equation with an ordered probit matching equation to address endogenous matching; however, a probit model does not capture the sorting that is the source of identification in matching models (Sørensen 2007). Prior literature has explored parametric estimation in non-nested matching games, that is, games without endogenous transfers (Sørensen 2007). The identification arguments presented in this chapter are based on maximum score estimators and are fully nonparametric.

Fox (2010a) shows that the maximum score estimator is consistent if the stochastic structure of the model satisfies a rank-order property, drawing on the literature on single-agent, multinomial choice (Manski 1975, Matzkin 1993). Rank-order property deals with inequalities arising from pairwise stability and probabilities of different equilibrium assignments and implies that I do not need to include all possible inequalities. Instead, I sample from a subset of them. Mathematically, it implies that

$$P[v\ sells\ to\ c\ and\ v'sells\ to\ c' \geq p[v\ sells\ to\ c'and\ v'sells\ to\ c]$$

This identification approach offers computational simplicity because it allows us to work with inequalities that I derive from pairwise stability conditions, instead of working with high-dimensional integrals over match-specific unobservable, and considering the entire assignment in addition to with whom the deviating players would match for each deviation (Pakes and Pollard 1989). The limitation of the approach is that client- or vendor-specific attributes, as well as transfers, are canceled out when I add inequalities (Akkus et al. 2015).

1.3 Model Development

1.3.1 Model Setup

I conceptualize the IT outsourcing market as a many-to-many market with vendors selling to multiple clients, and clients buying services from multiple vendors. The market is also conceptualized as a two-sided matching market with transfers, implying that payments (the contract price) or other benefits (value for money including after-sale support) must be given from one party to the other, although my estimation is independent of the transfer between firms. Matches are the outcome of a competitive market. I use the structure of the market to define the two-sided matching market boundaries, i.e., task types, to identify the outside options for each firm within that boundary (Becker 1973, Roth and Sotomayor 1990) and use characteristics of these rival agents in comparison to the actual vendor-client matching choices to structurally estimate joint value creation.

I collect data on the matches in each of several independent outsourcing matching markets defined by task type. I observe characteristics of vendors and clients in each task, as well as the sets of matches that occurred, and aim to simultaneously infer client and vendor preferences. I observe equilibrium outcomes that result from mutual choices and not choice sets so identification in this type of model cannot be based on the analysis of single-agent demand models (Fox 2010). Choice models, such as logit, are not entirely appropriate due to biased preference estimates (Ni et al. 2015). Because I adopt a two-sided matching model, the outcomes (here contracts between vendors and clients) should result from the two-sided matching process. I believe this is a proper assumption since both clients and vendors have the right to choose with whom to transact in competitive markets and have the discretion to negotiate their contracts. Specifically, I am dealing with customized services that vendors have not produced yet, and the successful development of these services requires the collaboration of both sides. This requirement implies that vendors have preferences for the preferred clients who are believed to have the prerequisites for being a good matching partner. Thus, regardless of what occurs before the selection process ends, the observed equilibrium matches are informative of sides preferences for the other party's attributes. Because vendors and clients choose their best potential matches, observed outcomes can be used to estimate the joint value created (Yang et al. 2009) from the outsourcing arrangement. A firm's best response indicates that the total utility of client c_1 from working with vendor v_1 exceeds its utility, instead of working with v_2 at a transfer level that would make v switch from its observed matched client c_2, which is weakly determined by the transfer level that makes v_2 indifferent between c_1 and c_2.

Clients and vendors assess the utility they derive from each other and choose a partner that maximizes their utility. In a two-sided matching model, the joint utility concept is used (Fox 2010) and is defined as the joint value that is perceived to be created as a result of working with each other. The value is conceptualized as the benefits that firms get from the match above the transfer fees such as the contract price. The joint value of the match is measured employing a "production function," which includes the interaction of observable characteristics of firms, which are selected characteristics that are believed to affect the value creation. Firms are matched such that they derive the highest possible joint value from the partnership compared to the counterfactual joint value created from matching with other pairs. The result is an equilibrium set of matches, and the matches between firms with certain types of characteristics are then more likely to be observed in the market. Matching games are cooperative games and use the cooperative solution concept of "pairwise stability" (Roth and Sotomayor 1990) instead of Nash equilibrium as the main solution concept. Pairwise stability allows only a single pair of potential partners to consider deviating from the proposed equilibrium at once. This could be a logical assumption since higher numbers of deviations require a larger number of firms to strategically coordinate their actions to do so (Pakes 2010). This process alleviates the need for solutions that must be considered, for each deviation, what the entire assignment would be in addition to with whom the deviating players would match.

Another important equilibrium concept of two-sided matching models is assortative matching. Pairwise stable equilibrium outcomes could be positively assorted, which means that agents' differential traits are complementary in the production function, resulting in matching with similar agents (marriage markets; Becker 1973), or negatively assorted, when a negative correlation between traits of agents exists and sides match with dissimilar agents (corporate alliances; Amaldoss and Staelin 2010). Mathematically, when dealing with only one trait for each side (Becker 1973), positive sorting means $\frac{\partial^2 f\left((v_1),(c_1)\right)}{\partial v_1\, \partial c_1} > 0\ \forall\ v_1, c_1$. Preferences also come into play when they affect the production costs, so if positive assortative matching is observed between firms and firms enjoy increasing returns to scale in the production function, the average costs of production fall at higher complementarities.

1.3.2 Formal Matching Model

The mutual choices that characterize contract formation can be analyzed using two-sided matching theory. In two-sided matching markets, the participants in one side of the market match with participants on the other side (Roth and Sotomayor 1990). The fundamental unit of analysis is the match which is a partnership between the two participants as a result of both mutually choosing each other. The sorting model controls the availability and impact of outside options for both sides.

In equilibrium, no pair has an incentive to deviate from the match. For example, no two firms are willing to exchange partners. The matching game assumes the created joint value after the exchange of partners in the new pairs is less than the created joint value in the original pairs.

A vendor might value a match with a specific client because of lower agency costs associated with dealing with that client. I assume that client size can capture this effect by changing the vendor's valuation of the deal. Similarly, clients might value a match with a vendor because of the vendor's ability to provide insight into their operations with regards to how to improve the clients. Clients also value vendors' agility in adopting new technologies and techniques. To maintain the pace of IT service changes, they prefer to work with vendors who adjust the process in a timely manner, since otherwise clients would need to start a new relationship with a better vendor from scratch, which would be costly. Hence, from the pool of vendor characteristics, I hypothesize that a vendor's social capital would change the valuation of the client.

To model the matching of vendors with clients, I use the nonparametric approach, called "maximum score estimator," developed by Fox (2008, 2010). In Fox's model, firms' fixed effects are not estimable, and it is only possible to put interactions in the production function to estimate these effects. To define equilibrium and my solution concept, I use "the local production maximization condition" developed by Fox (2008). The local production condition means that the total production of any two observed matches should exceed the total production from the exchange partners; otherwise, additional matches could also be formed without affecting the existing assignments in the market. An alternative solution concept would be core stability in which no group of agents would deviate from their assignments. The computational costs of estimating such a model is far greater than its benefits (Fox 2008). Therefore, I use the local production maximization for computation.

In the model below, R^V is the match payoff to the vendor, and R^C is the match payoff to the client. Transfer between agents is represented by t_{vc}.

$$R^V(v,c) = r^V(v,c) + t_{vc} \qquad (1)$$

$$R^C(v,c) = r^C(v,c) - t_{vc} \qquad (2)$$

Pairs are not willing to exchange partners in equilibrium, hence:

$$R^C\,(v_1, c_1) \geq\; R^C(v_2, c_1) \qquad (3)$$

Indifferent vendor v_1 who is in a match with c_1 receives $\tilde{t}_{v1c1}$ to be indifferent between firms c_1 and c_2.

$$r^V(v_1, c_1) + t_{v1c1} = r^V(v_1, c_2) + \tilde{t}_{v2c2} \qquad (4)$$

From (3) I get

$$r^C(v_1, c_1) + t_{v1c1} \geq r^C(v_1, c_2) + \tilde{t}_{v2c2} \qquad (5)$$

After some rearrangement and writing inequality (5) for pair (v_2, c_2) as well, I have:

$$r^C(v_1,c_1)+r^V(v_1,c_1) + r^C(v_2,c_2)+r^V(v_2,c_2) \geq r^C(v_1,c_2)+r^V(v_1,c_2) + r^C(v_2,c_1)+r^V(v_2,c_1) \quad (6)$$

Production to match is defined as

$$f(v_1, c_1) = r^C(v_1, c_1)+r^V(v_1, c_1) \qquad (7)$$

Inequality (6) summarizes to

$$f(v_1, c_1) + f(v_2, c_2) \geq f(v_1, c_2) + f(v_2, c_1) \qquad (8)$$

Since firm fixed effects cancel out from two sides of inequalities, it is only possible to put interactions in the production function to estimate the effects. Function is defined below:

$$f(v, c) = \beta \times [V \times C] + \varepsilon_{vc} \qquad (9)$$

V= $\{V_{SE}, V_B, V_{PE}\}$ is a vector of vendor characteristics which measure vendors' social capital. Moreover, C=$\{C_{sizedum}, C_{revenue}, C_{SE}, C_{PE}\}$ is a vector of client characteristics and includes client size dummy variables, the log transformation of client revenue, client total degree centrality and client positional embeddedness measured by eigenvalues.

Parameters of interest in β are estimated using Fox's (2010) maximum score estimator, which follows the equation below:

$$\max_f Q_H(\beta) = \frac{1}{H} \sum_H \sum 1[f(v_1, c_1) + f(v_2, c_2) \geq f(v_1, c_2) + f(v_2, c_1)]. \qquad (10)$$

H is the number of markets. Markets are defined at task types and year level. The pairs to the left of inequality (8) represent actual matches in the market, and the pairs to the right denote counterfactuals after considering pairs that switch partners. The objective function finds parameter β by maximizing the number of inequalities that hold. The estimation uses a recursive algorithm: I start with an initial value of beta and then calculate the value of the objective function for the maximum score estimator according to (10). Then, I apply the global optimization routine known as differential evolution (Storn and Price 1997) to find the optimal parameters. Maximum score estimation does not assume a distribution for error terms; hence, confidence intervals are calculated using subsampling. A large number of counterfactual matches can be created from matching data. These allow us to estimate β using Manski's (1975) semiparametric maximum score estimator. This estimator does not require specifying the probability distribution of the error terms.

Equation (4) means that for every market in the data, the matching value generated by each pair of observed matches is compared to that generated by each pair of counterfactual matches. $1[\cdot]$ is an indicator function that takes the value of 1 when the matching value generated by the observed matches is superior to the counterfactuals, and 0 otherwise. Thus, the estimator finds the optimal matching value that causes the maximum proportion of inequalities or comparisons to satisfy the equilibrium condition (Fox 2010). The maximum score can be thought of as the fit of the model: a higher maximum score implies that a larger number of the inequalities in the data satisfy the structural condition.

1.4 Methodology

1.4.1 Data

I use comprehensive data of IT procurement transactions (IDC, 2013). The database includes IT contract data signed between 1989 and 2013. Each of these contracts is a new contract and has a unique contract ID number. Data are right censored, and the outcomes of 1717 contracts are fully observable. The contract status for 47,355 projects is listed as “Current or unknown.” The average project size is $58m, and the average project lasts 69 months. Such large projects necessitate extensive pre-screening and underscore the need to develop a formal model to analyze the matching behavior when clients decide to gain more value from outsourcing.

I observe client, vendor, and deal characteristics in the data set. In general, there are four main macro markets: outsourcing, projects, support/training, and other services. Macro markets are further decomposed into 15 tasks (e.g., IT, business process, and system integration outsourcing). Figure 1.1 shows the distribution of tasks observed in the original data.

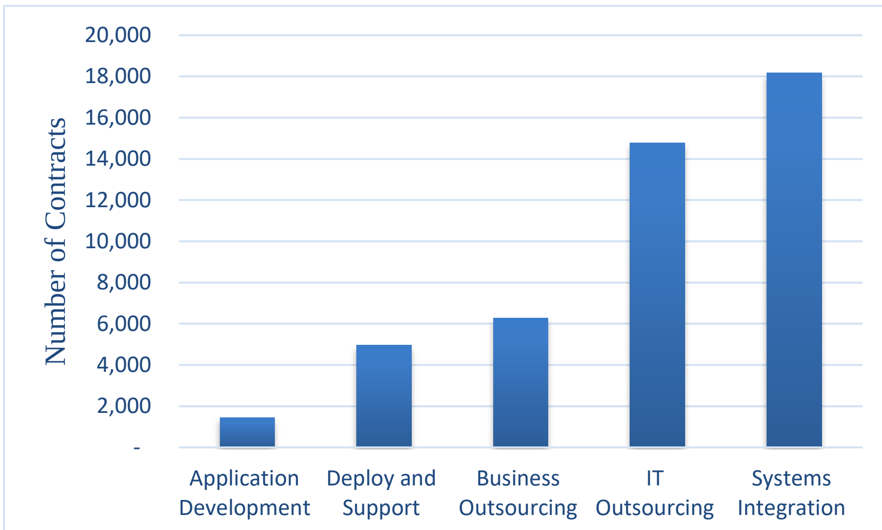


Figure 1.1: Distribution of Top five Engagement Types

Contracts can have values ranging from several hundred dollars to multi-million dollars. All of this depends on the characteristics of the products or services that are offered. For example, contract values for bridge construction or defense and aerospace equipment are quite substantial, whereas contracts for application development service solicitations are usually significantly less. Figure 1.2 shows the average contract value and length for different task types. Of all the auctioned projects, the identity of bidders is revealed for 1241 of them. Typically, contracts are signed for 34 months, but I observe very short (i.e., one month) and very long (i.e., ten years) contract duration. There are 7,816 unique vendors and 22,507 unique clients. On average, a vendor signs six contracts (mean=6, SD=41). Table 1.2 shows the basic statistic measures for vendor contract frequency. An average client signs two contracts in the data. Of 49,072 observations, 3.7% are coded as “off-shore/nearshore” and 50.5% are coded onshore. The status of the rest is not observable. Lastly, less than 1% of contracts are between a client and a subcontractor, and most contracts are awarded to primary vendors.

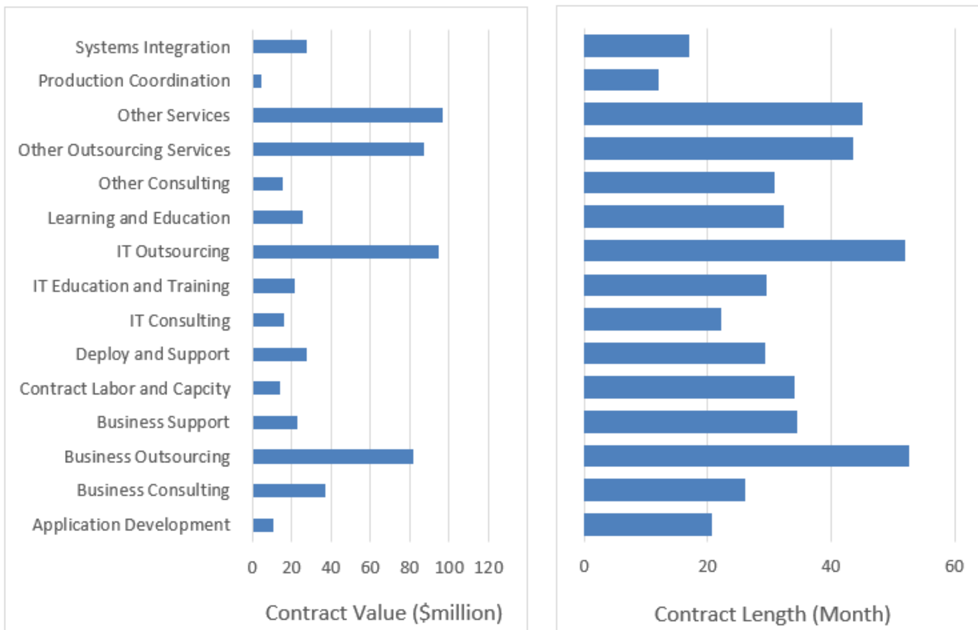


Figure 1.2: Distribution of Average Contract Value and Length by Task

I use vendor network measures of the three largest markets: IT outsourcing, system integration, and business outsourcing. The selected segments include 89% of the original 1717 contracts, totaling 1526 deals. Extracting a subset of observations from these 1526 contracts, so no missing data for contract type (auction or negotiation), global sourcing strategy (offshore or onshore) and with at least one observation in prior years leaves us with the final sample size of 657 contracts.

Table 1.2: Vendor Contract Frequencies

Statistic	Value
Mode	1
100% Max	2289
90%	7
75% Q3	2
50% Median	1
25% Q1	1
10%	1

1.4.2 Market Definition

A matching market contains all vendors who provide similar tasks in a specific year and all the clients who are observed transacting for that task in that year. I choose three tasks: IT outsourcing, system integration, and business outsourcing from contracts that have an identifiable end status and have no missing data for the contract type (auction or negotiation) and global sourcing strategy (offshore or onshore). Because contract networks evolve throughout a study, I use history measures to calculate the variables. For example, if IBM signed ten contracts in 2000, eight contracts in 1999, and seven contracts in 1988; I use the contracts signed in 2000 or before to calculate the variables for all the IBM contracts in 2001. In simple words, I used lagged network measures to predict the network and matches in the current time.

Figure 1.3 shows a typical market that is observed in the data. The market has two sides who compete with similar firms on their side to sign a contract with a better firm on the other side of the market. In the market for IT solutions in the year 2013, buyers and seller who transact with each other for a specific task (here IT Solution) constitute a market. So, a market is defined at the task and year level; because later in section 1.5 I will show pairwise stability assumption makes

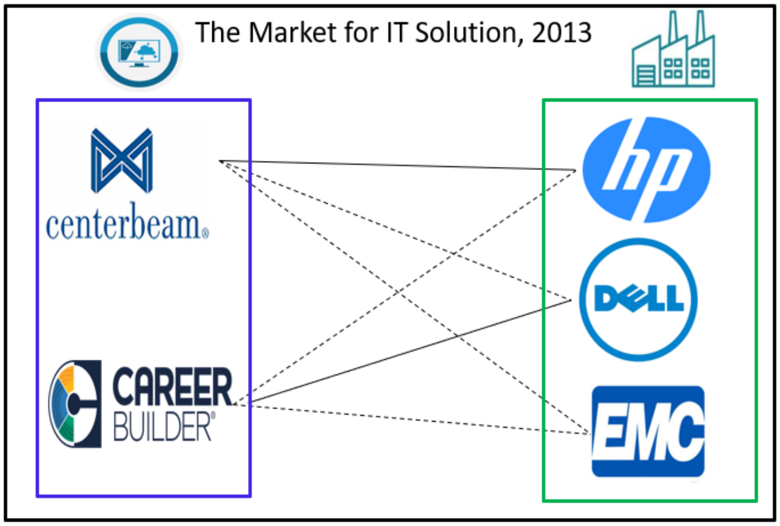


Figure 1.3: Two-sided Matching Market Conceptualization.

sense under this market definition. In Figure 1.3, HP, Dell, and EMC are vendors, and CenterBeam and CareerBuilder are buyers who were in the market for an IT solution. Solid lines show the observed matches, dashed lines show the counterfactual matches- matches that I did not observe but could have happened if they provided higher joined match values to both sides.

1.4.3 Measures

Structural Embeddedness (SE). Degree centrality (Freeman 1979) is used to measure structural embeddedness and is defined as the number of links incident upon a node (i.e., the number of ties that a node has). The degree can be interpreted in terms of the likelihood of the node to catch the information flow of the network.

$$SE = cumulative\ observations\ per\ client/vendor\ \ so\ far \quad (11)$$

The number of links incident upon a node captures the number of unique firms a node has transacted with and is different from cumulative contracts a firm has signed because the strength of a tie with between a pair is not incorporated in the current SE measure. The strength of a node can be calculated by adding relationship strength- the number of prior collaborations in a dyad-to the calculated SEs.

Positional Embeddedness (PE). I use eigenvector centrality to measure positional embeddedness (Bonacich 1987). Eigenvector centrality measures the importance of the vendor (client) by giving weight to the centrality of the client (vendor) to which the agents are connected. This measure calculates the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that not all connections are equal, and that connections to firms which are themselves central will be given a higher score than connections to less central firms. The weight takes a value between 0 and 1, and a value of 0 means the importance of the partner, measured by the number of their connections in the network, does not matter. I use the value of 1 for ω.

The centrality of unit i is calculated by the following expression (Bonacich 1987):

$$\lambda e_i = \sum_j R_{ij} e_j,$$

where R is the matrix of relationships, λ is a constant that makes the equation have a nonzero solution. The matrix notation for the equation is,

$$\lambda\ e = R\ e,$$

where e is an eigenvector of R, and λ is its eigenvalue. Bonacich's proposed measure follows the following formula:

$$c_i(\alpha, \omega) = \sum_j (\alpha + \omega c_j) R_{ij}$$

Equivalently,

$$c(\alpha, \omega) = \alpha (I - \omega R)^{-1} R 1,$$

where I is an identity matrix, and 1 is a column vector of ones. The coefficient α is a scaling parameter and is set such that the sum of squared scores equals the number of nodes.

Junctional Embeddedness (JE). I calculate the brokerage potential of vendors by using vendor "betweenness centrality," which measures how well the vendor can mediate the information and how well the vendor catches the experience, learning, and information spillovers from other pairs in the network (Freeman 1977). Betweenness is the sum of all the shortest paths that pass through a vendor and reflect the ease with which the vendor can transfer information flow in the network. Betweenness centrality is an approximate measure for how the flow of information in the network between pairs is dependent on the presence of the actor, and if removing the actor disturbs the flow of information more, the dependence of the whole network on that actor for information diffusion is higher. The log transformation of JE ranges from 0 to 4.8 in my selected sample used in the estimation. The shortest-path betweenness of a node, v, is calculated by the formula below:

$$C_{JE}(v) = \sum_{i,j: i \neq j, i \neq v, j \neq v} \frac{g_{ivj}}{g_{ij}}$$

where g_{ivj} is the number of shortest paths from i to j from v. High betweenness vendors lie on a large number of non-redundant shortest paths between other firms; thus, they can be thought of as "bridges."

Example of Network Measures. To have a clearer picture of centrality measures, I develop an example of a client-vendor network and calculate the network measures. The network graph is presented in Figure 1.4. The centrality measures for the same graph are calculated using R software and reported in Table 1.3. Degree centrality is the number of edges in/out of a node. For example, vendor 4 and vendor 2 have the highest degree of centrality.

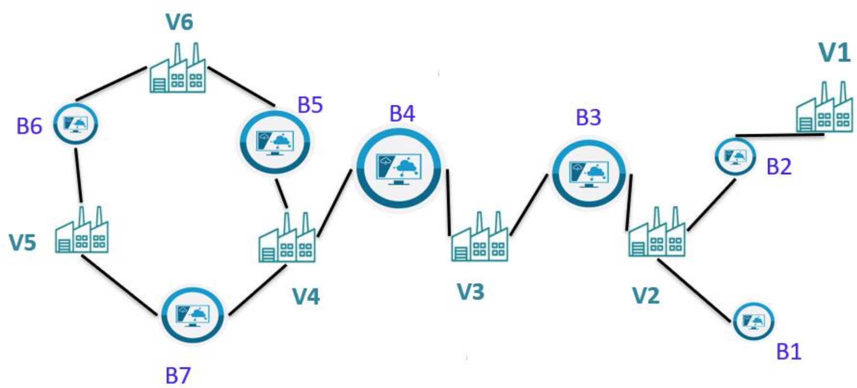


Figure 1.4: An Example of Buyer-Vendor Network.

Vendor 4 has the highest betweenness because it brokers between more pairs. Vendor 1 has the lowest betweenness score since it brokers between no pairs. Eigenvalue captures the position in the network and is higher if agents are connected to other well-connected firms. For example, clients 5 and 7 have the highest eigenvalue since they are connected to central vendor 4.

The "network measures" used in the main model are calculated based on all data (i.e., 49,072 observations). I must sample from the contracts to estimate the matching model because it is computationally impossible to do it on the whole data. The sampling procedure is explained in section 4.1; the final number of observations used in estimation is 657.

Table 1.3: An Example of Network Measures

	Degree	Betweenness	Eigenvalue
V1	1	0	0.096
V2	3	29	0.356
V3	2	35	0.504
V4	3	37	1.000
V5	2	5.5	0.592
V6	2	5.5	0.592
B1	1	0	0.164
B2	2	11	0.209
B3	2	32	0.397
B4	2	36	0.695
B5	2	12.5	0.735
B6	2	2	0.547
B7	2	12.5	0.735

To offer a visual description of how networks evolve, the network graphs are drawn for specific tasks in certain years. Figure 1.5a shows the network of firms who transacted in IT outsourcing market in the year 1997. Firms' JE values scale size of the nodes. Bigger nodes represent higher JE. IBM, EDS, and CSE are top JE vendors and Imperial chemical co. is top JE client. Figure 1.5b shows the same network, except that firms' PE values scale the size of the nodes. Bigger nodes represent higher PE. EDS is shown to have the top PE, followed by IBM. Figure 1.5c shows the network in year 2004. Bigger red circles represent firms with higher PE. Figure 1.5d shows how the 2004 network has evolved in year 2013.

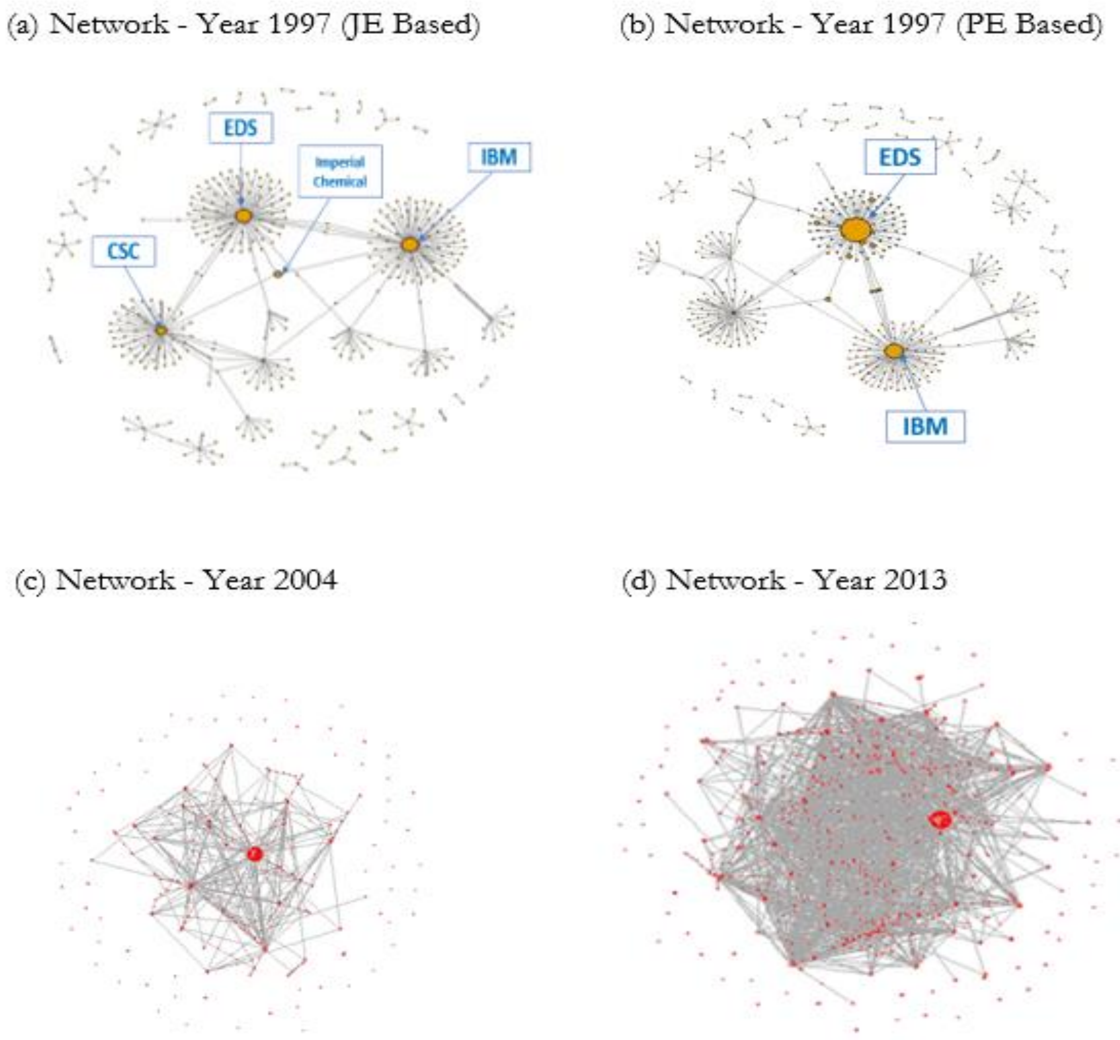


Figure 1.5: Buyer-Vendor Network

Client Size. I use dummy variables to code client size based on the number of employees and classify them into four categories: small (0-99), medium (100-999), large (1000-4999), and enterprise (5000+). Banerjee and Dufflo (2000) use firm size as a proxy for their risk aversion. Larger firms are assumed to be more risk-averse. In my context, client size determines their bargaining power (Gopal et al. 2003) and outside options. Agency frictions and transactional costs are higher for very large clients. Large clients can impose higher operating costs because transaction uncertainties are higher for complex projects. Figure 6 shows how the distribution of client sizes evolve.

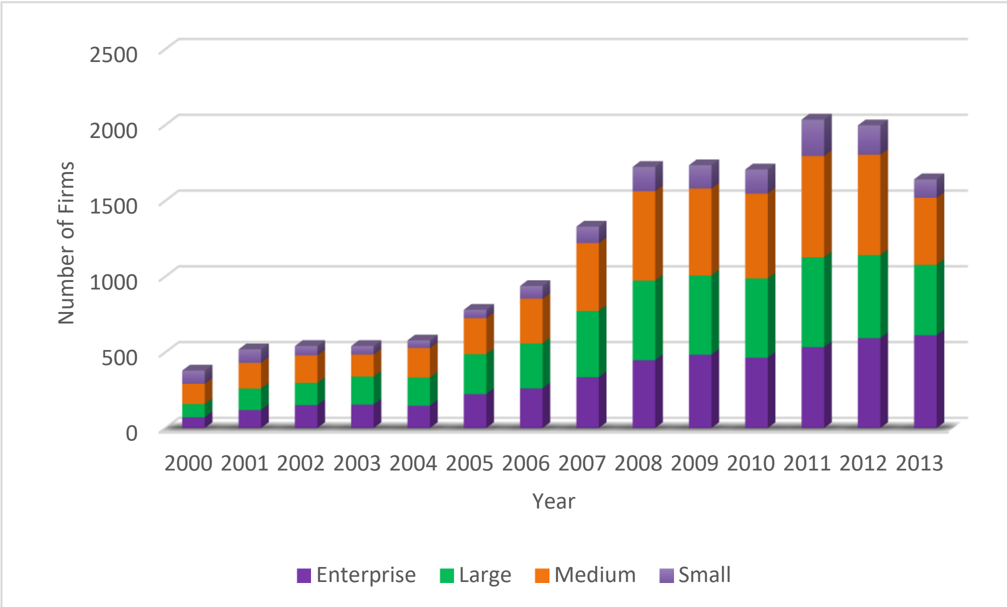


Figure 1.6: Client Size Distribution over Time

Client Revenue. In the sourcing market, clients with higher revenue are considered low risk due to their better ability to pay; thus, they are considered more attractive (Ni and Srivinsan, 2015). On the other hand, larger clients have higher negotiation power and could change the contract terms to their benefit. They can also have more outside options, and search frictions due to search cost are lower for clients with higher revenue. I use the log transform of client revenue in dollars.

Table 1.4: Client Network Descriptive Statistics

Segment	Variable	Mean	S.D.	Minimum	Maximum
All Clients (n=22507)	Revenue($b)	5.51	250.45	0	36,600
	C_SE	3.12	6.30	1	132
	C_JE	10.08	2.18	0	15.77
	C_PE	0.01	0.02	0	0.71
Enterprise (n=4812)	Revenue($b)	19.30	529.17	0	36,600
	C_SE	4.13	7.80	1	132
	C_JE	9.66	1.98	0	14.78
	C_PE	0.01	0.02	0	0.33
Large (n=5056)	Revenue($b)	2.98	77.37	0	5,266
	C_SE	2.50	5.05	1	96
	C_JE	9.66	1.98	0	14.78
	C_PE	0.01	0.02	0	0.71
Medium (n=5519)	Revenue($b)	1.92	75.30	0	5,402
	C_SE	2.12	3.50	1	61
	C_JE	9.61	2.43	0	15.27
	C_PE	0.00	0.01	0	0.15
Small (n=1670)	Revenue($b)	0.05	0.40	0	14.08
	C_SE	1.57	2.20	1	21
	C_JE	9.18	2.99	0	14.07
	C_PE	0.00	0.01	0	0.06

Two statistics are found in Table 1.4 and Table 1.5. Table 1.4 shows network measures calculated for clients of different sizes. For example, there are 5,519 medium sized clients, the average values for client revenue, SE, JE, and PE are $1.92b, 2.12, 9.61 and 0.001, respectively. The maximum PE belongs to large clients (0.71); the value is much larger than the maximum PE observed forenterprises (0.31). Table 1.5 shows statistics of the vendor network calculated at match level. For example, medium size clients, on average, have transacted with vendors whose SE, JE and PE are 55.19, 9.83 and 0.08, respectively.

Table 1.5: Descriptive Statistics Observed in Client and Vendor match

Segment	Variable	Mean	S.D	Minimum	Maximum
All Clients (n=22507)	V_SE	62.52	108.70	1	735
	V_JE	10.21	3.97	0	16.35
	V_PE	0.10	0.27	0	1
Enterprise (n=4812)	V_SE	69.84	114.57	1	735
	V_JE	10.56	3.65	0	16.35
	V_PE	0.12	0.28	0	1
Large (n=5056)	V_SE	65.65	109.62	1	735
	V_JE	10.32	3.93	0	16.35
	V_PE	0.12	0.28	0	1
Medium (n=5519)	V_SE	55.19	101.12	1	735
	V_JE	9.83	4.25	0	16.35
	V_PE	0.08	0.23	0	1
Small (n=1670)	V_SE	43.67	85.24	1	735
	V_JE	8.96	4.54	0	16.07
	V_PE	0.07	0.23	0	1

Task Type. Different IT services may need different investments. According to Gopal (2013), the required investments are associated with project uncertainties. I use IDC's classification of tasks for this matter. For my final model, I use data from three task types: IT outsourcing, system integration, and business outsourcing. The average contract values for the three tasks are $106.4 m., $35 m., and $96.1 m, respectively. The selected tasks are the top three frequent tasks observed in the data. System integration, IT outsourcing, business outsourcing, and business deployment tasks each constitute 37%, 30%, and 13%, respectively, of the total recorded deals.

Vendor network measures are calculated separately for each task. However, client network measures are calculated based on all outsourcing deals, since according to my theoretical arguments, client embeddedness impacts the general experience with outsourcing (KBV). In contrast, vendor embeddedness is associated with transaction-specific uncertainties that are assessed at the task level.

1.5 Model Estimation

In this section, I explain the estimation steps. Since the maximum score estimator is a nonparametric estimation method, the approach first identifies the optimal parameter by applying Differential Evolution (DE) method on the objective function and then subsamples vectors of parameters to create confidence intervals. The related steps include:

Step 1. Define markets at the task- and transaction-year level and form original and counterfactual pairs within each market. For example, market i includes vendors and clients that transacted for the business process outsourcing task in year 2009. This guarantees that I do not allow for all possible pair swaps in the data, and only logical swaps are allowed.

Step 2. Construct interaction terms for both original pairs and counterfactuals. Then sum up the interactions for original pairs to form equilibrium interactions. Similarly, I sum up the interactions for counterfactual pairs to form counterfactual interactions. For identification, the model requires setting one of the interaction parameters to −1 or +1 (Fox, 2010). I keep +1 for the interaction term Vendor Eigenvalue × Client Degree since the model's maximum score is higher with +1 compared to -1.

Step 3. Set the initial values for the parameters in the production function f by producing random numbers. In this step, a matrix M of random numbers is generated: with D parameters to estimates, a matrix of D by 5D is generated.

Step 4.a Use DE method on objective function f to find the optimal joint utility estimates. Calculate the maximum score value. This is the percentage of times the local production condition ($f_{equil} - f_{counterfactual} > 0$) holds, where f_{equil} is the sum of joint utility functions for original pairs and $f_{counterfactual}$ is the sum of joint utility functions for swapped pairs.

Step 4.b DE first chooses three random vectors−the so-called population vectors−from matrix M. DE then generates a new parameter vector by adding the weighted difference, created by multiplication with a real and constant number F [0,2], between two population vectors to the third vector. This operation is called mutation. The mutated vectors' parameters are then mixed with the parameters of a predetermined vector which is called the "target vector" −in my case the ith column of matrix M−to create the trial vector. The target vector changes from i=1 to the size of matrix M. The objective function is calculated for the trial vector and is called "score." It is also calculated for the target vector i and is called "cost." If score > cost, then the trial vector replaces the target vector. Otherwise, the target vector is kept, and the loop continues to check the comparison for higher values of i. I set F=0.5. I repeat the whole process by inserting step 4.b in a larger loop for G=2000 iterations to find the global optimum.

Step 5. Calculate the confidence intervals using subsampling. I generate point estimates by running the differential evolution routine for ten times and choose the coefficient vector that yields the highest value for the maximum score objective function. For inference, I follow the procedure suggested by Santiago and Fox's (2008). Specifically, from all of the data used for estimation, which includes *n (*=567) contracts, I generate 190 subsamples (1/3 total sample size) of size n-1 by removing one random contract each time. Then, I run the optimization for each subsample and create the optimal parameter estimates. These computed parameters create the empirical distribution from which I compute the confidence intervals and statistical significance.

1.6 Results

The observed matches are the market's equilibrium outcome, and in equilibrium, no firm should be willing to switch partners. Consequently, observed matches are more likely to occur than unobserved matches. I generate unobserved matches (i.e., counterfactual matches) from the data by combining the participants observed to match those who did not. Since counterfactual matches never occurred, I can say that the observed matches should have provided firms with a higher matching value; otherwise, swaps would have been made. For identification, the model requires setting one of the interaction parameters to -1 or $+1$ (Fox, 2010). I keep $+1$ for the interaction term Vendor Eigenvalue × Client Degree since the model's maximum score is higher with that value. Operationalization of the variables that are used in the model is found in Table 1.6. Estimated values of the parameters used in equation (10) and their confidence intervals are presented in Table 1.7.

The interpretation of betas is somewhat different for “continuous × discrete” and “continuous × continuous” interaction terms. Positive coefficients for vendor SC × client dummy variables (i.e., client size) mean that the match will add more to the generated joint value compared to a match with the baseline cases, which is a match involving small size clients. Negative coefficients for vendor SC × client size mean that the match will add less to the joint value compared to a match involving small clients.

Table 1.6: Variable Operationalization

Variables	Operationalization	Notes
Dependent Variable		
Match Outcome	Match joint value to pair ij with their vector of specific characteristics V and C in the market	Match Likelihood
Independent Variables		
Social Capital	Measured by three variables: SE, PE, B	Information reach and richness
Structural Embeddedness (SE)	Firm degree centrality=$cumulative\ observations\ per\ client/vendor\ so\ far$	Network Quantity
Positional Embeddedness (PE)	Eigenvector centrality, which represents that the firm is more central if it is connected to other centrals	Network Quality
Junctional Embeddedness (JE)	Freeman Betweenness Centrality (BC), which represents the number of disconnected firms that are connected due to the vendor's location in the market	Network Quality
Client Size	Number of employees: small (0-99), medium (100-999), large (1000-4999) and enterprise (5000+)	Client Bargaining Power & Outside Options
Client Revenue	Log transformation of client annual revenue in dollars	
Matching Markets		
Task Types (i.e., Engagements)	IDC's Worldwide Services Taxonomy (App Development, Business Consulting, Business Outsourcing, Business Support, IT Consulting, IT Outsourcing, System Integration, Production Coordination)	

These effects are beyond the effects of client SC and revenue and capture a different aspect of the matching that is not reflected in client outsourcing experience (reflected in SC) and financial strength (reflected in client revenue). Client size determines firms' *opportunity* to increase vendor transaction costs after the deal is made, by using "bargaining power" to take advantage of vendor lock-in or "agency costs" associated with managing such clients. Higher client revenue is-

Table 1.7: Two-Sided Matching Model Results

		Vendor Characteristics		
		Structural Embeddedness (SE)	Junctional Embeddedness (JE)	Positional Embeddedness (PE)
Client Characteristics	Enterprise	-19.26**	0.24**	-4.73**
		(-20.05, -17.80)	(0.22, 0.49)	(-10.23, -3.97)
	Large	9.06**	0.151**	5.93**
		(7.91, 9.63)	(0.12, 0.179)	(5.18, 10.93)
	Medium	0.21	0.04	23.50**
		(-0.9, 1.15)	(0.05, 0.08)	(20.80, 26.29)
	Small	0	0	0
	Revenue ($m)	3.79**	0.047**	24.21**
		(3.41, 3.95)	(0.046, 0.048)	(22.22, 28.99)
	PE	1.10	71.50**	1.32
		(-0.84, 5.19)	(71.43, 74.43)	(-1.16, 4.65)
	SE	1.36**	-0.252**	1
		(1.14, 1.56)	(-0.28, -0.250)	-

Model Score: 64.6%, Sample Size: 657

Note 1: Confidence intervals are shown in brackets.
Note 2: The cell values represent coefficients of interactions for V and C. Positive values mean a positive joint value of the match.

correlated with high project sizes, and higher levels of investments (relationship breadth: the number of unique tasks in a dyad and relationship depth: cumulative expenditure in time). Investments in dyad have been shown to affect a firm's *motivation* to collaborate in the future. As explained in KBV Section 1.2.2, since project development in this context is a collaborative process, client *ability* to manage the process by better anticipation and communication of requirements-which is hypothesized to be reflected in its SC- can affect the value generation by decreasing ex-post transaction cost which will arise in case of costly amendments.

Positive beta for interactions of two continuous variables-vendor SC and client SC- mean that the generated matching value is higher than if the interaction was insignificant (and, thus, valued at zero), with the opposite interpretation for negative estimates[2]. It also means high type firm extracts more value from the match compared to low type firm, and that the match is more valuable as the sum of firm networks increases. It implies actors' traits on two sides of the market are complements. Negative coefficient means the match is more valuable when high type matches with low type, which occurs when traits are substitutes. Thus, one can expect that in equilibrium, positive beta means high types match with high types, medium with medium, and low with low (Becker 1973).

On the other hand, negative coefficient means high type matches with low type and that actors' traits are substitutes. In sum, a positive beta for interactions of two continuous variables-vendor SC and client SC- means symmetry in traits is valued. Negative beta for these interactions

---

[2] An increase in vendor SE, JE, PE *increases* the added value to the match with a client that has *positive* interaction sign. An increase in SE, JE, PE *decreases* the value of the match with a firm that has *negative* interaction sign. SE, JE, PE do not have any effect on the value of the match with a firm that has *insignificant* interaction sign. For example, if we could estimate match specific parameters (i.e., the parameters that are not client or vendor specific and instead are match specific, factors such as distance between firms, project size, relationship history, etc.), one can expect the estimated beta for distance is negative because monitoring and collaborating on projects is harder when firms are not geographically close. Thus, the more distant a firm is from its partner, the lower is the value that can be created.

means asymmetry in traits is valued. When this explanation on how to interpret the coefficients, I now turn to the main findings. The detailed results are presented in two separate parts: first, the impact of vendor network and client network on match value is discussed. Second, the association between vendor network and client revenue and size is explained.

Conclusion 1. Table 1.7 shows that the coefficient estimates on the interaction between vendors network, and client network is statistically significant in most cases; the exception is the interaction between vendor JE and client SE. These findings suggest that highly embedded vendors generally match with highly embedded clients, and these matches are revealed to be more valuable compared to the counterfactual matches that did not occur. In other words, there is evidence of positive assortative match on both PE and SE. For example, the last coefficient in the first column (i.e., the coefficient estimated for interaction) is positive and significant. This finding suggests that an increase in client's structural embeddedness makes the high SE vendor derive more value from the joint match compared to low SE vendor. Similarly, the coefficients in the 6th row of the table (client PE row) are positive, which suggest that an increase in vendor's social capital measured by SE, JE, and PE is associated with an increase in the effect of one unit increase in client PE on match value. These findings suggest that a network of vendors and clients is more valuable as the size of their network grows (i.e., higher joined social capital is more valued) and that clients with higher social capital extract disproportionally more value from a vendor's network. As summarized in Table 1.1, firms with high SC are more capable of managing the development process, have higher abilities in dealing with outsourcing tasks, are more competent in meeting the project requirements, and are less likely to make costly mistakes, these result in a decrease in the ex-post transaction costs. Since these projects are collaborative by nature, in the presence of outside options, both sides prefer to work with other highly capable partners but sorting in the markets makes types match with partners who are similar in their network characteristics. This finding shows that high SC can impact ex-post TC by increasing firms' *ability* to coordinate knowledge exchanges by better communication and a common understating of what needs to be done and how it will be done.

The exception is the interaction of Vendor JE and client SE, the last coefficient in the 2nd column of Table 8b, which shows negative assortative matching. The coefficient is -0.252. The negative sign implies that lower SE clients derive more value from a one-unit increase in vendor JE. Also, negative estimate means that vendor JE and client SE are substitute attributes in the match joint value function. The finding could be because vendors with higher JE are selling in more niche markets that cover these low SE clients[3]. A higher value is created by keeping the vendor network open and less densely tied which can be achieved by matching with high PE low SE clients (i.e., clients connected to information hubs in the network).

All in all, I find evidence for the positive assortative matching in this market. The first set of findings show positive assortative matching on firms' network quantity and network quality. This finding signifies that vendors and clients derive significantly more match value if they are

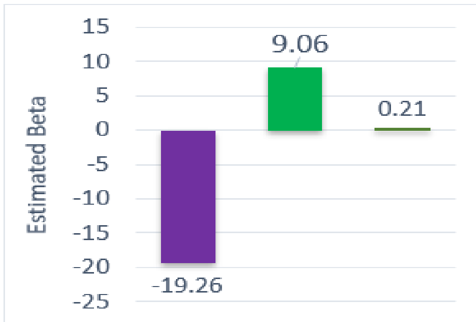
---

[3] Figure 1.4 is a good example to show the difference between SE and JE because there is good JE variation at similar levels of SE. The figure shows that while v2 and v4 have similar SE, v4 has a higher JE (37 compared to 29). It also shows V2 is covering more excluded markets. And although v3 has lower SE=2, its JE=37 is higher than v2_JE=29. V2 is covering niche markets compared to left part of the graph, where more vendors and clients are connected. V5 and V6 have very low JE because they are selling to clients who have themselves transacted with other connected vendors. An open and not densely tied network helps v2 have a higher JE than v5 and v6.

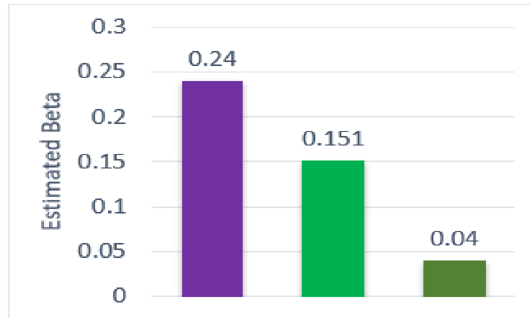
similar regarding network quantity. The similarity of network quality is also warranted, if statistically significant. Mismatches in firms' quantity and quality of network can be value destroying, highlighting that from vendors' perspective, the optimal value maximizing strategy to ensure growth is to position their products such that they appeal to clients who are similar to them in terms of their network strength. However, vendor network quality (JE) and client network quantity (SE) show evidence of substitutionary traits. The implication of this finding can be reflected in a vendor's decision to cover niche vs. general markets. Also, the parity of network in similar dimensions (vendor SE and client SE, vendor PE and client PE) and disparity of network strength in dissimilar dimensions (vendor JE and client SE) implies that market structure and matching process in the presence of outside options values vendors that emphasize network quality over network quantity. In general, marketing managers should note that the market I studied here rewards JE over SE and PE (more values are generated when JE increases or when a firm has high JE), and if this reward structure is stable over medium term, which is revealed to be the case in my long panel data, firms need to invest in projects that increase their JE in order to be more competitive in future and experience growth.

Conclusion 2. Next, I discuss how the client's non-network characteristics affect created the joint value in the match with vendors with different levels of SC. Specifically, I discuss the first four rows of Table 1.7 that include interactions between client size dummies/revenue and vendor SE, JE, and PE. Client size determines firms' *opportunity* to increase vendor transaction costs after the deal is made, by using "bargaining power" to take advantage of vendor lock-in and "agency costs" associated with managing such clients. Client revenue affects project size and is associated with higher levels of investments- in my data, this is reflected in higher relationship breadth (unique tasks in dyad) and higher relationship depth (i.e., cumulative expenditure in time). Firms' investments in a relationship are associated with firms' *motivation* to collaborate with their matched firm in future.

Figure 1.7 shows how the vendor network and client size determine match value. Purple, light green and dark green bars respectively show enterprise, large, and medium clients. The vertical Baseline cases have match values of zero and include matches with small clients. Coefficients show how much more (or less) match value is created as compared to the base cases. The positive coefficients imply that vendors with higher social capital extract more match value form selling to medium and large clients relative to small and enterprise clients, but the strength of effects differs. Increasing PE enhances the chances of a match with medium, large and small clients more than the match with Enterprise clients. Because with increasing PE, a firm can produce more value if matched with a medium firm. Each estimate should be interpreted as the value of increasing the social capital one more unit. The results as shown in Figure 1.7b indicate that increasing JE enhances a vendor's chance of matching with Enterprise, as for such matches one increase in vendor JE is estimated to bring an additional value of 0.24. Also, one can say a 10% increase in vendor PE increases the match value with a medium size client 30% more relative to the increase in match value with a large client. However, an increase in vendor SE has a different effect on match values: a 1% increase in vendor SE is associated with 42% increase in match value obtained from selling to a large firm compared to an increase in match value with a medium client. Findings show that larger scale and information reach can help large clients and vendors to extract higher match value (Beta=9.06 in Figure 1.7a), this may mean that successful development of services in larger firms is contingent on scale economies of the vendor.



(a) Network Size: SE



(b) Network Quality 1: JE

(c) Network Quality 2: PE

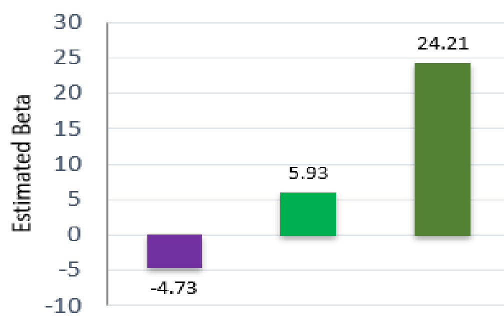


Figure 1.7: Effect of Client Size and Vendor Network on Match Value

The interpretation of negative betas is that the created match value is less than that of the base case. These values are conceptualized as benefits that accrue to firms above and beyond the transfers. The reduction in generated values can be due to incurred costs due to characteristics of the matching partner; transaction costs or operating costs could be among such value decreasing factors. Although the positive and significant coefficients for the interaction of client revenue and vendor social capital show evidence of strong assortative matching, the coefficients for enterprise clients show negative assortative matching with vendor SE and PE (coefficient -19.26 and -4.73 in Figures a.7a and 1.7c respectively). Vendors with high values of SE and PE have a disadvantage when targeting enterprise firms. The finding implies that an increase in client revenue enables vendors with higher social capital to extract higher match value compared to vendors with lower social capital. In the sourcing market, clients with higher revenue are considered low risk due to their better ability to pay; thus, they are considered more attractive (Ni and Srivinsan 2015). More outside options and lower search frictions due to lower search cost help clients with higher revenue reap more match value from the pool of high SC vendors. On the other hand, larger clients have higher negotiation power. Matching with enterprise clients can be costlier due to long delays in finalizing the deals, these costs show up in the negative coefficients. It is much easier for vendors' selling teams to approach smaller firms' top executives, whereas in larger clients there are more bureaucratic steps. Agency frictions and transaction costs are higher in enterprises compared to smaller clients, which can negatively impact the match value.

To sum up, in the presence of *transactional* and *behavioral* uncertainties in the software development process and the presence of *rivalry* in forming matches that create higher *values* for both sides, vendors and clients with similar network positions show evidence of higher joint match value relative to firms with dissimilar network positions. Results show that the market is not strictly positively assortative, as the highest matching value seems to be generated by matches that involve high JE, irrespective of the client's size and social capital. This pattern may suggest that high JE have high complementarities with other clients. It may also reflect the additional value generated by a wider portfolio breadth when firms collectively provide a comprehensive product that spans across markets. Pairwise stable equilibrium outcomes are positively assorted in IT outsourcing markets, which means that firms' differential traits are *complementary* in the production function, resulting in matching with similar agents (marriage markets; Becker 1973),

Table 1.8: Contributions and Results Implications

Theoretical Implications	Managerial Implications
• Extend TCE to include information spillovers. • Demonstrate the dynamics of past selections which serve as a stock of assets whose governance properties (through affecting reliability and certainty) extend beyond the focal match. • Motivation, ability and "*opportunity*" can be thought of as drivers of TC. • I characterize two-sided matches such that firm "bundle of traits" can be evaluated and valued, and I show that successful matches occur between "buyers" and "seller" who have *similar* traits in most cases.	• Firms "shared interests" (i.e., mutual investments) and firms' competencies (i.e., social capital) are not enough to avoid collaboration risks, and opportunities to cause inefficiencies (i.e., bargaining power and agency issues) can affect the value created in a match. • To be more competitive in the matching process, firms can *reposition* their network structure by watching the process and "matches made." ○ The market I studied here rewards JE over SE and PE (more values are generated when JE increases or when a firm has high JE), and if this reward structure is stable over medium term, which is revealed to be the case in my long panel data, firms need to invest in projects that increase their JE in order to be more competitive in future and experience growth. For example, in enterprise selling, high vendor JE (high information richness) more than offsets the agency costs associated with these sales.

as opposed to negatively assorted, which is observed when a negative correlation between traits of agents exists and sides match with dissimilar agents (corporate alliances; Amaldoss and Staelin 2010). Moreover, although synergies created from the similarity in network positions are valuable, *agency frictions* and *transaction costs* decrease the match value with clients who have high social capital, but also possess high levels of these costs. The policy implication of this finding is that removing these costs can significantly increase the joint match values that can be realized in markets. The summary of my contributions can be found in Table 1.8.

1.7 Model Comparison

Table 1.9 shows that the suggested model-which is called full model- has a 64.5% accuracy rate. This model includes "Network, Revenue, and Size" measures. Other applications of maximum score estimator reported similar accuracy levels (Banal-Estañol et al. 2018). *Model Score* is calculated by taking the percentage of pairwise inequalities that hold, given the model estimates (see ineq. 8). The pairs to the left of inequality (8) represent actual matches in the market, and the pairs to the right denote counterfactuals after considering pairs that switch partners.

Table 1.9: Model Comparison

Independent Variables	Score
Revenue Only	53%
Size Only	24%
Network Only	62%
Size + Network	62%
Full Model	64.5%

1.8 Simulating In-Sample Performance Measures

To evaluate the in-sample performance of my model, I compute *samematch*- the fraction of matches for which the realized match is the same as optimal matches implied by the solution to the linear programming problem, given the model's main estimates. Also, I calculate *percentoptimum*- a measure for % match value that observed matches generate relative to the match value created in solution to the linear programming (Shapley and Shubik 1971).

Table 1.10 shows that the actual matches produce 20-95% of the value of the matches that are created as the solution to the linear programming optimization model. For brevity, the results for four markets (out of 25 markets) are shown below:

Table 1.10: In-Sample Fit

Same Match	% Optimum	Observed Value (m$)	Optimal Value	Year
0.5	0.4289	226.50	528.11	2005
0.2	0.53349	206.95	387.95	2007
0.5	0.6884	258.24	375.12	2008
0.2	0.5692	3347.35	5880.46	2009

1.9 Simulating Out-of-Sample Performance Measures

To evaluate the out-of-sample performance of the model, I used pooled estimates from the main model to predict matches for system integration engagement in 2013. In this hold-out sample, there are 357 observations without missing data on predictable variables. Then, I compute *samematch* and *percentoptimum* at time T. Simulation steps involved in the LP model are as follows:

First, I calculate the payoff matrix. Within each simulated market, I calculate the total payoff for all the possible matches between a buyer and a seller using the pooled estimates from the last periods.

Second, I construct the matching problem in the current period using the payoff matrix calculated above and solve it by the linear programming for the selected market. Table 1.11 shows out of sample performance of the model for system integration task and year 2013. The out-of-sample performance of the model is worse than in-sample performance; other applications of the Fox matching model also report low out-of-sample fit results (Hortaçsu et al. 2015).

Table 1.11: Out-of-Sample Fit

Same Match	% Optimum	Observed Value (m$)	Optimal Value	Year
0.006	0.2942	435.2	528.11	2013

1.10 Other Checks

First, there may be a concern that the decision to form a match with a particular client is mostly driven by the presence of other vendors in the project rather than the preference for the client's characteristics. However, from the total sample of 49072 contracts, 9% are the consortiums between multiple vendors and clients and the majority of deals are between clients and primary vendors. Second, another critical issue which is worth considering is the number of mergers.

Mergers are important because they can impact firms' social capital by accessing the target's stock of information which can lead to changes in network positions. I observe 88 unique acquirers acquire 106 unique vendors. Table 1.12 shows acquisitions statistics based on the total number of contracts a target had signed before being merged; higher numbers indicate higher social capital of the acquired firm. As seen in the table, the largest merger in the IT outsourcing history is EDS acquired by HP in 2008; thus, if a merger changes network position of the acquirer, the strongest effect can be found by examining how HP network position evolves in time. Figure 1.8 shows the change in HP's standardized network measures from 1995 to 2013. I used standardized measures instead of absolute to see the change in HP's competitive position better. The change in firm JE is most pronounced after the merge. According to public announcements, HP acquired

EDS to increase its portfolio of services and become the largest service provider in depth and breadth of products. Its second objective was to expand the reach and penetrate to new markets. Before the merger, standardized SE, JE, and PE are 6.5, 6.5 and 2.8 for HP and 8.4, 8.2, and 4.3 for EDS. It takes four years for the merge to see the peaked values of 11.3, 12.5, and 6.4.

Table 1.12: Vendor Merger Statistics

Acquirer Company	N. Prior Contracts Before Merge	Percent
EDS (Electronic Data Systems Corporation) (acquired by Hewlett-Packard Company) (now HP Enterprise Services, LLC)	677	19.00
Science Applications International Corporation (SAIC) (SAIC, Inc.) (now Leidos Holdings, Inc.)	489	13.72
Atos Origin S.A. (now Atos S.E.)	430	12.07
Nokia Siemens Networks (NSN) (now Nokia Solutions and Networks)	322	9.04
Motorola, Inc. (now Motorola Solutions, Inc.)	203	5.70
SBC Communications (now AT&T, Inc.)	181	5.08
EDB Business Partner A.S.A. (Divested) (now EVRY ASA)	155	4.35
TietoEnator Corporation (now Tieto Oyj)	131	3.68
Andersen Consulting (now Accenture plc)	83	2.33
Other firms	892	25.04
Total	3563	100

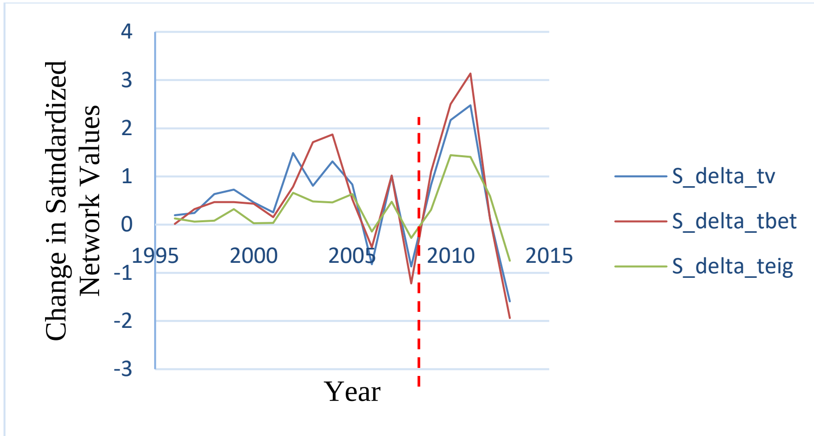


Figure 1.8: HP Network Evolution Before and After the Merge in 2008

Third, one can argue that service deals are awarded based on vendor's expertise in specific tasks, for example, a vendor may be more experienced in software products, whereas another may have higher social capital in markets for hardware. First, to the extent that I observe engagement types (IT outsourcing, System integration, business process outsourcing, etc.), I can control for vendor's task-specific knowledge. Second, data shows that on average, projects include 81% services (SD=26%), the median and mode is 100%.

1.11 Conclusion

Upstream channel structure has gone through considerable changes in the past two decades and designing proper marketing strategies is a crucial determinant of success and survival. The strategy design raises several research questions that have not received significant attention. First, a large body of literature on Transaction Cost Economics investigates contracting efficiencies to solve transaction uncertainty concerns. Transactions have been analyzed at the dyad level, and they have been assumed to hold the governance properties at the onset of the relationship throughout the partnership. Exceptions are Klein (1996) and Wathne et al. (2018), who use a two-wave survey to test if firms' governance choices have properties that extend over time. But to my knowledge, the assumption that past governance choices, outside the focal match, restrict and affect current transactions by mitigating conflicts and helping to realize mutual values have not been empirically tested. Setting the level of analysis at transaction ignores production costs and horizontal market effects such as knowledge/capability differences between firms, but information differences matter since firms operate in competitive markets and their competitiveness impacts their potential to overcome reliability and uncertainty concerns, which will be reflected in the final matching outcome.

Second, TCE assumes zero spillover effect. However, firms can redeploy the information and assets to their next projects and continuously seek to build up their stock of information through learning by doing and investments. Thus, they are likely to consider these spillover effects when deciding to work on projects. In this chapter, I have taken a step to bridge these gaps. I consider a model in which the "with whom" decision of the outsourcing and the pre-entry competition are addressed simultaneously by applying the interfirm network in a two-sided matching model.

From a practical standpoint, the question of network dynamics has important implications. Evidence of on-going network effects- as reported in the results section- means that a given network structure has properties of governance economies, by affecting competency and reliability concerns, thereby changing the created joint value in a match. Conversely, the lack of such effects could mean that the properties of the focal dyad, such as selection efforts or governance mechanisms employed to safeguard the focal relationship in the screening stage, are more relevant to answer the question of with whom to ally. The lack of aggregate network effect could also mean that these effects are time-related (i.e., decay faster than what I have assumed, or more recent networks explain the matches better than the aggregate network constructed for all data for all the time points). By filtering the clients who can communicate and meet the development requirements of projects, a vendor will be motivated to invest in value-enhancing activities that benefit both.

Sources of transaction costs for firms include partner ability, motivation, and opportunities to reduce conflicts. Ability-motivation-opportunity framework can be the other side of transaction and behavior uncertainty framework found in Table 1.1. The presence of transaction uncertainty (e.g., technological, requirement and requirement risks) impacts partners ability to harvest values from the match, and the presence of behavioral uncertainty (i.e., opportunism) impacts partners motivation to invest in a partnership and thus affects the value creation. Client's ability to cooperate and prevent termination has transaction cost reducing effects for a vendor. A client may be motivated to act well in a partnership due to its investments. However, this client may still lack the required competencies to manage the development process, or the organizational contexts provide it with opportunities to deviate. The lack of client competency results in lost efficiencies in the form of ex-post transaction costs.

In this chapter, I show that firms with higher firm social capital are more able to manifest specific behaviors that are conducive to the success of a collaboration. I also find that the size and revenue of the client predict their motivation to take advantage of the other party's lock-in conditions and increasing their transaction costs. A less motivated client does not have the power (e.g., bargaining power due to outside options or less reputational consequences) to exploit the vendor's investments. Thus, while higher social capital increases a vendor's ability to produce value in the match, an incompetent business partner can decrease the value by imposing inefficiencies.

Joint value creation can be compromised by firms' ability and motivation to collaborate during the project development process. Past selections motivate the vendors to invest ex-ante by having a solid base to rely on, created by client IT social capital, which ensures the firm has appropriate capabilities, is less likely to make costly errors and needs less corrective action on vendor's side. A high social capital firm is more capable of realizing the value from suppliers' investments, and thus supplier will be more motivated to do so, this will safeguard them ex-post (the governance effects of good selections are far-fetching and carry over time). Good selections help firms preserve and increase their social capital faster by avoiding lost efficiencies due to client inability thus decreasing value-creation risks. Resources can be spent where higher values are expected to be generated.

In summary, I show that how firm social capital, whose primary purpose is to guarantee the ability to collaborate, and firm size/revenue, whose primary focus is to ensure opportunity and motivation to collaborate, affect the matching outcome. Specifically, I investigate the effect of client and vendor interfirm embeddedness on the matching outcome, by including horizontal market factors, such as the presence of rival agents, and structural constraints that firms face, including the knowledge, information, and resources accumulated in a firm at a specific location in the network. My theoretical arguments suggest that in the "market for outsourcing partners," a process of sorting occurs whereby firms will tend to partner with other firms that are reliable and add value by improving their capabilities. This sorting is suggested to be partly based on network characteristics of the partnering firm and the *advantageous social capital* it provides. Considering these characteristics may help *reduce transaction uncertainty* caused by technological and requirement uncertainties, and behavior uncertainty caused by the geographically distributed environment and unfamiliar service providers (i.e., reliability and opportunism). This selection pattern of exchange is also, in part, driven by and leads to capability improvement over time. This selection pattern will positively impact firms' preference for selecting a firm that possesses technical capabilities such as better product mix, breadth, depth, and functional capabilities like an appropriate proposal, comprehensive contract, and good communication skills. Finally, the willingness of the partners to agree on contract terms and requirements also depends on the competition in the market since firms have heterogeneous levels of SC.

Equilibrium matches are the best each party in the dyad can achieve given the characteristics of the other buyers and sellers, and ex-ante these matches are optimal. I show that evaluating trading partners based on a network perspective provides an alternate metric for vendor evaluation. Using a comprehensive sample of IT outsourcing deals signed in years 1989-2013, I find evidence of a positive effect of social capital (embedded in network positions) on matching outcomes. The results suggest firms with equally competitive social capital tend to form the equilibrium matches because compatible levels of SC are revealed to produce higher joint value than matches between firms with dissimilar levels of SC. Specifically, a vendor with high SC is better off if it is matched with a client with a higher SC, but if a party increases her own SC, she can derive more match value from a specific matched partner compared to the party who keeps her SC constant.

The vendor social capital reduces the perceived contracting uncertainties, and positively affect the matching outcome, particularly by increasing the match likelihood with medium- and large-sized clients. However, the positive effect of social capital on matching decisions diminishes for both small and enterprise firms. These effects are best explained by transaction costs and agency frictions that are more pronounced in extreme-sized transactions. The traditional line against selling to small firms is that they are costly to sell to, for example, because they exhibit higher product requirement uncertainness (*hbs.org).* Also, according to hbs.org, the industry petitioners believe that creating the most appropriate product for the most demanding and bureaucratic clients can lead to long transition and production cycles. These costs negatively affect the generated joint value in the software development process. Removal of these costs can significantly increase the realized value from matches.

The finding that firm embeddedness explains who contracts with whom has other important implications. For vendors, results have implications for their resource allocation and targeting strategies that impact long-term outsourcing growth strategies and will help them win larger contracts. For the client side, results help them evaluate the competitiveness of the sellers, reducing the costs of soliciting bids from less promising suppliers. In this sense, a model based on a set of relationships which have been stable over some time is informative. This is because their network embeddedness and social capital determine how competitive they are willing to be in winning business with certain types of clients. From a marketing strategy perspective, "initial resource environments have lasting effects on subsequent organizational structures, processes, strategies, and survival likelihood" (Romanelli 1989). Winning larger contracts and successfully matching with large clients provides the required resources for further growth since firms' status in the network of clients and vendors is strategic and hard-to-imitate. For vendors, an implication from this study is the role of firms' network position in determining growth strategies, revenue generation efforts, and winning larger contracts.

CHAPTER 2

THE EFFECT OF PRICE PROMOTIONS AND TARIFF STRUCTURE ON REVENUE AND CHURN

2.1 Introduction

In wireless service industries, consumers choose a plan considering their expected consumption. Services have distinct characteristics of perishability in the sense that consumers cannot transfer unused allowances to next period. Many firms offer menu-based pricing and flexible contract terms to cater to customers with various preferences for usage and price. Pricing of these services with the objective of maximizing revenue is a crucial decision for many firms (Iyengar et al. 2011). Internet service industries are characterized by simple plans, two-part and three-part tariffs. In simple plans, a one-time fee is applied to gain access to subscription and unlimited usage allowance. In two-part tariff (2PT) plans, customers pay both a fixed subscription fee to get the service and a usage fee for each unit of consumption. In three-part tariffs (3PT) customers are entitled to specific usage allowance, and they are charged for usage in excess of that limit. An even more complicated offer is to bundle the tariffs with some compliments such as a modem, software, and new innovative services. A key underlying assumption of offering menu-based pricing is that consumers choose the plan that maximizes their utility. Overspending caused by being on low credit plans has been positively associated with customer churn in telecom industries (Lambrecht and Skiera 2006).

Few studies focus on the effect of pricing on post-purchase use. The evidence is mixed regarding the implications of each pricing format (2PT, 3PT, etc.), complex bundling and their influence on consumption and revenue (Shiv et al. 2005). In other words, it is not clear whether specific pricing and promotion formats will benefit the firm in the long-run. Limited attention has been given to factors contributing to increased or decreased usage under each contract setting with specific features (i.e., price, duration, complexity, time-pressure, etc.), though the inclusion of such information can guide firms in targeting right customers at the right time (Schweidel et al. 2011.) An extensive body of literature exists on predicting customer lifetime value and extend/churn decisions given observed usage/renewal behavior, yet few existing empirical works distinguish between the influence of selection and contract features on people's post-purchase behavior. Ascarza et al. (2012, 2015) document that contract features can alter consumer behavior in an undesirable direction.

While research on non-linear pricing uses marginal price perspectives to advocate the profitability of tariffs, works on consumer psychology find other determining factors unjustifiable by mere economic aspects (Bertini and Wathieu 2008). There are two schools of thought in the psychology of pricing regarding how pricing affects consumers' use of products and services. The first argument rests on the assumption that higher prices stimulate more use through sunk-cost effect (Thaler 1980; Ester 2002). An interpretation of this effect is that the more people feel out of pocket, the higher is the chance of using up the product or service that they have paid for (Gourville and Somen 2002; Ashraf, Berry, and Shapiro,2007). This argument implies that any pricing policy that masks the individual prices of items in a bundle or tariff decreases people's attention to consumption because it "imposes less pressure on people to get their money's worth." Gourville and Somen (2002) also argue that more usage is positively related to more repeat purchase.

The classic example of the sunk-cost impact of pricing on post-purchase usage is the case of health clubs and monthly or annual membership plans. It has been reported that people who are on monthly billing cycle are more likely to attend the gym regularly and this, in turn, increases the likelihood of extending their membership (DellaVigna et al. 2006). In practice, these duration-based pricing instances are abundant. For example, as shown in Figure 2.1, Grammarly.com website also uses different price promotions for monthly, quarterly, and annual plan. This research aims to answer how each of these pricing plans impacts consumer post-purchase consumption and churn decision.

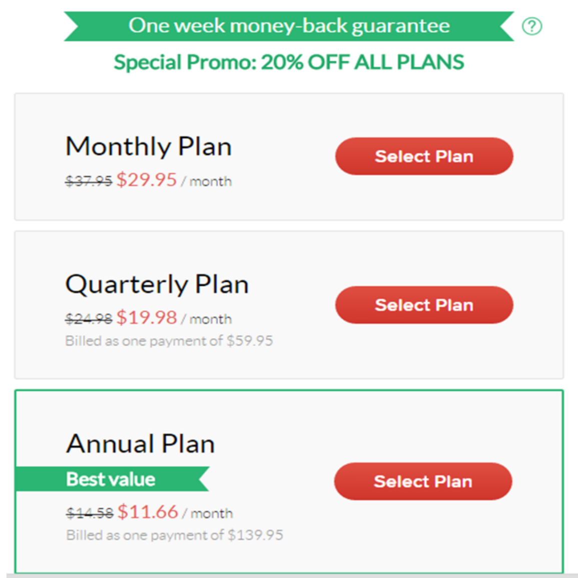


Figure 2.1: Grammarly.com Pricing Plans

Another interpretation of the assumption "higher prices lead to higher use" is that people with more expected use self-select themselves into higher-priced offerings (Oster, 1995). Friedman, Daniel, et al. (2007) state that laboratory experiment evidence is mixed with regards to the presence of the sunk-cost effect. Empirical evidence suggests that higher prices screen people with a higher value for the offering (Ashraf et al. 2007). As shown in Figure 2.2, increased price sensitivity of customers was among top three strategic pricing challenges faced by practitioners (RSR Report, April 2014); but retailers believe it's the pressure from price-sensitive consumers that make them more promotional and not the other way around. That is, as opposed to the finding in Ahsraf et al., (2007), anecdotal evidence suggests that pricing does not have the simple effect of screening customers based on their willingness to pay because price changes can change customer behavior as well.

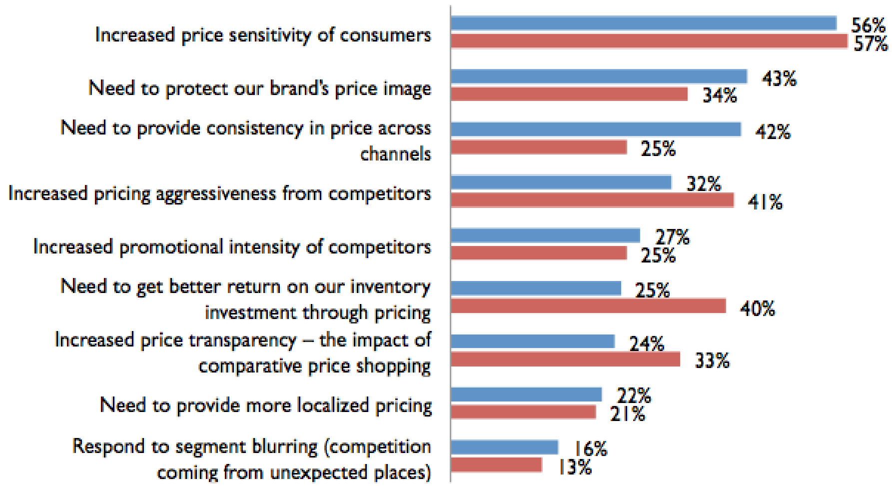


Figure 2.2: Top Three Strategic Pricing Challenges

As opposed to the sunk-cost interpretation of the mentioned assumption, other scholars believe that positive affective mood associated with free units in 3PT or bundles helps increase demand for purchasing extra credit; meaning that even after using up the allowance, positive mood spillovers cause consumers to demand more units (Ascarza et al. 2012). Authors found that three-part tariffs can increase consumers' consumption and firm revenue because consumers value free allowance as a gift and positive affective mood associated with free items in tariff increases their consumption. The effect of the positive mood generated by free gifts has also been demonstrated by Shampanier et al. (2007). As it seems, the placebo effect of pricing-i.e., higher prices leading to more consumption- which is investigated and demonstrated in previous research can be at odds with theories that support the positive impacts of bargains.

Ashraf et al. (2007) find that the willingness to pay (WTP) is a stronger predictive of use than observable household characteristics. If the screening effect holds, I should see consumers who are willing to pay higher prices are more likely to consume. These researchers vary offered prices to different groups and also offer random discounts to customers who have already agreed to pay a higher price to investigate whether WTP is more predictive of use than sunk cost effect. Elsewhere, other scholars apply a similar two-stage experiment design to distinguish between sunk cost and screening effect (Arkes and Blumer, 1985; Karlan and Zinman 2006). In Arkes and Blumer's (1985) research conducted on 60 observations with hypothetical choices, "discounts" play the role of transaction price in Ashraf 's design. Eyster (2002) reports that Arkes and Blumer's (1985) experiment is the only field study of sunk cost effect. Two possible explanations are justifying that higher offer prices lead to more use. First, full prices compared to discounted prices screen people based on observable covariates. Second, full prices select consumers with higher discounted utility and expectation of future use. Identifying this is important both from a theoretical perspective as it yields insight into the underlying mechanism of price promotion effects and from an empirical viewpoint since it helps us target customers based on readily available observables. In this research, I will proceed with the first explanation, and I will assume selection is based on observables.

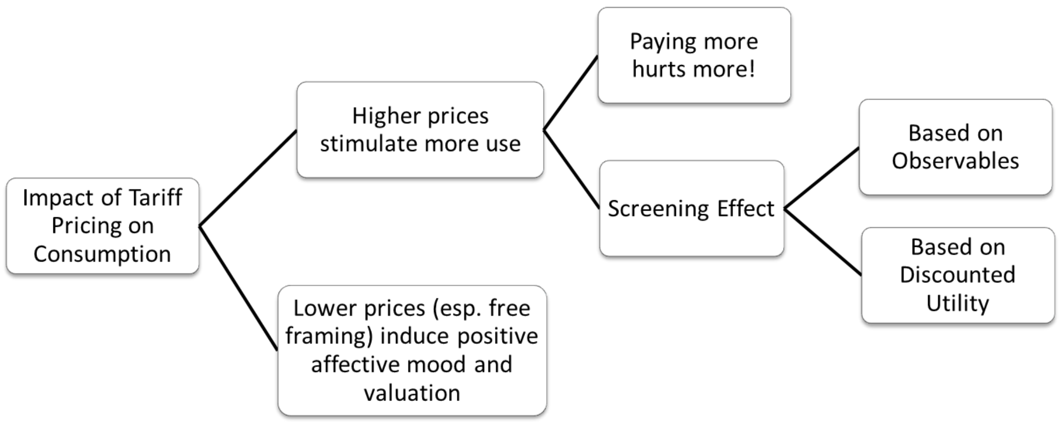


Figure 2.3: Tested Theories of Pricing Psychology

Price promotions have been largely studied by researchers, and many benefits and disadvantages have been reported (Gupta1988; Neslin 2002). For instance, price promotions attract new buyers and make consumers feel good about themselves (Lee, Leonard, and Claire I. Tsai,2014). Many firms combine price promotions and bundles to attract more new customers, to encourage existing ones to switch to longer contracts and to secure a longer financial flow of customers. In this study, I investigate whether promotional bundle (PB) buyers post-purchase consumption behavior can be explained by screening or contract features. I also aim to identify whether the sunk-cost effect or positive mood associated with discounts explains the underlying mechanism. To the best of my knowledge, no empirical research has answered these questions using a large field experiment. The answer to each question can have theoretical and managerial policy implications. Figure 2.3 shows a summary of the theories tested in this chapter.

In my first study, I use data collected from a natural field experiment done by an ISP to compare purchase behavior of two groups of customers: the first group (treatment) are exposed to promotion bundles and the second group (controls) are not. I intend to examine which of the pricing effects mentioned above prevails in the context of PB offerings in the wireless industry. I find that even after controlling for selection, promotion bundle buyers spend less compared to their matched counterparts in the control group who are not exposed to promotion. Specifically, results show that these promotions invite current customers who are already aware of their consumption preferences to switch. Still, controlling for selection does not rule out all the differences between control and treated group. Hence, I conjecture contract features (favoring sunk-cost assumption) and the mere act of offering PBs change customers' post-purchase behavior. Contract features that mask the pain associated with payment (i.e., complexity, length, etc.) significantly affect post-purchase consumption.

In their paper, Iyengar et al. (2007) report that with learning consumers leave less unused utility on the table and this phenomenon increases customer retention and CLV. While I find evidence in favor of some learning effects on PB choice in this context, results also show that PB buyers churn significantly more compared to consumers with similar matched observed characteristics in the control group. I postulate that exposure to PBs can lower customers' inertia. Once customers become aware of saving plans within the company, they might also search for better offers of competitors (Ascarza et al. 2015). This hypothesis is in line with Mela, Gupta, and Lehmann's work (1997) that suggests price promotions can have adverse effects on demand by lowering price expectations and increasing price sensitivity. The negative effect of offering PBs is reinforced if the firm does not have an attractive plan for customers who are at the end of their promotion contract. Lowering inertia mitigates the effect of self-selection to cost-saving plans; hence these customers should be targeted with different attractive retention plans before their promotion ends.

I contribute to the literature by ruling out the positive role of emotions associated with promotional bundles on consumer's post-purchase usage and by explaining what behavioral factors affect consumer's post-purchase consumption. I also incorporate customer heterogeneity in responding to promotions and tariff design into my model. My model accounts for customer heterogeneity by including "customer tenure" and "level of price-sensitivity," comparing the behavior of old and new segments as well as high and low price-sensitive ones. Promotions are offered to increase market share and to encourage current customers to switch to long-term plans to secure financial flow for a longer term. However, I believe offering promotions can lower existing customers' inertia and increase their churn rate especially in the segment that proves to be more careful in choosing cost-minimizing plans. This effect can be very important in markets where people are already very inertial, but the firm's offering of promotions can make them aware of the possibility of searching for better plans with other firms. Increasing customer base and decreasing customers' inertia is a trade-off that needs to be considered when offering and pricing promotion bundles and tariffs. Finally, I suggest how firms should utilize the information contained in these self-selected customers to design retaining policies as well as new post-promotion packages. To my knowledge, my study is the biggest natural experiment to date to test for the presence of behavioral pricing effects.

2.2 Research Setting

To assess the effect of offering promotion bundles on revenue and customer churn, the desired research setting needs to satisfy several stringent conditions. First, I need a sample of customers who are exposed to promotions (treatment group) and a group who are not (control group). Second, I need to have enough observable variables that explain the choice of bundles among old users in the treated group. Lack of enough explanatory variables leads to misleading purchase propensity scores (PS) estimates. Third, I need to ensure I have enough "overlap," i.e., the distribution of covariates is similar for treated and control group. I will elaborate on these conditions using technical expressions when I present my model.

I utilized a unique dataset from a wireless communications company in Iran to meet these criteria. Due to license and law restrictions, this company does not offer promotional bundles in some specific centers throughout the country. This natural variation in data helps us proceed with experimental techniques to estimate the impact of such offerings on customers usage and churn behavior. Managers offer discounted bundles with the hope of attracting new customers and retaining the existing ones for a longer time. However, there is uncertainty regarding how customers' behavior is affected by such offering and what needs to be considered when designing new offerings when promotions end. It is also not clear how existing customers are affected by exposure to promotions: Will offering constant promotions trigger price comparison search and make customers switch to competitors by making them aware of promotional packages? Do low price-sensitive and high price-sensitive customers react similarly to exposure to promotions that are always available for every customer? These are some of the questions that I investigate in this research.

2.2.1 Natural Field Experiment

Promotional Bundle: A specific promotional bundle that includes both complement and substitute items were selected for this study. This bundle is comprised of a 12-month annual subscription to 8 Mbps Internet, monthly credit allowance, free time-constrained credit (TCC)- each unit of this should be used within each day- and Anti-virus software. The company follows mixed bundling strategies and items are also offered individually in the company's product assortments. The difference between these two kinds of credit is that the first one allows for a much longer period for use, but the second one expires after a few days of using it. Because of its time-pressured nature, it is priced much less than other credit. Managers conjecture this type of credit changes consumption pattern and make people buy more credit in excess of their allowance in the long run. I test this hypothesis in section 5.

Competitive Environment: Most Internet Service Providers (ISPs) offer the same services with similar features and pricing plans. TCC is the innovation of this particular company and serves to attract users with more variational usage, for instance, movie watchers or game players who download films or play games at random times. There is almost no price variation in plans offered by firms in this industry within my observation period, and services features are fixed; this enables researchers to estimate revenue elasticities of plan features and usage.

Customers: Customers can be heterogeneous in responding to marketing efforts. Identifying characteristics that determine customers' reaction to promotions is important from an empirical standpoint since it helps us understand how these activities affected customers and how I can design future promotional policies. Customer relationship management (CRM) stream of research suggests that specific consumer characteristics inform us most about their possible reaction to marketing efforts and their demand for services in contractual settings. For instance, Lemmens and Croux (2006) list change in monthly minutes of use, the base cost of a plan, total revenue over life and mean overage revenue among most important predictors determining renewal or churn decisions. Overage is defined as extra credit purchased in 3PT plan when usage allowance is exceeded. The authors also find evidence that decreased trend of consumption increases propensities of churn. Ascarza and Hardie (2013) also find similar results by using a joint model of usage and churn. Ascarza et al. (2015) use three main observables to segment customers: 1) Overage, 2) Coefficient of Variation in individual usage during months before treatment, 3) trend of revenue measured by percentage change. Train et al. (2007) report that consumers often make mistakes in choosing the proper plan that maximizes their utility considering their future use and they find evidence of overspending. In my context, this overspending is captured by the variable "overage." Overspending is generally not good for customers because overage should be purchased at higher rates than default allowances.

2.2.2 Data

This study draws its data from a primary source of information: click-stream data stored in the customer database of an ISP firm. I reshaped data to common panel formats to perform the analysis. Users are identified by unique ids. For each user, I observe the following: 1. Tariff choice (Services and promotions), 2. Service features (price, duration, default allowance, speed, etc.), 3. start, extend, change and churn decisions and their timing. Monthly revenue can be broken down into three subcategories: a) tariff fee, b) overage sum (extra credit purchased), c) time-constrained overage sum. Table 2.1 shows the descriptive statistics of two groups: control groups (excepted centers who do not see promotion bundles) and treated group (who see all PB).

I choose bundle buyers with long enough history for my analysis to get more precise estimates of explanatory factors for bundle choice. To this end, I omitted new customers with no observed covariates and selected treated customers who purchased at least one plan during the observation period.

Table 2.1: Descriptive Statistics of Control and Treatment Groups

Exception	N. Obs	Variable	Mean	Std Dev	Min	Max
0	101,783	Total Sum (IRR)	1,691,920	1,670,624	0	258,134,501
		Total Freq	16.88	19.69	1	495
		TotalOverageFreq	13.74	18.62	0	492
		TotalOverageSum (IRR)	869,941	1,071,728	0	44,751,002
		Churn Rate	0.08	0.27	0	1
1	20,112	Total Sum (IRR)	2,049,177	1,540,942	94,500	52,101,001
		Total Freq	19.4	20.25	1	314
		TotalOverageFreq	14.9	19.24	0	313
		TotalOverageSum (IRR)	956,519	1,197,300	0	50,310,001
		Churn Rate	0.06	0.24	0	1

a Note 1: IRR (Iranian Rial) is the official unit of Iranian currency.

b Note 2: Exception coded “zero” represents customers in exposed (treatment) group.

There are 101,783 treated and 20,112 control customers. Descriptive statistics show that treated customers (exception equals 0) are less valuable in terms of the total sum paid for plans and overage. They are less frequent for purchasing plans, but they churn more. Testing for overlap is complicated in this context because I see the PB for the treated groups during all months of observation period; as a result, I cannot check for the similarity of covariate distributions before introducing the promotions. However, there is no behavioral targeting for these promotions. Customers are not assigned to these groups based on marketing related indexes but are grouped based on restrictive regulatory laws. In some areas of the country, offering specific services are prohibited according to regulations. Hence, I can assume there was enough overlap between these two groups before exposure to treatment.

2.3 The Impact of Promotion Bundles on Customer Behavior

In this section, I analyze the impact of treatment on two important outcomes:1) usage and revenue (total money paid for access fee, allowance, and overage), 2) churn. First, I estimate the average treatment effects of PB on customer behavior. Second, I explain possibly heterogeneous behaviors among treated customers. Finally, I suggest explanations for the observed results.

2.3.1 Aggregate Effect of Promotion Bundles

If I only compare treated PB buyers and treated non-PB customers in the first group and do the analysis in the absence of those who were not exposed to bundles, I would find that bundles increase revenue in forms of the total sum paid by customers and increase customer loyalty exhibited by less churn rate (Table 2.2). Consequently, I would wrongly infer that PBs are successful in meeting marketing success indicators (revenue, customer retention). However, I need to take into account that PB buyers may have self-selected themselves into promotional plans and ignoring self- selection can considerably bias my estimates of any of the outcomes mentioned. It is also possible that the action of offering promotions changes the sensitivity of customers to prices and increases their propensity to search for better alternatives elsewhere. If I only had one group, separating these possible effects could be complicated; however, with the help of a control group, I can identify each of these effects.

Table 2.2: Comparing PB-59 Bundle Buyers with Non-Buyers in Treatment Group

Group	N. Obs	Variable	Mean	Std Dev	Minimum	Maximum
Control	101,783	Total Sum (IRR)	1,691,920	1,670,624	0	258,134,501
		Total Freq	16.9	19.7	1	495
		TotalOverageFre	13.7	18.6	0	492
		TotalOverageSu	869,941	1,071,728	0	44,751,002
		Churn Rate	0.079	0.270	0	1
Treatment	750	Total Sum (IRR)	3,068,843	1,041,289	2,100,000	9,124,002
		Total Freq	12.4	15	1	146
		TotalOverageFre	11.3	15	0	145
		TotalOverageSu	917,456	996,759	0	7,024,002
		Churn Rate	0.016	0.126	0	1

A possible way to control for selection issue here is using a matching technique. Matching is used to process data to improve causal inferences in observational data (Ho et al. 2007). According to the authors, there are two possible approaches to follow after matching. First, I can compare differences in means without controlling for confounding variables. Second, I can control for such confounds assuming that matching may not have been done precisely and proceed using a parametric form to improve the causal inference. In this research, I use both approaches and compare the results.

2.3.2 The Model

I use propensity score matching (PSM) to control for possible self-selection in choosing the promotion bundles. The propensity score was first introduced by Rosenbaum and Rubin (1983) as a tool to help infer the causal effect of marketing interventions. According to the authors, the propensity score is the conditional likelihood of selecting a treatment given observable covariates; thus, this method helps us eliminate any bias present due to observable predictors. Propensity score has been used in marketing literature to compare sample responses in “the effect of CRM programs on customer knowledge and satisfaction" (Mithas et al. 2013), "DVR adoption" (Bonnenberg 2010) , "impact of online channel use on increasing customer revenue and decreasing cost to serve" (Gensler 2012), "the effect of referral program participation on loyalty of existing customers" (Garnefeld, et al. 2013); to name just a few.

Rubin and Waterman (2006) state that few empirical papers study the effect of promotions, advertising and other marketing efforts using correct empirical methods to infer causation and the use of traditional techniques such as data mining and least square estimation on data leads to the wrong conclusion. There are two general approaches to propensity score matching: a) match samples based on single score, b) cluster samples into subgroups with close propensity scores and then calculate the average treatment effect. In this research, I apply the second method.

I proceed in three steps: First, I consider all customers in the treatment group and estimate the probability of buying the promotion bundle based on observed covariates by running a probit model. Second, using the calculated parameter estimates, I calculate the probability of buying the promotion bundle for all customers, including those in the control group who do not see the bundle. Third, I estimate the purchase behavior of customers in control and treated group by using a regression adjustment with propensity scores of purchasing the bundle. In this approach, I regress the outcomes on a dummy indicator for receiving the treatment, estimated treatment scores and the interaction of purchase probabilities and treatment. The effect of treatment is determined using estimated regression coefficients with treatment and scores covariates. Austin (2011) explains why the use of regression adjustment with propensity scores is preferred to simple regressions when working on observational data to estimate treatment effects. Specifically, he states that it is easier to determine if the model using PS is well specified than assessing adequate specification of a regression model with observational covariates and treatment indicator as in the latter case R square does not testify whether the outcome model is well specified.

3.2.1. Stage 1: Propensity Model. For customer i, let Buy_i be an indicator variable that takes the value of 1 if a consumer purchases the promotion bundle and 0 otherwise. Then, the probability of selecting the promotion bundle is modeled using pooled probit as

$$prob(Buy_i | \alpha, \beta, X_i) = Prob(\alpha + \beta X_i + \epsilon_i), \quad (1)$$

where α is a constant, X_i is a vector of user-specific covariates and ϵ_i is i.d. ~ N (0, 1). I leveraged past studies on customer relationship management that focused on usage and renewal decisions in contractual settings (Kumar et al. 2014; Ascarza and Hardie 2013). The vector X_i contains customer tenure duration (months passed since starting a contract with the firm), activity duration (time difference between minimum and maximum month of activity within observation period as a proxy to control for learning effect of repeated purchases), plan frequency, overage frequency, a dummy variable indicating whether customer has purchased time-constrained credit-(assuming that the presence of this type of credit in the promotion bundle possibly attracts this segment), and coefficient of variation for plan and overage sum and frequency.

I hypothesize that more active customers are more likely to update their usage beliefs through repeated decisions they make each month. Plan frequency is the number of times a customer extends, changes or signs up for a new plan. Note that activity duration is different from frequency measures since frequency can be high in one month but zero in the others. In this case, activity duration will be recorded zero. Collinearity diagnostics inflation index (4 and 5 for duration and frequencies respectively) also show that frequency and activity duration are measuring different explanatory factors. If the coefficient for activity duration is positive, I can say that more aware customers will choose this bundle. Table 2.3 shows parameter estimates of the propensity model. I tried many transformation and interactions and chose the model with the best fit. The fit of the model was checked using multiple fit measures produced in SAS output (MC Fadden R square, Veall-Zimmermann, etc. and in most cases the value was higher than 0.9).

Table 2.3: Propensity Model Results

Variable	Estimate	Standard Error	t Value	Pr >|t|
Intercept	31.513	0.192	163.75	<.0001
Tenure	0.001	0.016	0.07	0.944
Activity duration	0.590	0.103	5.71	<.0001
Overage freq	-0.256	0.063	-4.06	<.0001
Plan freq	-35.987	0.192	-186.99	<.0001
TCC treat	1.084	0.368	2.95	0.0032
Overage freq cv	-0.008	0.013	-0.61	0.541
Overage sum cv	-0.016	0.012	-1.35	0.177

Estimates show that activity duration is positively associated with the propensity to purchase the PB and total plan frequency significantly reduces the likelihood of PB purchase. The latter result has proper justification because as shown in Table 2.4, a high percentage of customers (%57) is inertial in choosing short-term monthly or quarterly contracts (compared to medium contracts with a duration of 3 or 6 months and long contracts with a duration of 12 or 15 months). That is, within my observation period, they always choose short-term plans. Dual types are switchers and are of mixed types. Activity duration estimate can be learning effects as mentioned earlier. My interpretation of this parameter is consistent with prior research that has found consumers learn about their usage within nine months (Iyengar et al. 2007). Since these long-term bundles require commitment to the firm, cost-minimizing customers must have enough knowledge of their expected usage in order to select these long contracts or they could either overspend by underestimating their usage (accepting lower monthly allowance) and being forced to spend more on overage or they could leave unused utility on table by having signed up for more expensive contracts (they do not have the opportunity to downgrade as short-term customers). Consistent with what I had conjectured, those customers who have previously purchased TCC (denoted by "TCC treat" dummy in Table 2.3) are more likely to buy this promotion bundle because they get increased utility from consumption of this item.

Table 2.4: Descriptive Statistics of Customer Types

Customer Type	Frequency	Percent
Dual	11,573	6.35
Long	50,553	27.74
Medium	15,787	8.66
Short	104,317	57.24

3.2.2. Stage 2: Computing Propensity Scores. At this stage, I used probit estimates of the propensity of buying for both control and treatment groups as calculated by equation (1). I calculated the propensities by multiplying probit estimates and X co-variates and then added estimated intercept. Since I used a probit model, I cannot directly get the propensity scores. I computed the normal CDF of the sum of intercept and X-Betas and used it as propensity scores for all customers including both control and treated group. The formula I used is as follows:

$$prob(Buy_i|\alpha,\beta,X_i) = F(\alpha + X'\beta), \quad (2)$$

where F is the cumulative distribution function (φ).

3.2.3. Stage 3: Effect of promotion bundles on Revenue and Churn. Finally, I used two different methods to estimate the effect of promotion bundles on revenue and churn. First, I ran an adjusted regression model with treatment dummy and propensity scores on both treated and controls; second, I used the average treatment effect model to estimate the same effect. Both the average treatment effect in Stata and regression adjustment in SAS had some interpretable results in the same direction. The revenue generated from bundle buyers and counterfactual "potential" buyers (i.e., controls) is significantly different, so is their churn rate and overage frequency.

Method 1: Adjusted Regression Model with *Propensity Scores.* At this step, I include all observations from control and treatment groups. In the case of continuous DV such as total plan sum and overage revenue, I describe the regression model below as

$$Revneue_i | \alpha, \beta, X_i, Score_i = \alpha + \beta T_i + \gamma X_i + \phi\, Score_i + \delta\, T_i Score_i + \epsilon_i, \qquad (3)$$

where T_i is a dummy variable that takes the value of 1 if the customer is treated (exposed to promotional bundles), 0 otherwise, X_i contains all observable characteristics (tenure, activity duration, frequencies, variability measured by CVs), $Score_i$ is the individual's propensity to purchase the bundle and ϵ_i is the error term normally distributed with mean 0 and unit variance. In regression equation (3) i indexes individuals from both treatment and control group. Table 2.5 shows the main results after running the regression in equation 3. Coefficient of variation is calculated by dividing means of parameters to their standard deviation. Finally, δ is the parameter I am interested in since it measures both treatment and promotion bundle propensity effect through the interaction defined. This parameter captures the impact of purchasing promotion bundles on revenue generated after controlling for selection.

Table 2.5: The Effect of Promotion Bundles on Customer Revenue and Churn

Variable	Total Sum	Overage Sum	Churn Odds Ratio
Intercept	-634,707***	-54,356***	-
Treat	-18,417*	-67,922***	1.026
Propensity scores	1,467,484***	364,031***	0.32***
Treat * Propensity Scores	-767,968***	-247,793***	3.133***
Tenure	1,966***	3,058***	0.985***
Activity Duration	84,081***	73,153***	0.761***
Total overage freq	30,351***	35,971***	1.004***
Total plan Freq	240,150***	-37,951***	0.984*
TCC treat	239,802***	144,230***	1.231***
Overagefreq cv	-2,008***	718***	1.002**
Planfreq cv	2,961***	-3,744***	0.999
Overagesum cv	-630**	-1,577***	0.999
Plansum cv	1840***	4179***	1.001

I checked the fit of the first two models that are shown in Table 2.5 by adjusted R-squares (0.45 and 0.56 for the first two models). Then, I use a logit model with the same covariates as in the regression revenue model to estimate the impact of offering promotion bundle on my discrete independent variable. The probability of churn is modeled as:

$$prob(Churn_i|\hat{\varphi}, T_i, X_i, Score_i) = prob\left(\varphi^0 + \varphi^T T_i + \varphi^S Score_i + \varphi^X X_i + \varphi^{ST} TScore_i + \varsigma_i\right), (4)$$

where φ^0 is constant and ς_i follows logistic distribution. Other parameters are the same as variables used in equation 3. Both the control and treatment group are included in Table 2.5 estimations. In calculating the predicted probability of promotion purchase, I used past observed behavior of buyers prior to the time of purchase. However, in running adjusted regressions, I considered their behavior after the bundle purchase.

Adjusted Regression Model Results. The values estimated for the interaction term in Table 2.5 are most important to us. The results show that even after controlling for selection on observables (by inserting "propensity scores" as individual variables in the regression), promotion bundle buyers are much more likely to churn (odds ratio 3.13) than people with the same level of price sensitivity[4] in control group and are less valuable both in terms of plan sum (-767,968 IRR) and overage money (-247,793 IRR) paid for extra credit. The odds ratio for variable "Treat" is 1.026, which means ceteris paribus, exposed customers are 2.6% more likely to churn. On average, this percentage of churn accounts for 47,055,098 (I RR/Month)[5] revenue. The saving from this leads to a significant improvement in profit in the firm's financial statements. Customers with higher propensity scores are less likely to churn (odds ratio is 0.32). Note that this includes propensity scores for both treatment and control groups. However, when I take the interaction of treatment (exposure to promotions) and promotion bundle propensity scores into account, I observe a drastic increase in odds ratio (3.133).

These findings are contrary to the descriptive results of Table 2.2 in which after comparing promotion bundle buyers and non-buyers in the treated group, I observe buyers are more valuable and churn less. This warns of possible misleading interpretation of treatment effect while the only treated group is considered. Holding all other things constant (by controlling for tenure, activity duration and etc.), customers who are exposed to promotions and are more price-sensitive(higher

---

[4] Throughout this research, I assume purchase propensity of promotions is a proxy for price sensitivity

[5] Hereafter I use 'Mo' instead of 'Month'

propensity score) are two times more likely to churn (odds ratio 3.13) than those who are not exposed to promotions but have the same level of price sensitivity. Results show that promotion bundle buyers spend significantly less (-767,968 IRR) on purchasing extra credit compared to their matched counterparts. The treated group is less profitable than the control group in terms of the total sum. If after controlling for screening effect I find that lower prices lead to lower usage (hence less demand for extra credit), I can discriminate between the effect of screening and sunk-cost effect in customers' post-purchase usage behavior. The level of sunk cost is determined by the form of contract features and tariff structure (i.e., pricing, duration, complexity). This implies more expensive, less complex or longer contracts lead to more sunk-cost considerations present in individuals' decision making. I can test for this implication using estimates presented in Table 2.5.

The higher propensity score is positively associated with higher monetary value and much lower retention rate; however, if I include interaction term (treat × propensity score), I see significant negative values for the plan and overage sum paid by customers and signification increase in the odds ratio for churn. That is, I find that customers who are not exposed to promotions spend more on overage and plan purchase even after I control for possible selection of more price-sensitive customers in the treated group who are exposed to promotions. This is evidence in favor of sunk cost effect: paying more leads to more consumption. I postulate promotion affects churn and decreases retention rate beyond the level that can be explained by customers' intrinsic preferences and base price sensitivity. Not all customers are price-sensitive, forward-looking and cost-minimizing to the same degree. If I assume that more price-sensitive customers buy the bundle, and then if I match customers in treatment and control group based on observables that predict bundle purchase (i.e., if I believe selection is based on observables) then

I can compare the effect of exposure to promotions. I conclude that constant exposure to random promotions increases churn rate and decreases revenue at the cost of attracting new customers. It is a trade-off and firms need to be more aware of the prices they set because it fuels consumers' sensitivity to better deals.

Method 2: Population Average Treatment Effect. In the adjusted regression model with propensity scores and treatment dummy, I consider the effect of a specific promotion bundle for analysis. However, in this section, I consider all promotion bundles that are offered and estimate their average effect on customer revenue and churn. If I have data on potential outcomes of treatment for both control and treated group, a simpler approach to estimate the effect of treatment is comparing the means of two groups. In this approach, instead of comparing outcomes of matched users in treatment and control group using PB propensity scores, I compare aggregate treatment effect (ATE) in two groups. Potential outcomes are the data on all observables that help us estimate a causal relationship. To begin with, let's assume I run the simple regression below to get the ATE:

$$Y_{i1} = \alpha + \beta\, T_i + \varepsilon_i\ , \qquad (5)$$

where

$$T_i = \begin{cases} 1, & if\ exposed\ to\ promotion\ bundles \\ 0, & otherwise \end{cases}$$

and

$$Y_i = \begin{cases} Y_{i1}, & if\ exposed\ to\ promotion\ bundles \\ Y_{i0}, & otherwise \end{cases}$$

In my research, I am interested in the population average treatment effect for the treated:

$$E(Y_{i1} - Y_{i0} | T_i = 1), \quad (6)$$

which measures how the same people would behave with and without the promotion bundles, but even in my randomized experiment, I cannot observe $E(Y_{i0}|T_i = 1)$. Thus, it is common to use equation (7) to estimate ATE:

$$ATE = E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 1), \quad (7)$$

Sul (2015) shows the bias and inconsistency of the formula above in estimating ATE. Specifically, I have:

$$E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 1) = E(Y_{i1} - Y_{i0}|T_i = 1) + E(Y_{i0}|T_i = 1) - E(Y_{i0}|T_i = 0)$$

$$= ATE + [E(Y_{i0}|T_i = 1) - E(Y_{i0}|T_i = 0)]$$

$$\neq ATE$$

If treatment effects are not endogenous, there will be no bias. Under endogeneity, I do not have conditional independence, and unobserved variable affects both treatment and outcome. In my case, treatment takes value 1 when customers are forced to be in the exposed group and 0; otherwise. Table 2.6 shows the estimates of promotion exposure effect using equation 11. I used random effect estimator and assumed promotion offering in one group is independent of potential outcomes (revenue, churn).The solution to inconsistency is including control variables, X_i, in the regression (5) as

$$Y_{i1} = \alpha + \beta\, T_i + \gamma\, X_i + \varepsilon_i \quad (8)$$

However, under unconfoundedness, the OLS estimation of (8) becomes consistent since I have

$$E(Y_i|X) = E(Y_i) \qquad (9)$$

In my population ATE model, I run the following model:

$$Y_{i1} = \alpha + \beta\, T_i + \varepsilon_i \qquad (10)$$

Table 2.6 presents the results of estimations for equation (10).

Population Average Effect Results. As results show, compared to customers who do not see the promotions, the exposed group is much less profitable in terms of the total sum (plan, overage, etc.) and 1.7% more likely to churn. However, this does not take into account any heterogeneity in response to exposure and promotion purchase probability. To take this simple model further, I use clustered propensity score matching as mentioned in section (3.2).

Table 2.6: Population Average Effect of Exposure to Promotions

	Coef.	Robust Std. Err.	z	P>z
Total Sum				
Treat (1 vs 0)	-347,185	12,054	-28.8	0.000
Pomean Treat 0	2,049,177	10,865	188.6	0.000
Total Overage Sum				
Treat (1 vs 0)	-86,231	9,081	-9.5	0.000
Pomean Treat 0	0.017	0.002	9.04	0.000
Churn				
ATE Treat (1 vs 0)	0.017	0.002	9.04	0.000
Pomean Treat 0	0.062	0.002	36.32	0.000

Method 3: Cluster Propensity Score Matching. In this section, I use the predicted probabilities of promotion bundle purchase (propensity scores) and treatment/control dummy to see the impact of exposure to promotion bundles on the behavior of two groups. To see how the average treatment effect is identified in this method; I utilize two assumptions of "unconfoundedness" and "overlap" as discussed in section (2). If I define the *average treatment effect conditional on x* as:

$$\tau_x = E(Y_{i1} - Y_{i0}|X = x) = \mu_1(x) - \mu_0(x),$$

where Y represents our potential outcomes (revenue, churn, etc.) and $\mu_g(x)$= E(Y_g|X = x), g = 0, 1. In this part, I am conditioning on propensity scores for purchase of promotion bundles.

The function τ_x represents customers heterogeneity (described by estimated PS) in response to the effect of promotions. This function by itself sheds light into how different segments of people with various promotion purchase propensities respond to exposure to promotions in exposed and not exposed group. By iterated expectations and without any assumptions, we always have

$$\tau_{ATE} = E(Y_1 - Y_0|X = x) = E(\tau_{PS}) = E(\mu_1(PS) - \mu_0(PS)).$$

It follows that τ_{ATE} is identified if $\mu_0(PS)$ and $\mu_0(PS)$are identified over their support of PS. To see μ0(.) and μ1(.) are identified under unconfoundedness and overlap, we expand $E(Y|PS, T)$ by

$$E(Y|PS, T) = (1 - T)E(Y_0|PS, T) + TE(Y_1|PS, T)$$

$$= (1 - T)E(Y_0|PS) + TE(Y_1|PS) \quad by\ unconfoundedness$$

$$= (1 - T)\mu_0(PS) + T\mu_1(PS).$$

Under unconfoundedness, μ0 (.) and μ1 (.) are non-parametrically identifiable over PS because I assume a random sample on (Y, PS, T) is available. After demonstrating that the *average treatment effect* is identifiable in my context using propensity scores, I proceed in two steps. First, I cluster samples such that:

$$\begin{cases} 0 \le p_i - p_j \le 0.5 & PS\ for\ the\ 1st\ group \\ 0.5 \le p_i - p_j \le 1 & PS\ for\ the\ 2nd\ group \end{cases}$$

where the first group includes customers whose purchase probability is less than 0.5 and second group covers potential buyers with purchase probability of larger than 0.5 and pi (pj) represents propensity scores of customers in treatment (control) group. Second, I calculated average treatment effects by taking

$$\acute{\tau}_m = \frac{1}{S}\sum_{s=1}^{S}\{\frac{1}{n_s}\sum_{s=1}^{n_s}(Y_{is}(1) - Y_{is}(0))\}$$

Cluster Propensity Score Matching Results. Table 2.7 shows some interesting results. After matching on PB propensity score proximity, I see outcomes are quite the same in control and treated group for customers who are not likely to purchase promotions (I can assume these people are less price-sensitive). However, churn rate and total sum are significantly different between control and treatments in segment 2, i.e., customers who are estimated to be likely to buy promotion bundles show different behaviors depending on which group (control or treatment) they belong.

Table 2.7: Cluster propensity Score Matching

	Group 1: purchase prob < 0.5	Group 2: purchase prob > 0.5
Churn		
Control	0.062	0.061
Treatment	0.07	0.104
ATE	0.008	0.043
Overage Sum		
Control	957,434	864,005
Treatment	857,141	906,987
ATE	-100,293	42,982
Total Sum		
Control	2,052,514	1,711,800
Treatment	1,953,074	1,001,135
ATE	-99,440	-710,664

For instance, churn rate is 70 % more if a customer is exposed to the promotion and is more likely to purchase the bundle. Moreover, the total sum paid by price-sensitive customers (purchase prob > 0.5) is 71% more in control group compared to treatment while this difference is only 5% in group1 that includes people who are not likely to purchase the promotion bundle. If I assume, I have these two segments of customers (one segment is more price-sensitive and one segment less price-sensitive), I can see the impact of exposure to promotions on these segments. I speculate offering promotions lowers existing customers' inertia and increases their churn rate especially in the segment that proves to be more careful in choosing cost-minimizing plans. This effect can be very important in markets where people are already very inertial, but firm's constant and random offering of promotions can make them aware of the possibility of searching for better plans with other firms.

The ATE row in Table 2.7 represents the difference in mean outcomes for samples. Except for the sign of estimate for the overage sum in group 2, the direction of effect for other parameters is consistent with results of methods in previous sections. Treated customers of both low and high price-sensitivity spend less overall; however, price-sensitive customers who are exposed to promotions spend less on overage.

2.4 The Impact of Contract Structure on Revenue

In this section, I am going to build on the argument presented for the effect of pricing and contract features on customers' post-purchase behavior. As before, I am going to focus on revenue as my potential outcome. I look at contract duration, complexity and time pressure as explanatory variables that change post-purchase consumption level. The consideration of service usage behavior allows researchers to differentiate customers with decreased or low usage versus the ones with increased or higher usage. The former represents inertial customers, and the latter typifies the learning segment that has realized the value of the service through usage. Distinguishing these segments helps firms design optimal retention policies aimed at each group (Schweidel et al. 2011). Moreover, I am interested in studying the effect of diffusion of a new innovative service (Time-Constrained Credit or TCC) on customers' trend of usage. The exploratory analysis shows that there is a significant difference in outcomes for customers who have purchased TCCC (or received it as a gift) and those who have not.

Table 2.8: Comparison of TCC Users and Non-users

Variable	Non-users	Users
Overage Sum	385,645	1,246,855
Churn	0.0928	0.0561
Total Sum	1,206,947	2,212,331
N. Observations	111,290	70,940

* All differences are significant at the p-value of 0.0001.

Table 2.8 shows the results of t-statistic that is done to check the difference between TCC buyers and non-buyers. Specifically, after comparing people who are heavier users of this new service with the rest of the customers, I see heavy TCC buyers significantly spend more on plan sum, overage sum and churn much less. This is not expected because TCC is cheaper than the default allowance or overage. The result of such comparison is present in Table 2.9.

Table 2.9: Comparison of Heavy TCC Users and Non-users

	Non-users	Users
Overage Sum	615,327	1,916,603
Churn	0.082	0.038
Total Sum	1,474,886	2,996,384
N. Observations	167,445	14,785

* All differences are significant at the p-value of 0.0001.

Thus, I hypothesize that the diffusion of TCC may have increased people's consumption and their spending on extra credit. In the next session, I am going to study if the adoption of this service at period t helps increase credit consumption at later periods and investigate if the results are robust to different contract features (length, complexity, time-pressure of TCC units).

2.4.1 Estimating Treatment Effect in Panel Data with Multiple Arbitrary Treatments

If I argue the adoption of TTC changes consumption pattern and causes positive trends of usage (or simply heavier aggregate usage) among adopters versus non-adopters, I need to rule out selection. In other words, if people with heavier credit usage self-select themselves into treatment, then I cannot argue that the introduction and diffusion of TTC have changed consumption pattern and fueled more usage. Moreover, it would be interesting to take customers' heterogeneity into account when analyzing their response to treatment to distinguish segments that show a stronger effect in terms of steeper increase in demand. I model revenue generated by the purchase of TTC for each (i,t) as

$$Y_{it} = c_i + \beta X_{it} + u_{it} \quad t = 1 \ldots T, \qquad (11)$$

where c_i denotes unobserved individual effects, X_{it} is the set of explanatory variables, u_{it} is idiosyncratic error term. I add unobserved effect ci to allow for some elements of X_{it} to be correlated with unobservable heterogeneity. However, according to Woodridge (2010), most often the introduction of fixed effects makes the estimation of average population effect (β) computationally difficult and estimation of ci along with β leads to incidental parameters problem. There are other methods that help us control for unobserved heterogeneity in panel data. For instance, in Chamberlain's (1980) approach, c_i and X_i are correlated by "assuming a conditional normal distribution with linear expectation and constant variance." Mundlak (1980) assumes a different structure for the correlation between unobservable effects and explanatory variables as

$$c_i|x_i \sim Normal(\psi + \bar{X}_{i\eta}, \sigma_a^2) \quad t = 1 \dots T, \qquad (12)$$

where $c_i = \psi + \bar{X}_{i\eta} + a_i$ and X_i is the average of X_{it} , t=1...T. In Chamberlain's device, I use the vector of all explanatory variables across all time periods instead of the average. In this section, I proceed with the Mundlak approach for simplicity and refer to it as the random effects probit model. In equation (12), I assume a distribution for unobserved effect given X_i and by this assumption, I allow dependence between the unobserved effects and explanatory observables. Intuitively, using the average of all the explanatory variables as controls for unobserved heterogeneity means I am estimating the effect of X_{it} holding the time average constant.

A kind of sample selection is present in my model because people decide not to purchase TTC at some period but not at the other. This case is referred to as incidental truncation in the literature. In my study, I also deal with a rotating panel because at each period some of the original customers may drop, and others join. Let's assume for each individual i from the sample, $S_i \equiv (S_{i1}, \dots, S_{it})'$ represents the T × 1 vector of selection indicators where

$$S_{it} = \begin{cases} 1, & if\ they\ buy\ TTC\ at\ time\ t \\ 0, & otherwise \end{cases}$$

Suppose that, for each t, S_{it} is determined by the probit equation

$$S_{it} = 1[\delta_t X_i + \nu_{it} > 0],\ \nu_{it}|X_i \sim Normal(0,1) \quad (13)$$

where X_i contains all the variables that affect customers' purchase of TTC. This includes the average of total, plan and overage frequencies for all time periods (as suggested by Mundlack); first and second lags of treatment and coefficient of variation (CV) calculated up to each time period. I also included a trend coefficient for TTC purchase frequency from t = 1 to t calculated by OLS (regressing frequency for TTC purchase on month) and trend coefficient for overage frequency. For t = 1 and t = 2, one or both lags do not exist, so I avoid including missing lags. If ν_{it} represents total sum paid for TTC, it is observed only if the binary selection indicator Sit equals unity. I assume

$$E(c_i|X_i,\, \nu_{it}) = E(c_i|\, \nu_{it}) = \rho\nu_{it}\; t = 1 \ldots T, \quad (14)$$

hence, the error structure of equations (11) and (13) is given by

$$\begin{pmatrix} u_{it} \\ \nu_{it} \end{pmatrix} \sim Normal \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & \rho\sigma_u \\ \rho\sigma_u & 1 \end{pmatrix} \right]$$

If assumption $E(\, u_{it}|X_i, S, c_i) = 0$ holds, the inverse Mills ratio (IMR) obtained from running equation (13) should not appear significant in model (15). To estimate the whole panel selection model, I proceed in two steps. First, I run a probit model to estimate δ and store IMRs. Then I use the IMRs, $\hat{\lambda}_{it}$ for all time periods, explanatory variables X_{it} and run the pooled OLS regression using the selected sample:

$$y_{it} = c_i + X_{it}\, \beta + \sum_{t=1}^{T} D_t \hat{\lambda}_{it} + u_{it} \qquad t = 1 \ldots T, \qquad (15)$$

or

$$y_{it} = \psi + \bar{X}_{i\eta}\, \beta + \sum_{t=1}^{T} D_t \hat{\lambda}_{it} + a_i + u_{it} \qquad t = 1 \ldots T, \qquad (16)$$

where X_{it} is a vector of contract features that aims to control for customer observed heterogeneity in different plans and includes promotion dummy, *default TTC in the plan* (allowance*), first and second lagged variables of promotional TTC*, *complexity of contract* (dummy for bundles), *contract duration*, *service speed* and *time pressure of TTC* in the plan. Time pressure is measured by the ratio of the number of units purchased over expiration day. The sensitivity of the outcome with respect to the explanatory variables is shown by a vector of coefficients β. I also add tenure and interaction terms in X_{it} trying to see if new customers adopt the service first and whether the usage trend is steeper for old versus new customers. I include time dummies in Dt and $\hat{\lambda}_{it}$ is the IMRs calculated in step 1. Finally, as explained in the previous section, to avoid incidental parameters problem, I include the mean of explanatory variables to control for unobserved heterogeneity.

If Gourvuille and Somen's (2002) argument holds, I expect complexity to reduce or delay the post-purchase usage of credits. Moreover, I expect the length of the contract to be negatively associated with the purchase of overage. The last two hypotheses are logical because the length and complexity of contract mask the effect of the price paid for each element of contract (subscription, default allowance, modem or other compliments, etc.). Since the parameter of interest is the effect of the introduction of new service on the change of the trend of usage, I do not model the choice of contract and TTC purchase simultaneously. Thus, selection is only considered in the latter case, and I use contract features as controls. That is, I am going to study if I hold these contract features fixed, how new service introduction affects customers' post-purchase behavior.

I used both the random effect probit model as equation (16) and the fixed effect model in equation (15) to estimate the parameters of interest and present a comparison of results. In the random effect probit model, I am going to assume individual specific unobserved heterogeneity can be partially represented by the average of all explanatory variables. In the fixed effect model, I am going to estimate fixed effects directly.

Table 2.10: The Effect of New Product and Contract Features on Overage Revenue

Variable	Variable
Treat	126705***(788)
Complexity * tenure	-1141***(105)
Duration * Allowance	7.01(5.62)
Duration * TCC Allowance	-4.52(2.82)
Treat * Complexity	-1693(3196)
Treat * Tenure	-1693***(41)
Treat * Duration	836***(257)
Treat * Complexity * Tenure	-248(155)
Mean overage sum	0.965***(0.008)
Mean TCC freq	-7861***(242)
Tenure	261***(26.8)
L. TCC Sum Trend	0.121***(0.009)
L.TCC Sum	-0.044***(0.007)
L.Overage Sum	0.039***(0.005)
Complexity	13080***(1868)
Duration	806***(167)
Time-Pressure	19796***(419)
Service Speed	0.57***(0.062)
Service Gig	-234***(71.4)
Service TCC Gig	55 (50.5)
Number of Obs = 560,058 Prob >F = 0.0000 R-square = 0.5893	

It could be that on long or complex plans, people are confused and overspend. By allowing for learning about preferences, I can partially control for this confusion effect. I expect longer-tenured customers to be more aware of their needs. Consequently, more tenured customers on complex plans are expected to spend less. If unobserved preferences derive selection, having IMRs and FE in the equation will produce inconsistent estimates (for proof, see Woodrige 2010). In addition to this consideration, overall R-square for the fixed effect model was low, and I only present the results for equation (15) in the appendix. Table 2.10 presents the results of the estimation of equation (16). The coefficient for selection variables in most periods (except first and last one) was significant, which means selection exists.

One of the interesting results is the significant positive coefficient of “L.TCC Sum Trend” - lagged coefficient of Trend calculated for the sum paid to purchase TCC at each time. For instance, TC Sum variable represents the slope of trend line fitted on the total sum paid for TCC by individual i at time t. According to estimates in Table 2.10, the higher the slope, the more the propensity to spend more on extra credit within the next period (coefficient is positive 0.121). This indicates after controlling for selection of this new innovative service (TCC); I observe those who purchase this service at period t are likely to spend more at following periods. This effect is beyond the level that could be explained by selection because I control for selection at each period using IMRs as previously explained. I conclude that the adoption of this service has a positive effect on increasing people’s consumption of extra credits, and overall it increases overage purchase.

Next, I examine possible customer heterogeneity in absorbing this effect. People on complex or long plans tend to purchase more extra credit. Time-pressure also has a considerable effect on the increased overage purchase. If more units should be consumed in a shorter time, I observe higher revenue generated from overage sales. The interaction terms show that higher tenure reduces the positive effect of new service on increased consumption (coefficients for interactions are negative). To wit, new customers are more likely to be affected by the impact of new service on consumption behavior.

2.5 Conclusion

Quantitative models that explain price promotions, tariff structure and their impact on a firm’s marketing success indicators (revenue, profit, customer retention, etc.) have ignored the role of behavioral mechanisms leading to outcomes. Instead, they have mainly focused on justifying profitability of such policies by offering economical justifications. However, behavioral work on pricing suggests that other factors beyond apparent economic variables affect customers' reaction to pricing and tariff design (Bertini and Wathieu 2008). In this research, I bridge this gap and bring together the quantitative and behavioral work on pricing in an experimental framework to examine which of screening or sunk-cost theories derive consumers' reactions to pricing and tariff design decisions.

In this study, by using a natural field experiment, I quantify the impact of price promotions, tariff (contract) features and new product introduction on consumer's post-purchase behavior (revenue and churn). I separate screening and sunk-cost effects of promotions and find that price promotion reduce consumption beyond the level that can be explained by screening effects. Results show that these promotions attract customers who are already aware of their consumption preferences. However, controlling for selection does not rule out all the differences between control and treated group. I show how contract features (favoring sunk-cost assumption) and the mere act of offering promotion bundles change customers' post-purchase behavior. That is, I find contract design that masks the pricing or alleviates the pain of payment reduces consumption. Interestingly, I find that low price-sensitive customers are not much affected by constant exposure to price promotions, but high price-sensitive segments are adversely impacted in terms of the total sum and churn rate. Findings can have managerial implications; firms should not offer frequent random promotions since it can have a detrimental impact on customers by encouraging them to search for better deals.

There are several directions in which the current research can be extended. First, a structural model that accounts for firms' interactions with customers can provide insight into customers' future behavior and can help us design optimal pricing policies that are in equilibrium, considering the trade-off mentioned. Incorporating time-varying marketing efforts can aid us in including financial considerations when deciding on marketing policies. Second, one can study if the effect of exposure to promotion on decreased inertia is the same for heterogeneous products and services in terms of hidden information and search costs. The quality of some offerings can be evaluated easier than services mainly because it does not need post-purchase experience with them (quality can be somewhat determined before consumption through inspection). Moreover, it would be interesting to see if the result holds for other offerings which different level of switching costs. Finally, the inclusion of competitive information can assist us in finding the equilibrium for both firms when they jointly decide how much to promote the product.

CHAPTER 3

THE ROLE OF MOTIVATIONS AND CONTENT EMOTIONAL TONE IN MESSAGE VIRALITY: EVIDENCE FROM TWITTER

3.1 Introduction

In recent years, social media have emerged as one of the main channels for disseminating information products among ordinary people, challenging the manipulation of the mass media by corporations. Sharing online content is very common in modern life; people send news articles to friends, share YouTube videos with their families, and forward reviews to their friends. Studies report that 59% of people frequently share online content with others (Allsop et al. 2007), and someone tweets a link to a New York Times story once every four seconds (Harris 2010). From a marketing perspective, this pattern has shifted digital marketing strategy away from an emphasis on "paid" media, where a brand pays to advertise, to "earned" media, where the customers themselves become the channel of delivery (Corcoran 2009). As advertising platforms, micro-blogging websites have two main advantages for advertisers. First, they allow advertisers to study drivers of content virality among platforms' users. Second, they allow advertisers to identify potential users who respond more favorably to specific content and target advertising messages to them (Vaynerchuck 2013).

Word of mouth (WOM) and social media are viewed as a more cost-effective tool than traditional media, but their utility depends on people distributing content that helps firms. If targeted individuals do not share a company's content, the benefit of the media is lost. Consequently, understanding what encourages users to share content and which segment of users respond to specific content with designed features can help organizations create content that is contagious and received well among the targeted segment. Moreover, platforms themselves need to understand what motivates users to be active since their profitability hinges on how active users are on the platform (Toubia and Stephen 2013)

However, extant marketing research on social media and online word of mouth has primarily focused on the outcomes of user activity and less on drivers of user activity (Katona et al. 2011). Similarly, less is known on what content characteristics help it go viral. Specifically, academic research on Twitter and similar micro-blogging platforms has been limited (Lambrecht et al. 2014). Exceptions include Tucker et al. (2014) who examine the effectiveness of targeting early trend propagators vs. late ones on Twitter and Stephen et al. (2012) who study how user activity predicts the virality of shared URLs in tweets. Measuring influence is more straightforward in Twitter since unlike most other contexts, the network over which WOM influence spreads is generally unobservable. Hence, influence is less difficult to attribute accurately, especially in instances where diffusion propagates for multiple steps (Bakshy et al. 2011.)

One reason people may share stories, news and information is because they contain useful information. Consumers may share such practically useful content for altruistic reasons (e.g., to help others) or for self-enhancement motivations (e.g., to look knowledgeable or clever, Wojnicki and Godes 2008). Users may also share useful content to benefit from social exchange value (Homans 1958) or to generate reciprocity (Fehr et al. 1998). When content contribution is not monetized, the literature categorizes the mentioned incentives into two types of utility that may motivate noncommercial social media users to contribute content: intrinsic utility and image-related utility. Users receive direct intrinsic utility from contributing content when they do it for the inherent satisfaction of the activity itself (Ryan and Deci 2000.) The image-related utility, on the other hand, assumes that users are motivated by the perceptions of others (Fehr and Falk 2002.) The related concepts to image-related motivations are status or prestige seeking needs. The image-related utility is also related to status-seeking or prestige motivation (e.g., Glazer and Konrad 1996; Harbaugh 1998a, b). These researches are mainly in the context of charity and donations. In this research, I build on the previous works to explain how intrinsic-and image-related motivations make people contribute free content on a more recent phenomenon, microblogging platforms.

To account for the heterogeneity of users on this platform, I argue that users with different levels of social capital (SC), measured by number of followers, have different motivations that underlie their posting behavior. Social capital is the product of each user's investments in her social network, which is made by posting activity. In a broader sense, social capital is the resources and values available to a user by being part of a network; and is obtained after some forms of investment is made. On Twitter, users need to invest time and contribute content to gain followers, while following people requires insignificant investment. Hence, followers are a user’s social capital. Moreover, I choose number of followers as the factor that categorizes users’ motives and intentions to post content since this information is readily available to marketers and they can segment the targeted population of their campaigns and craft clear marketing strategy for each segment.

Content characteristics, such as emotional aspects, may also affect whether it is shared (Heath, Bell, and Sternberg 2001). For example, there is a common belief that people are more likely to pass along negative news (Godes et al. 2005). Hence, determining which segment will propagate each type of content and what underlying motivations account for that, have an important economic implication. To this end, this research explores users' motivations to post tweets that have positive or negative emotions embedded in them. I do so through a quantitative analysis of questionnaires answered by Twitter users. I analyze users' motivations and discuss how their social capital can play important roles in what and why users will retweet. The closest work to this research is the paper by Toubia and Stephen (2013). The authors conduct a filed study and use a dynamic discrete choice model to find that both intrinsic and image utilities are key drivers of posting activity for users and that utilities are concave in the number of followers. They find that though both types of utilities drive users, image-related utilizes are stronger for most people. The contribution of the research is two-fold: first, instead of implicit modeling of users' motivations by imposing assumptions and inferring them from posting activity, I directly ask each segment of users about their motivations, and I delve deeper into each group of latent incentives by breaking them down to their components. Second, I investigate how emotions embedded in each content activate different underlying motives to post for each segment and what posters expect to get in terms of involvement of other users who see their tweets.

Next, I review the literature on online sharing behavior and explain the research gap observed in previous works. Then I present my data collection procedure and summary statistics. Finally, I discuss the model results.

3.2 Online Sharing Behavior and Network Size

This research builds on three streams of literature. The first stream is a small but growing marketing literature that studies consumer sharing behavior on online platforms such as Twitter. For example, Bakshy et al. (2011) used 1.6M Twitter users and 74M diffusion events to examine the attributes of influencers who are more successful in spreading a message, and find that the size of influencers' network predicts information cascade and that the most cost-effective segment to identify are ordinary influencers with average or even less than average effect. In this line of research, some studies have examined the relationship between the size of a user's network and the drivers of sharing behavior and report that utility from status can be modeled as a non-decreasing concave function of the number of followers. Baumeister and Leary (1995) review empirical evidence to demonstrate that humans express the need for forming a minimum number of social bonds, but the formation of further social attachments beyond that minimal level is subject to diminishing returns. Thus, I expect that people will experience less satisfaction in the creation of new relationships beyond a point. Likewise, DeWall et al. (2008) experiment to show that the more the need for social acceptance is satiated, the less is the motivation to satisfy that need.

3.3 Online Virality and Content Characteristics.

The second stream of research I build on is the works on content characteristics and strategies that encourage virality. In a randomized field experimental setting involving the 1.4 million friends of 9,687 experimental users on Facebook.com., Aral and Walker (2011) show that allowing users to spread a message at their discretion is more effective in increasing the adoption rate per message, but forcing consumers to spread a customized message is more effective in increasing contagion. Chen et al. (2011) use an experiment to disentangle observational learning from positive/negative WOM and report that social influence is more pronounced at the beginning of a product life-cycle. Berger and Milkman (2012) use a data set of online *New York Times* published over three years and find that positive content is more viral than negative content. They also examine how emotions and psychological arousal impact virality and report that online news is more likely to be shared if it evokes high or negative arousal as opposed to low-arousal or deactivating emotions such as sadness. Eckler and Bolls (2011) also emphasize that a more positive tone encourages more virality. Porter and Golan (2006) suggest outrageous content such as sexuality, humor, violence, and nudity encourage virality. Brown et al. (2010) emphasize the importance of comedic violence and outrageous appeals as key drivers of these ads. Consumers may share content to present themselves (Wojnicki and Godes 2008) and to express their identity, and for these reasons, positive content is more likely to be shared due to its positive effect on self-presentation. This happens because people prefer to make others feel good and also because sharing positive information could increase potential rewards from recipients (e.g., recommending a restaurant).

3.4 Motivations and Contribution to Online Platform.

The last body of research related to my work is the domain that investigates users' motivations to contribute to open platforms. Social studies have extensively studied intrinsic and image-related utilities (e.g., Glazer and Konrad 1996; Bénabou and Tirole 2006, Ariely et al. 2009). Lerner and Tirole (2005) argue that open-source developers derive intrinsic utility from working on "cool" projects and also unpaid open source contributions may lead to peer recognition and job offers in the long-run. Several other works in the context of open-source development provide survey-based evidence that both intrinsic and image related utilities play a role in developers' contribution decisions (e.g., Hars and Ou 2002, Roberts et al. 2006). Bughin (2007) surveys users of online video-sharing sites and reports that motivations to contribute are both image-related ("I seek fame") and intrinsic ("it is fun"). Hennig-Thurau et al. (2004) survey 2000 online opinion platforms to investigate what motivates people to articulate themselves on the internet. The authors found that motivations tend to be either intrinsic (e.g., "It is fun") or image-related (e.g., "My contribution make me look clever”). To my knowledge, no previous study has examined what motivates different segments of users with various potential influence (i.e., SC measured by number of followers) to post positive/negative content on micro-blogging platforms like Twitter and what are their perceived consequences of sharing positive/negative content with other community members.

3.5 Hypothesis Development

Boyd et al. (2012) provide survey evidence on what motivates Twitter users to post content. Respondents mention ten main motivations; I classify them into two groups of “intrinsic” and “image” -related motivations. I argue that a user should derive more intrinsic utility from spreading content as SC increases. Likewise, there is a monotonic non-decreasing relationship between image drivers and the size of SC. Followers need to be earned; hence, SC is an informative signal of social status on online platforms. Thus, number of followers is the most straightforward *measure* of social capital.

The most common form of social capital is information Access. By being in a social network, users can gain individual benefits such as receiving useful pieces of information. Network also affects the timing of information access, and the earlier and quicker access is more desirable. Furthermore, social networks are sources of referral, which add value and credibility to the information and the particular users who disseminate it (Burt 1992, Recuero et al. 2011.) Accordingly, I argue that the “user experience” on the platform and “user authoritativeness” are two latent constructs that change the stock variable of interest “social capital.” User experience, measured by years active on the platform and the intensity of platform activity, affect the ability of a user to filter and receive more valuable pieces of information. Similarly, users with high authoritativeness- i.e., elite, convincing and credible members, will likely have higher social capital.

Prior research shows that emotional tone of the message impacts virality and follower involvement with the message. However, the evidence is mixed as what emotional tone leads to higher levels of virality. For example, there is a lay belief that people are more likely to pass along negative news (Godes et al. 2005), whereas Berger and Milkman (2012) use a data set of online *New York Times* and find that positive content is more viral than negative content. In this research, I am testing to see which pattern is more pronounced on microblogging platforms such as Twitter. I hypothesize that since compared to neutral contents both positive and negative emotions have been reported to affect users, the presence of these emotional tones is more likely to increase the engagement of followers, but beyond a point, posters do not care as much about followers' reaction to their post as before. As such, they are less likely to post contents to engage others emotionally. Therefore, even though higher levels of expressed emotions are more likely to be conducive to tweet virality (liked, retweeted, and replied), high SC users will be less motivated to express strong emotions because they will care less about others' reaction to their posts. Compared to low SC users, high SC users will be more likely to increase others engagement if they post negative content because, on Twitter, one main reason people may share content is that they contain useful information. Taking together, my arguments lead to the followings:

> H1: Due to the station of image-related needs with higher numbers of followers, higher SC users will less likely post due to image-motivated reasons and more likely post due to intrinsic-motivated reasons.

H2: The level of expressed emotions in a tweet is associated with the number of followers in a nonlinear (diminishing returns) manner, such that the effect of additional followers on expressed emotions is positive at low levels of followers but becomes smaller as followers increase.

H3: Compared to positive emotions, negative emotions in tweets are more likely to increase user engagement with the post.

3.6 Data Description

The data comes from survey responses answered by hired tweeter users on Amazon Mechanical Turk. I hired 181 unique workers in three different consequential surveys. To screen the workers who are indeed Twitter owners, I used a screening criterion which is available for an additional charge of five cents per response. In the last round, I paid $0.7 per subject for a three-minute survey. Of these 181 responses, 118 responses were coded "confident." I check the responses that took less time than the average time of researchers as unconfident since the assumption is that the researcher who is already familiar with the objectives and constructs used in the survey records the minimum threshold time of response. In round 1, the survey included two main blocks of "activity pattern" and "tweet emotions," and 18 responses were collected. In the second round, three main blocks of questions were used: "activity pattern," "tweet emotions," and "motives;" and 18 responses were collected. In the last round, 82 respondents answered questions on "activity pattern," "tweet emotions," "motives" and "authoritativeness."

I am investigating the response to my research questions from users with different levels of social capital, namely the number of followers, thus I need to segment users based on this factor.

Hence, I suggest six classes of users with different range of followers. The six classes and the size of each of them are shown in Table 3.1.

Table 3.1: Follower Distribution Based on Final Sample

Number of followers	Frequency	Percent
G1-fewer than 10	10	8%
G2-more than 10, fewer than 50	33	28%
G3-more than 50, fewer than 100	29	25%
G4-more than 100, fewer than 200	24	20%
G5-more than 200	22	19%
Total # Observations	84	100%

Since the last class with the "number of followers more than 1000" contains only a single user; I add class 5 and 6 to have 15 responses. I see that according to this segmentation; number of followers has a somewhat uniform distribution.

3.6.1 Model-Free Evidence

The correlation between tweet emotions and motivations are shown in Table 3.2. The correlations between tweet emotions and motivations show that expressing emotions and the belief about consequences of such expressive behaviors have a higher correlation with intrinsic motives than image motives; the exception is the belief that "negative emotions in tweets can increase users' involvement," which is more correlated with image than intrinsic incentives. Moreover, the higher the number of followers, the more likely it is that users believe negative emotions can increase "involvement" and "number of new followers" compared to positive emotions. However, users in general, and more specifically the top segments (i.e., segments with higher social capital) are more likely to post positive tweets than negative tweets. Finally, both motivations have significant correlations with number of followers, but as number of followers increases, users are more likely to post due to intrinsic reasons than image-related motives.

Table 3.2: Descriptive Statistics and Correlation Matrix

Variable	Image	Intrinsic	N. Followers
N. Followers	0.221	0.275	1
Express Pos Emotions	0.480	0.591	0.452
Express Neg Emotions	0.357	0.475	0.311
Positive Increases “Followers”	0.334	0.378	0.054
Negative Increases “followers”	0.283	0.354	0.293
Positive increases “involvement”	0.351	0.445	0.120^{ns}
Negative increases “involvement”	0.183	0.176^{ns}	0.195

Note: "ns" identifies insignificant correlations.

In Table 3.2, number of followers is a categorical variable that takes the value of 1 to 5, where 1 represents the first segment with the lowest number of followers (i.e., fewer than 10) and 5 represents the segment with the highest number of followers (i.e., more than 200). In Figure 3.1, the average of latent constructs is calculated for each motive construct. The order of groups from left to right is g1, g2, g3, g4, and g5. The data uses 100 observations in this graph. There are 9, 26, 27, 20, 18 observations in the groups, respectively. Figure 3.3 shows the breakdown of the construct for intrinsic motives into its components for each group of users. The distribution of responses to motivation questions shows that motivations such as “it is fun to tweet,” “validate others’ thoughts,” “spread news” and “identify with a group” are most common. More than 60%

of the sampled users are very likely to tweet for the above reasons. The order of groups from left to right is g1, g2, g3, g4, and g5. Next, I present my formal model and results.

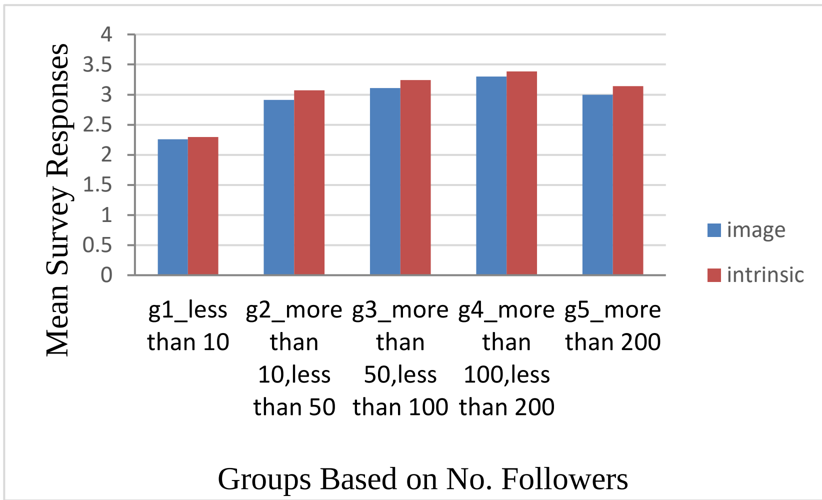


Figure 3.1: The Average of Motivations to Post on Twitter by Group

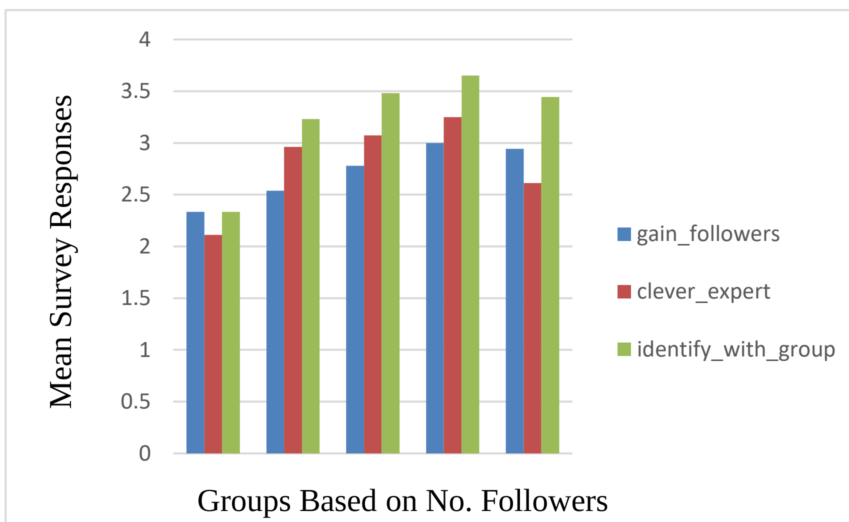


Figure 3.2: The Average of Image Motives by Group

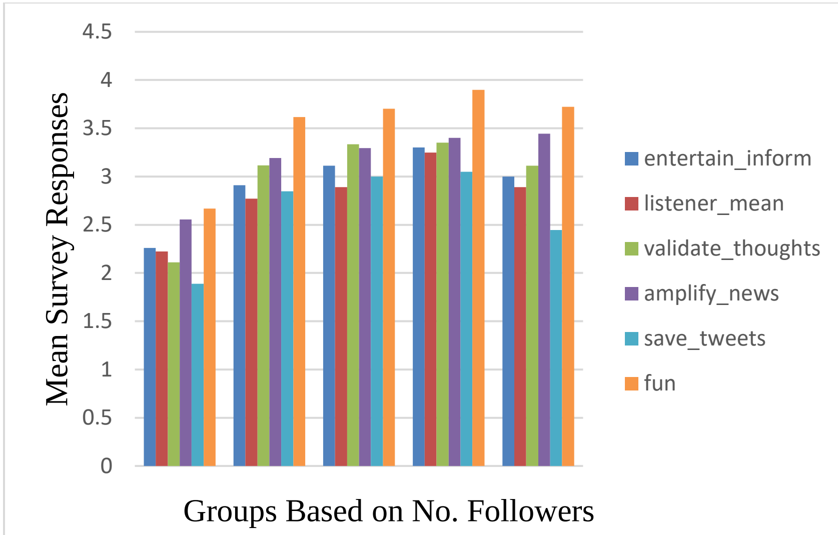


Figure 3.3: The Average of Intrinsic Motives by Group

3.7 Data Analysis and Discussion

I recategorize the segments into two classes of high and low social capital. This classification done for manipulation check purposes and the main mediation analysis is not dependent on it. In other words, I aim to check to see if there is a difference in reactance patterns between high and low SC groups. Low social capital (SC) includes the first two segments, with number of followers fewer than 50, and high SC includes groups 3-5 with number of followers more than 50. I choose 50 as the threshold for analysis because according to research (Bakshy et al. 2011), the median number of followers that users have on Twitter (eighty-five) falls in the third group of this research (Table 3.1). Statistics show that the average number of followers per user is 208. (Jeff Bullas 2015)

Positive and negative contents. Participants shared more positive tweets in the high SC (M=3.67, SD=0.63; F=2.43, p-value=0.006) than in low SC (M=2.93, SD=0.99). However, there is no significant difference in sharing negative tweets between high SC and low SC.

Image motives. There was a main effect of SC ($F= 2.26$, $p = .01$), such that participants in the high SC condition reported a stronger image motive to tweet ($M = 3.12$, $SD = 0.73$) than those in the low SC condition ($M = 2.7$, $SD = 1.1$).

Intrinsic motives. There was a main effect of SC ($F= 2.17$, $p = .02$), such that participants in the high SC condition reported a stronger intrinsic motive to tweet ($M = 3.32$, $SD = 0.63$) than those in the low SC condition ($M = 2.9$, $SD = .92$).

Mediation analysis. To test whether users' image and intrinsic motives mediate differences in positive versus negative emotions expressed in tweets for different segments of SC, I conducted two analyses using the likelihood of expressing positive information and the likelihood of expressing negative emotions, respectively, as my dependent variable (Hayes 2013)

I first tested whether image and intrinsic motives mediated the effect of SC on posting positive content on Twitter. Two separate regressions revealed that SC predicted both image motives to tweet ($B=.15$, $t(82)=1.96$, $p=.05$) and intrinsic motives to post *positive* content ($B=.16$,$t(82)=2.54$, $p=.01$). Next, a regression including SC and two motives (adj-Rsquare=.42) revealed that intrinsic motives significantly predicted the likelihood of posting *positive* content ($B = .54$, $t(82) = 3.4$, $p = .002$) whereas image motives do not ($B = .03$, $t(82) = .23$, $p = .8$). In addition, SC significantly predicted the likelihood of sharing positive contents ($B = .21$, $t(82) = 3.59$, $p =.000$). These results indicate that intrinsic motives successfully mediate the impact of SC on the likelihood of sharing positive content whereas image motives do not.

Next, I tested whether the differences in the likelihood of sharing negative information as a function of SC could be attributed to any of the motives. A regression including SC and two motives (adj-Rsquare=.24) revealed that intrinsic motives significantly predicted the likelihood of posting negative content (B = .58, t(82) = 2.85, p = .005) whereas image motives do not (B = -0.06, t(82) = -0.34, p = .73). In addition, SC significantly predicted the likelihood of sharing negative contents (B = .16, t(82) = 2.06, p =.04).

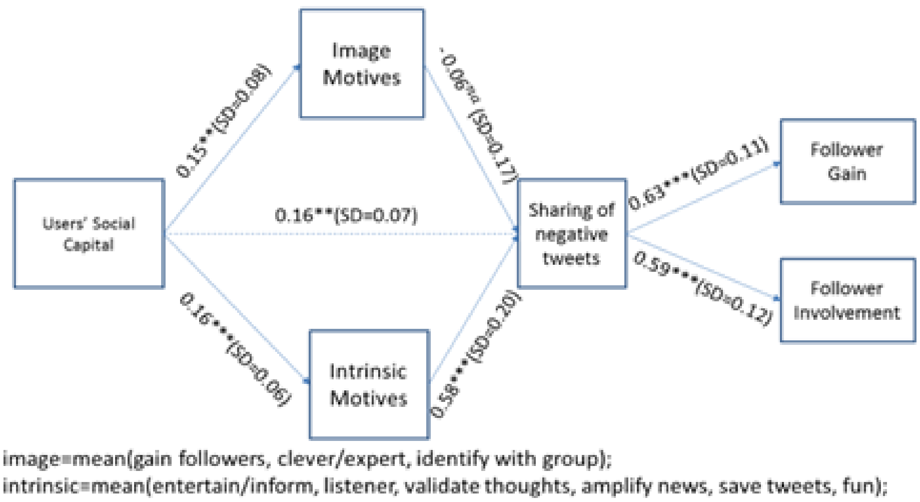


Figure 3.4: The Mediation of "Negative" Content

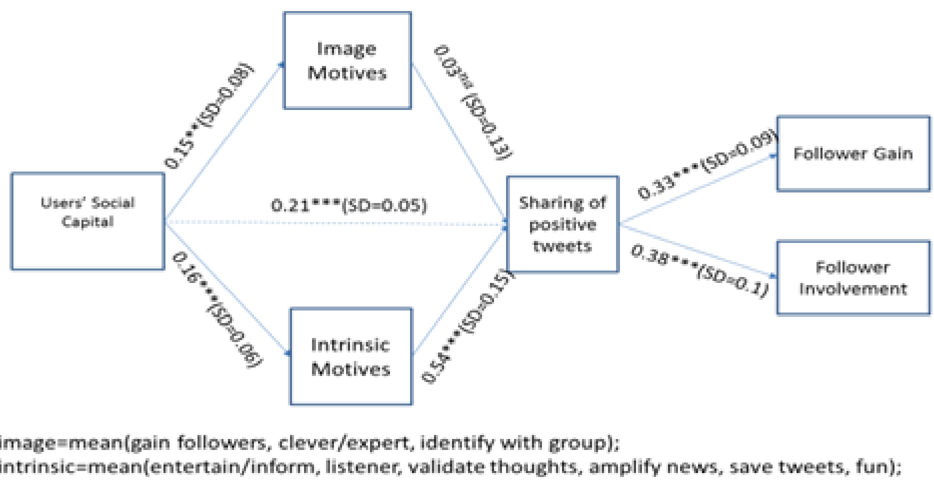


Figure 3.5: The Mediation of "Positive" Content

These results indicate that intrinsic motives successfully mediate the impact of SC on the likelihood of sharing negative content whereas image motives do not. Overall, the results of the mediation analyses show that higher SC triggers intrinsic motivations whereas low SC activates image motivations to post content on the platforms and that higher SC encourages posting more positive content than lower SC.

Positive/negative content and Users involvement. Two separate regressions reveal that sharing positive content is believed to encourage users' involvement more in terms of likes, replies and comments (B=.38) than in terms of attracting new followers (B=.33). On the contrary, sharing negative content is believed to encourage users' involvement more in terms of attracting new followers (B=.63) than in terms of users' involvement (B=.59).

3.7.1 Simultaneous Equation Modeling

To test the whole model simultaneously in a single analysis instead of testing separate regressions, I use Simultaneous Equation Modeling (SEM, also called covariance structure analysis). Measurement error is a potential concern in mediation testing because of "attenuation of relationships," and the SEM approach can address this problem by removing measurement error (Baron and Kenny, 1986). First, I tested the model shown in Figure 3.6 in SEM; I used the mediation analysis that is like the one I used in the previous section. Items used in scales for image and intrinsic motives are as explained before. Items used in "activity frequency" scale include login/read, tweet original, retweet, comment/ reply. The scale loadings and other statistics are presented in the appendix (Table 3.4).

Second, I used "Preacher and Hayes" method (P&H) and bootstrapped standard errors to get the significance of parameters (Preacher & Hayes 2004). The significance of indirect effect demonstrates whether mediation exists in the model. The R code for this analysis can be found in the appendix. With bootstrap confidence intervals, I empirically approximate the sampling distribution of indirect effects by repeatedly resampling the data with replacement and estimating coefficients in each resample. I then use the empirical distribution to generate a confidence interval for the indirect effect and other coefficients. Finally, I also tested the model using Preacher and Hayes mediation analysis but omitted bootstrapping in the model and instead used standard methods to calculate standard errors (SEM version 2). The results of this model are presented in SEM version 3.

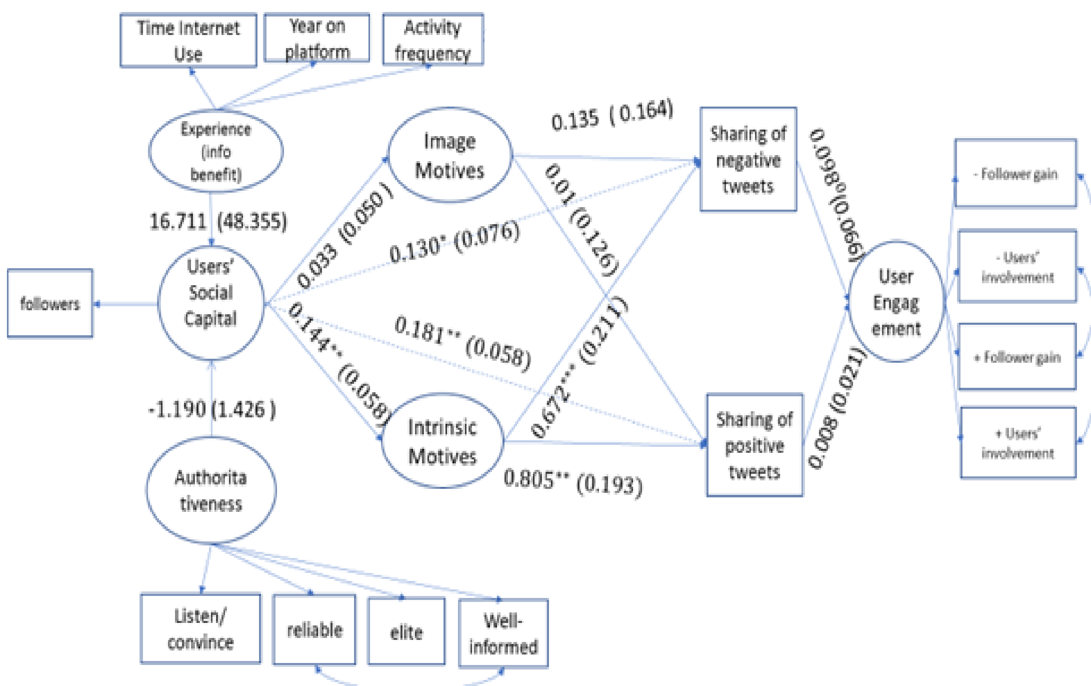


Figure 3.6: Full SEM Model

3.8 SEM Results

The first model shows that intrinsic motives (in order of importance determined by standardized loading coefficients on latent constructs: validate thoughts, fun, listener, amplify news, entertain/inform, save tweets) are more conducive to positive expressions (B=.805, SE=.193) as opposed to negative expressions (B=.672, SE=.211); whereas image motives (in order of importance: look clever/expert, identify with a group, gain followers) are more conducive to negative expressions (B=.135, SE=.164) but not positive expressions (B=.01, SE=.126, p=.94). Results also indicate that expressing negative emotions (i.e., critical news, serious posts) are believed to encourage higher levels of engagement (B = .098, SE=.066, p =.13). Note that the standardized loading coefficients on engagement construct show that "negative_followers" (StB=.866), "negative_involvement" (StB=.821), "positive_followers" (StB=.220), and "positive_involvement" (StB=.189) are accordingly ordered based on the magnitude of the effect. In this scale, "negative_followers," refers to "the belief that posting negative contents leads to new followers gained"; the rest of the items are interpreted similarly.

Higher number of followers is associated with intrinsic motives (B = .144, SE=0.058, p = .013) but has no effect on image motives (B = .033, SE=.050, p = .5). High SC also encourages more positive expressions (B=.181) than negative ones (B=.130). In the presence of SC and intrinsic motives, image motives are not predictive of negative emotional tone. This implies that image motives do not mediate the use of negative tone, but they mediate the relationship between number of followers and positive tone. On the other hand, intrinsic motives mediate the relationship between SC and emotional tone. The full P&H analysis in Table 3.3 shows that the total effect of intrinsic motives (i.e., the sum of direct and indirect effects) is stronger on positive expressions (B=0.297, SE=0.089, P=0.000) compared to negative expressions (B=.226, SE=.080, p=.011). The table is part of the output for SEM version 2. However, the full mediation analysis of P&H with bootstrap shows that total effect of image motives on negative expressions is insignificant (B=.134, p=.129), while their effect on positive expressions is significant (B=.181, p=.004)

Table 3.3: Full P&H Median Analysis

	Estimate	Std.E	z-value	P(>|z|)	Std.lv	Std.all
image_pos_ind	0.000	0.031	0.011	0.991	0.000	0.001
image_neg_ind	0.004	0.032	0.140	0.889	0.006	0.006
intrinsic_pos_ind	0.116	0.063	1.828	0.068	0.148	0.177
intrinsic_neg_ind	0.097	0.069	1.392	0.164	0.124	0.133
image_pos	0.181	0.063	2.879	0.004	0.232	0.276
image_neg	0.134	0.088	1.517	0.129	0.171	0.184
intrinsic_pos	0.297	0.080	3.693	0.000	0.380	0.452
intrinsic_neg	0.226	0.089	2.537	0.011	0.289	0.310

Note 1: Image_pos_ind stands for the indirect effect of image motives on the expression of positive emotional tone.

Note 2: Image_pos stands for the total effect of image motives on the expression of positive emotional tone (direct + indirect effects)

3.9 Conclusion

In this research, I demonstrated that there is an inverted U-shaped relationship between number of followers and motives. This finding is in line with the literature that suggests satiating the need for image and intrinsic utilities leads to a reduction in the drive to satisfy the need. (Baumeister and Leary 1995, Toubia and Stephen 2013). The breakdown of motives also revealed that both image and intrinsic motives are highest for users with a medium number of followers. Moreover, I showed that high SC users are more "intrinsic" driven as opposed to "image" driven, and are more likely to post positive posts, although they are more likely to believe that negative content has higher propensity to increase followers. Generally, users believe that negative content is more conducive to virality, both in terms of gaining user involvement (likes, comments, replies) and attracting new followers, but as SC increases, they are less likely to be concerned with users' involvement with their posted content. In sum, as the number of followers increases, users are more likely to post due to intrinsic reasons than image-related motives, this explains that even though they believe negative content can lead to more virality, they do not post negative content because with high SC they do not care as much about the reciprocity from followers as they care about the intrinsic motivations to post content. These findings have implications for marketing campaign designers. We live in an information overload century, and firms need to find ways to stand out and to get attention; otherwise, their tweets can easily be ignored. This research proposes that paying attention to the emotional tone of the contents and tapping on Twitter influencers with a medium number of followers are two separate strategies firms can work on to stand out like advertising billboards on the roads. Considering that the average Twitter user follows five business, % 80 of users have mentioned a brand in their tweets, and other Twitter marketing statistics[6], I believe that there are huge opportunities available to firms, but they need to design

[6] https://www.brandwatch.com/

effective contents and show them to appropriate targeted segments. This research took one step in this direction.

APPENDIX

QUESTIONNAIRE FOR CHAPTER 3

Q1. Do you have a Twitter account?

[ ] Yes [ ] No

Q2. How often do you post tweets for the following reasons?

[ ] Never [ ] Rarely [ ] Sometimes [ ] Very often [ ] Always

To gain followers or reciprocity (e.g., retweets, likes, replies) from more visible participants
To entertain or inform a specific audience
To make my presence as a listener visible
To validate others' thoughts
To look clever, expert and knowledgeable
To identify with a group, you belong to or associate with
To amplify or spread tweets to new audiences (e.g., world news)
To save tweets for future personal access
To have fun by communicating this way with other people in the community

Q3. How often do you do the following activities in your Twitter account?

[ ] Never [ ] Less than once a month [ ] 2-3 times a month [ ] Weekly [ ] Daily

Login/Read posts
Tweet original posts
Retweet
Comment/reply

Q4. Choose the number of your Twitter followers:

[ ] fewer than 10
[ ] more than 10, less than 50
[ ] more than 50, less than 100
[ ] more than 100, less than 200
[ ] more than 200, less than 1000
[ ] more than 1000

Q5. When did you join Twitter?

[ ] 2006 [ ] 2007 [ ] 2008 [ ] 2009 [ ] 2010 [ ] 2011

[ ] 2012 [ ] 2013 [ ] 2014 [ ] 2015 [ ] 2016 [ ] 2017

Q6. How many hours a day do you spend on social media (Facebook, Twitter, Instagram, Soundcloud, Spotify, etc.)?

[ ] Less than 30 minutes
[ ] 30m-1h
[ ] 1h-2h
[ ] 2h-3h
[ ] more than 3h

Q7. How often do you post a tweet to express ***positive*** emotions about things (e.g., good/interesting world/local/personal news)?

[ ] Never [ ] Rarely [ ] Sometimes [ ] Very often [ ] Always

Q8. How often do you post a tweet to express negative emotions about things (e.g., satirical news/opinion, critical posts, complaints, bad world/personal news, etc.)?

[ ] Never [ ] Rarely [ ] Sometimes [ ] Very often [ ] Always

Q9. In your opinion, how likely are you to *increase* the number of your followers by expressing ***positive emotion*** *in* tweets? (e.g., good world/local/personal news)

[ ] Extremely unlikely [ ] Unlikely [ ] Neutral [ ] Likely [ ] Extremely likely

Q10. In your opinion, how likely are you to *increase* the number of your followers by expressing ***negative*** emotions in tweets? (eg. satirical news/opinion, critical posts, complaints, bad world/personal news, etc.)

[ ] Extremely unlikely [ ] Unlikely [ ] Neutral [ ] Likely [ ] Extremely likely

Q11. In your opinion, how likely are you to increase the <u>involvement</u> of your followers (retweets, likes, replies, etc.) with posting ***positive*** tweets? (e.g., good world/local/personal news)

[ ] Extremely unlikely [ ] Unlikely [ ] Neutral [ ] Likely [ ] Extremely likely

Q12. In your opinion, how likely are you to increase the involvement of your followers (retweets, likes, replies, etc.) with posting negative tweets? (eg. satirical news/opinion, critical posts, complaints, bad world/personal news, etc.)

[ ] Extremely unlikely [ ] Unlikely [ ] Neutral [ ] Likely [ ] Extremely likely

Q13. To what extent do your followers find you **<u>well-informed</u>** about the topics you post?

[ ] Not at all [ ] Slightly [ ] Somewhat [ ] Moderately [ ] Extremely

Q14. To what extent do your followers find you a **<u>reliable</u>** source of information for the topics you post? (i.e., they do **not** argue **against** what you post due to your *bias*, *wrong information* presented, etc.)

[ ] Not at all [ ] Slightly [ ] Somewhat [ ] Moderately [ ] Extremely

Q15. Are you considered an **<u>elite</u>** member (**vs**. an **ordinary** member) in your network of Twitter followers? (an elite member could be a social media guru, an organization, etc.)

[ ] Yes [ ] No

Q16. In your circle of the accounts you follow and those who are your followers on **<u>Twitter</u>**, what part would you be most likely to play? "**Listen**" to their ideas OR "**convince**" them of your ideas through your posting behavior?

[ ] You mainly listen to people
[ ] You are more likely to promote your opinion

Table A.1: Component Factor Analysis

Scale	Items	Factor Loadings	Cronbach	Two-sided Bootstrapped CI for alpha (%95)
Image Motivation	– Gain followers	0.54	0.72	(0.60,0.81)
	– Look clever/expert	1.01		
	– Identify with a group	0.55		
Intrinsic Motivation	– Entertain/inform	0.59	0.79	(0.68,0.87)
	– Listen/be visible	0.65		
	– Validate thoughts	0.72		
	– Amplify news	0.60		
	– Save tweets for future use	0.51		
	– Have fun	0.70		
Follower Engagement	– Positive emotions increase involvement.	0.76	0.62	(0.38,0.77)
	– Positive emotions increase followers.	0.16		
	– Negative emotions increase involvement.	0.85		
	– Negative emotions increase followers.	0.19		
User Platform Experience	– Time spent on social media	0.04	0.69	(0.55,0.78)
	– Years on platform	0.33		
	– Login/read	0.79		
	– Tweet original posts/	0.91		
	– Retweet	0.80		
	– Comment/reply	0.82		
User Authority	– Well-informed	0.73	0.53	(0.40,0.63)
	– Reliable source of news/information	0.93		
	– Elite platform member	0.28		
	– Listen (as opposed to convincing) *	0.00		

*Item is reserve-coded in the factor.

Number of variables is 22, number of observations is 82. GFI= 0.7032, where above 0.9 is good. RMSEA (the amount of unexplained variance or residual) is 0.1266 (0.06 is good, 0.08 is adequate). Chi-square is significant at 0.0001 level. CFI should be above 0.9.

Factor loadings are presented in table above. With few exceptions, most items have loadings above 0.5.

REFERENCES

Ackerberg DA, Botticini M (2002) Endogenous matching and the empirical determinants of contract form. *Journal of Political Economy*. 110(3):564-91.

Akkus O, Cookson JA, Hortaçsu A (2015) The determinants of bank mergers: A revealed preference analysis. *Management Science*. 62(8):2241-58.

Allsop, Dee T., Bryce R. Bassett, and James A. Hoskins (2007), “Word-of-Mouth Research: Principles and Applications,” Journal of Advertising Research, 47 (4), 388–411.

Amaldoss W, Staelin R. Cross-function and same-function alliances: how does alliance structure affect the behavior of partnering firms? *Management Science*. 2010 Feb;56(2):302-17.

Anderson, Chris. The long tail: Why the future of business is selling less of more. Hachette Books, 2006.

Aral S, Walker D (2011) Creating social contagion through viral product design: A randomized trial of peer influence in networks. Management Sci. 57(9):1623–1639.

Argyres NS, Zenger TR. Capabilities, transaction costs, and firm boundaries. *Organization Science*. 2012 Dec;23(6):1643-57.

Ariely D, Bracha A, Meier S (2009) Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. Amer. Econom. Rev. 99(1):544–555.

Ascarza, Eva, Anja Lambrecht, Naufel Vilcassim. 2012. When talk is free: The effect of tariff structure on usage under two-and three-part tariffs. *Journal of Marketing Research* 49(6) 882–899.

Ascarza, Eva, Bruce GS Hardie. 2013. A joint model of usage and churn in contractual settings. Marketing Science 32(4) 570–590.

Ascarza, Eva, Raghuram Iyengar, Martin Schleicher. 2015. The perils of proactive churn prevention using plan recommendations: Evidence from a field experiment. *Journal of Marketing Research*.

Ashraf, Nava, James Berry, Jesse M Shapiro. 2007. Can higher prices stimulate product use? evidence from a field experiment in zambia. Tech. rep., *National Bureau of Economic Research*.

Bajari, Patrick, and Jeremy T. Fox. Measuring the efficiency of an FCC spectrum auction. No. w11671. *National Bureau of Economic Research*, 2005.

Bakshy, Jake M. Hofman, Winter A. Mason, and Duncan J.Watts. Everyone’s an influencer: Quantifying influence on Twitter. In Proc. 4th ACM WSDM, 2011.

Banal-Estañol, Macho-Stadler, Pérez-Castrillo. Endogenous Matching in University-Industry Collaboration: Theory and Empirical Evidence from the United Kingdom. *Management Science*. 2018 April; 64(4)

Banerjee AV, Duflo E. Reputation effects and the limits of contracting: A study of the Indian software industry. The *Quarterly Journal of Economics*. 2000 Aug 1;115(3):989-1017.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology, 51, 1173-1182.

Becker GS. A theory of marriage: Part I. *Journal of Political economy*. 1973 Jul 1;81(4):813-46.

Beer R, Ahn HS, Leider S. Can trustworthiness in a supply chain be signaled?. *Management Science*. 2017 Sep 13.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. Amer. Econom. Rev. 96(5):1652–1678.

Benaroch M, Lichtenstein Y, Fink L. Contract Design Choices and the Balance of Ex-Ante and Ex-Post Transaction Costs in Software Development Outsourcing.

Bergen M, Dutta S, Walker Jr OC. Agency relationships in marketing: A review of the implications and applications of agency and related theories. *Journal of Marketing*. 1992 Jul 1:1-24.

Berger, Jonah, and Katherine L. Milkman. "What makes online content viral?." Journal of marketing research 49.2 (2012): 192-205.

Bonacich P. Power and centrality: A family of measures. American journal of sociology. 1987 Mar 1;92(5):1170-82.

Bönte W, Wiethaus L. Knowledge disclosure and transmission in buyer–supplier relationships. *Review of Industrial Organization*. 2007 Dec 1;31(4):275-88.

Borgatti SP, Halgin DS. On network theory. *Organization Science*. 2011 Oct;22(5):1168-81.

Boyd, Danah, Scott Golder, and Gilad Lotan System Sciences (HICSS), 2010 43rd Hawaii International Conference on. IEEE, 2010

Boyle, Michael P., et al. "Expressive responses to news stories about extremist groups: A framing experiment." *Journal of Communication* 56.2 (2006): 271-288.

Breschi S, Lissoni F. Knowledge spillovers and local innovation systems: a critical survey. Industrial and corporate change. 2001 Dec 1;10(4):975-1005.

Bronnenberg, Bart J, Jean-Pierre Dub´e, Carl F Mela. 2010. Do digital video recorders influence sales*? Journal of Marketing Research* 47(6) 998–1010.

Bughin JR (2007) How companies can make the most of usergenerated content. McKinsey Quart. (August):1–4.

Burt RS. Brokerage and closure: An introduction to social capital. *Oxford university press*; 2005 Aug 11.

Burt RS. The contingent value of social capital. *Knowledge and Social Capital* 2000 (pp. 255-286).

Burt, R. 2000. The Network Structure of Social Capital. In R. Sutton and B. Staw, Research in Organizational Behavior. Vol. 22. JAI Press (2000)

Cha, M., H. Haddadi, F. Benevenuto, and P. K. Gummadi (2010). Measuring user inuence in Twitter: The million follower fallacy. In W. W. Cohen and S. Gosling (Eds.), ICWSM. The AAAI Press.

Chen Y, Wang Q, Xie J (2011) Online social interactions: A natural experiment on word of mouth versus observational learning.J. Marketing Res. 48(2):238–254.

Christophe Van den Bulte and Yogesh V. Joshi. New product diffusion with influentials and imitators. *Marketing Science*, 2007.

Coleman, J. S. 1988. Social Capital and the Creation of Human Capital. American Journal of Sociology, n. 94, p. S95-S120.

Corcoran S (2009) Defining earned owned and paid media. Forrester Research (December 16).

DellaVigna, Stefano, Ulrike Malmendier. 2006. Paying not to go to the gym. The American Economic Review 694–719.

Dibbern, Jens, Jessica Winkler, and Armin Heinzl. "Explaining variations in client extra costs between software projects offshored to India." *MIS quarterly* 32.2 (2008): 333-366.

DiMaggio, Paul, and Hugh Louch. "Socially embedded consumer transactions: For what kinds of purchases do people most often use networks?." *American sociological review* (1998): 619-637.

Eyster, Erik. 2002. Rationalizing the past: A taste for consistency. Nuffield College Mimeograph .

Fox, Jeremy T. "Identification in matching games." *Quantitative Economics* 1.2 (2010): 203-254.

Fox, Jeremy T. Estimating matching games with transfers. No. w14382. *National Bureau of Economic Research*, 2008.

Freeman, Linton C., Douglas Roeder, and Robert R. Mulholland. "Centrality in social networks: II. Experimental results." *Social networks* 2.2 (1979): 119-141.

Friedman, Daniel, Kai Pommerenke, Rajan Lukose, Garrett Milam, Bernardo A Huberman. 2007. Searching for the sunk cost fallacy. *Experimental Economics* 10(1) 79–104.

Garnefeld, Ina, Andreas Eggert, Sabrina V Helm, Stephen S Tax. 2013. Growing existing customers' revenue streams through customer referral programs. *Journal of Marketing* 77(4) 17–32.

Gefen, David, Simon Wyss, and Yossi Lichtenstein. "Business familiarity as risk mitigation in software development outsourcing contracts." *MIS quarterly* (2008): 531-551.

Gensler, Sonja, Peter Leeflang, Bernd Skiera. 2012. Impact of online channel use on customer revenues and costs to serve: Considering product portfolios and self-selection. *International journal of research in marketing* 29(2) 192–201.

Glazer A, Konrad KA (1996) A signaling explanation for charity. Amer. Econom. Rev. 86(4):1019–1028

Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication" Marketing Science, 23 (4), 545–60.

Gopal, Anandasivam, et al. "Contracts in offshore software development: An empirical analysis." *Management Science* 49.12 (2003): 1671-1683.

Granovetter, Mark. Getting a job: A study of contacts and careers. University of Chicago Press, 1995.

Grant, Robert M. "Toward a knowledge-based theory of the firm." *Strategic management journal* 17.S2 (1996): 109-122.

Grewal, Rajdeep, Gary L. Lilien, and Girish Mallapragada. "Location, location, location: How network embeddedness affects project success in open source systems." *Management Science* 52.7 (2006): 1043-1056.

Gu, Xing Sam, Paul R Rosenbaum. 1993. Comparison of multivariate matching methods: Structures, dis- tances, and algorithms. *Journal of Computational and Graphical Statistics* 2(4) 405–420.

Gulati, Ranjay. "Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances." *Academy of management journal* 38.1 (1995): 85-112.

Gupta, Sunil. 1988. Impact of sales promotions on when, what, and how much to buy. Journal of Marketing research 342–355.

Harbaugh WT (1998a) The prestige motive for making charitable transfers. Amer. Econom. Rev. 88(2):277–282.

Harbaugh WT (1998b) What do donations buy? A model of philanthropy based on prestige and warm glow. J. Public Econom.67(2):269–284.

Harris, Jacob (2010), "How Often Is the Times Tweeted," New York Times Open Blog, (April 15), [available at http://open. blogs.

Hars A, Ou S (2002) Working for free? Motivations of participating in open source projects. Internat. J. Electronic Commerce 6(3):25–39.

Hennig-Thurau T, Gwinner KP, Walsh G, Gremler DD (2004) Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? J. Interactive Marketing 18(1):38–52.

Hirano, Keisuke, Guido Imbens, Geert Ridder. 2000. Efficient estimation of average treatment effects using the estimated propensity score.

Hitsch, Gunter J., Ali Hortaçsu, and Dan Ariely. "Matching and sorting in online dating." American Economic Review 100.1 (2010): 130-63.

Ho, Daniel E, Kosuke Imai, Gary King, Elizabeth A Stuart. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis* 15(3) 199–236

Homans, George C. (1958), "Social Behavior as Exchange," American Journal of Sociology, 63 (6), 597–606.

Inkpen, Andrew C., and Eric WK Tsang. "Social capital, networks, and knowledge transfer." *Academy of management review* 30.1 (2005): 146-165.

Iyengar, Raghuram, Asim Ansari, Sunil Gupta. 2007. A model of consumer learning for service quality and usage. *Journal of Marketing Research* 44(4) 529–544.

Kang, Min-Ping, Joseph T. Mahoney, and Danchi Tan. "Why firms make unilateral investments specific to other firms: The case of OEM suppliers." *Strategic Management Journal* 30.2 (2009): 117-135.

Karlan, Dean S. 2005. Using experimental economics to measure social capital and predict financial decisions. *American Economic Review* 1688–1699.

Katona Z, Zubcsek PP, Sarvary M (2011) Network effects and personal influences: Diffusion of an online social network. J. Marketing Res. 48(3):425–443.

Kogut, Bruce, and Udo Zander. "Knowledge of the firm, combinative capabilities, and the replication of technology." *Organization science* 3.3 (1992): 383-397.

Kokkodis, Marios, and Panagiotis G. Ipeirotis. "Reputation transferability in online labor markets." *Management Science* 62.6 (2015): 1687-1706.

Kumar V (2009) Why do consumers contribute to connected goods? A dynamic game of competition and cooperation in social networks. Working paper, Tepper

Kumar, V, Xi Zhang, Anita Luo. 2014. Modeling customer opt-in and opt-out in a permission-based marketing context. *Journal of Marketing Research* 51(4) 403–419.

Lambrecht, Anja, Bernd Skiera. 2006. Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *Journal of marketing Research* 43(2) 212–223.

Lambrecht, Anja, Catherine Tucker, and Caroline Wiertz. Should You Target Early Trend Propagators? Evidence from Twitter. Working paper, 2015.

Lambrecht, Anja, Catherine Tucker, and Caroline Wiertz. Should You Target Early Trend Propagators? Evidence from Twitter. Working paper, 2015.

Lee, Leonard, Claire I Tsai. 2014. How price promotions influence postpurchase consumption experience over time. *Journal of Consumer Research* 40(5) 943–959.

Lemmens, Aur´elie, Christophe Croux. 2006. Bagging and boosting classification trees to predict churn. *Journal of Marketing Research* 43(2) 276–286.

Lerner J, Tirole J (2005) The economics of technology sharing: Open source and beyond. J. Econom. Perspect. 19(2):99–120.

Levina, Natalia, and Jeanne W. Ross. "From the vendor's perspective: exploring the value proposition in information technology outsourcing." *MIS quarterly* (2003): 331-364.

Lilien, Gary L., et al. "Calculating, creating, and claiming value in business markets: Status and research agenda." *Marketing Letters* 21.3 (2010): 287-299.

Lin, Mingfeng, Yong Liu, and Siva Viswanathan. "Effectiveness of reputation in contracting for customized production: Evidence from online labor markets." *Management Science* (2016).

Lin, N. 2001. Social capital. A theory of social structure and action, Cambridge, Cambridge University Press.

Luo, Lan, Brian T. Ratchford, and Botao Yang. "Why we do what we do: A model of activity consumption." Journal of Marketing Research 50, no. 1 (2013): 24-43.

M. U. Ilyas and H. Radha. Identifying influential nodes in online social networks using principal component centrality. In Communications (ICC), IEEE International Conference. IEEE, 2011.

Manski, Charles F. "Maximum score estimation of the stochastic utility model of choice." *Journal of econometrics* 3.3 (1975): 205-228.

Matzkin, Rosa L. "Nonparametric identification and estimation of polychotomous choice models." *Journal of Econometrics* 58.1-2 (1993): 137-168.

Mayer, Kyle J. "Spillovers and governance: An analysis of knowledge and reputational spillovers in information technology." *Academy of Management Journal* 49.1 (2006): 69-84.

Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and P Krishna Gummadi. Measuring user influence in Twitter: The million follower fallacy. ICWSM, 2010.

Mela, Carl F, Sunil Gupta, Donald R Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing research* 248–261.

Mithas, Sunil, Mayuram S Krishnan, Claes Fornell. 2005. Why do customer relationship management appli- cations affect customer satisfaction? *Journal of Marketing* 69(4) 201–209.

Moon, Sangkil. "An empirical investigation of dual network effects in innovation project development." *Journal of Interactive Marketing* 25.4 (2011): 215-225.

Muller, Eitan, and Renana Peres. "The Effect of Social Networks Structure on Innovation Performance: *A Review and Directions for Research*." (2017).

Neslin, Scott A. 2002. Sales promotion. *Handbook of marketing* 13 311–338.

Neves, Alan, et al. "Quantifying Complementarity among Strategies for Influencers' Detection on Twitter1." *Procedia Computer Science* 51 (2015): 2435-2444.

Ni, Jian, and Kannan Srinivasan. "Matching in the sourcing market: A structural analysis of the upstream channel." *Marketing Science* 34.5 (2015): 722-738.

Ohanian, Roobina. "Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness." Journal of advertising 19.3 (1990): 39-52.

Pack, Howard, and Kamal Saggi. "Vertical technology transfer via international outsourcing." *Journal of Development Economics* 65.2 (2001): 389-415.

Packard, Grant, et al. "The role of network embeddedness in film success." *International Journal of Research in Marketing* 33.2 (2016): 328-342.

Pakes, Ariel, and David Pollard. "Simulation and the asymptotics of optimization estimators." Econometrica*: Journal of the Econometric Society* (1989): 1027-1057.

Pakes, Ariel. "Alternative models for moment inequalities." *Econometrica* 78.6 (2010): 1783-1822.

Paulk, Mark. "Capability maturity model for software." *Encyclopedia of Software Engineering* (1993).

Portes, Alejandro. "Social capital: Its origins and applications in modern sociology." *Annual review of sociology* 24.1 (1998): 1-24.

Ratchford, Brian T. "The economics of consumer knowledge." *Journal of Consumer Research* 27.4 (2001): 397-411.

Recuero, Raquel, Ricardo Araujo, and Gabriela Zago. "How does social capital affect retweets?." ICWSM. 2011.

Roberts JA, Hann I-H, Slaughter SA (2006) Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects. Management Sci 52(7):984–999.

Romanelli, Elaine. "Environments and strategies of organization start-up: Effects on early survival." *Administrative Science Quarterly* (1989): 369-387.

Rosenbaum, Paul R, Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1) 41–55.

Rosenbaum, Paul R, Donald B Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39(1) 33–38.

Roth, Alvin E., and Marilda Sotomayor. "Two-sided matching." *Handbook of game theory with economic applications* 1 (1992): 485-541.

Rouse, Anne C., and Brian J. Corbitt. "Business process outsourcing: the hysteresis effect and other lessons." *Information Systems Outsourcing*. Springer, Berlin, Heidelberg, 2006. 583-602.

Rubin, Donald B, Richard P Waterman. 2006. Estimating the causal effects of marketing interventions using propensity score methodology. *Statistical Science* 206–222.

Schweidel, David A, Eric T Bradlow, Peter S Fader. 2011. Portfolio dynamics for customers of a multiservice provider. Management Science 57(3) 471–486.

Shampanier, Kristina, Nina Mazar, Dan Ariely. 2007. Zero as a special price: The true value of free products. *Marketing science* 26(6) 742–757.

Shapley, L. S., and Martin Shubik. *Game Theory in Economics*: Chapter 3: the" rules of the Games." Rand, 1972.

Shiv, Baba, Ziv Carmon, Dan Ariely. 2005. Placebo effects of marketing actions: Consumers may get what they pay for. *Journal of marketing Research* 42(4) 383–393.

Shoemaker, Pamela J., and Stephen D. Reese. Mediating the message in the 21st century: A media sociology perspective. Routledge, 2013.

Shriver SK, Nair HS, Hofstetter R (2013) Social ties and usergenerated content: Evidence from an online social network. Management Sci., ePub ahead of print February 1, http://dx.doi .org/10.1287/mnsc.1110.1648.

Smelser, Neil J., and Richard Swedberg. "The sociological perspective on the economy." (1994).

Snir, Eli M., and Lorin M. Hitt. "Vendor screening in information technology contracting with a pilot project." *Journal of Organizational Computing and Electronic Commerce* 14.1 (2004): 61-88.

Sørensen, Morten. "How smart is smart money? A two-sided matching model of Venture Capital." *The Journal of Finance* 62.6 (2007): 2725-2762.

Srivastava, Rajendra K., Liam Fahey, and H. Kurt Christensen. "The resource-based view and marketing: The role of market-based assets in gaining competitive advantage." *Journal of management* 27.6 (2001): 777-802.

Stephen AT, Galak J (2012) The effects of traditional and social earned media on sales: A study of a microlending marketplace. J. Marketing Res. 49(5):624–639.

Susarla, Anjana, and Anitesh Barua. "Contracting efficiency and new firm survival in markets enabled by information technology." *Information Systems Research* 22.2 (2011): 306-324.

Susarla, Anjana. "Contractual flexibility, rent seeking, and renegotiation design: an empirical analysis of information technology outsourcing contracts." *Management Science* 58.7 (2012): 1388-1407.

Tiwana, Amrit, and Ashley A. Bush. "A comparison of transaction cost, agency, and knowledge-based predictors of IT outsourcing decisions: A US-Japan cross-cultural field study." *Journal of Management Information Systems* 24.1 (2007): 259-300.

Toubia, O. and A. T. Stephen (2013). Intrinsic vs. image-related utility in social media: Why do people contribute content to Twitter? Marketing Science 32 (3), 368{392.

Train, Kenneth E, Daniel L McFadden, Moshe Ben-Akiva. 1987. The demand for local telephone service: A fully discrete model of residential calling patterns and service choices. *The RAND Journal of Economics* 109–123.

Tucker, Catherine E. "The reach and persuasiveness of viral video ads." Marketing Science 34.2 (2014): 281-296.

Uzzi, Brian. "Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing." *American sociological review* (1999): 481-505.

Vaynerchuck, G. (2013). How to master the 4 big social-media platforms. Inc. Magazine.

Wang, Rui, Aditya Gupta, and Rajdeep Grewal. "Mobility of top marketing and sales executives in business-to-business markets: a social network perspective." *Journal of Marketing Research* 54.4 (2017): 650-670.

Williamson, Oliver E. "The vertical integration of production: market failure considerations." *The American Economic Review* 61.2 (1971): 112-123.

Williamson, Oliver E. "Transaction cost economics: The natural progression." *Journal of Retailing* 86.3 (2010): 215-226.

Williamson, Oliver E. "Transaction-cost economics: the governance of contractual relations." *The journal of Law and Economics* 22.2 (1979): 233-261.

Wojnicki, Andrea C. and Dave Godes (2008), "Word-of-Mouth as Self-Enhancement," working paper, University of Toronto.

Wooldridge, Jeffrey M. 2010. Econometric analysis of cross section and panel data . *MIT press*.

Wuyts, Stefan, and Inge Geyskens. "The formation of buyer—supplier relationships: detailed contract drafting and close partner selection." *Journal of Marketing* 69.4 (2005): 103-117.

Yang, Yupin, Mengze Shi, and Avi Goldfarb. "Estimating the value of brand alliances in professional team sports." *Marketing Science* 28.6 (2009): 1095-1111.

Yoganarasimhan, Hema. "The value of reputation in an online freelance marketplace." *Marketing Science* 32.6 (2013): 860-891.

Zamudio, César. "Matching with the stars: How brand personality determines celebrity endorsement contract formation." *International Journal of Research in Marketing*33.2 (2016): 409-427.

BIOGRAPHICAL SKETCH

Fereshteh Zihagh was born in Tehran, Iran. After completing her schoolwork at Sampad High School in Tehran in 2006, Fereshteh entered Iran University of Science and Technology in Tehran, Iran, and received a BS in Industrial Engineering in 2011. She then entered the Graduate School of Sharif University and obtained her MBA in 2014. Fereshteh entered the Management Science Program of the Naveen Jindal School of Management at The University of Texas at Dallas in August 2014 to obtain her PhD in Management Science with a concentration in Marketing.

CURRICULUM VITAE

Fereshteh Zihagh

The University of Texas at Dallas
Naveen Jindal School of Management
800 West Campbell Road
Richardson, TX 75080-3021

Email fereshteh.zihagh@utdallas.edu

Education

2014-present	PhD in Marketing (Expected August 2019) University of Texas at Dallas, Naveen Jindal School of Management
2011-2014	Master of Business Administration (MBA) Sharif University of Technology
2006-2011	BSc in Industrial Engineering Iran University of Science and Technology

Research Interests

Substantive	Social Networks, User Engagement, Behavioral Pricing, CRM
Methods	Econometrics, Matching, Field Experiments

Dissertation Essays

Zihagh, Fereshteh, Brian T. Ratchford, and Xiaolin Li, "*Structural Embeddedness and Business Partner Selection: A Network Perspective*," To be submitted to *Management Science.*

Zihagh, Fereshteh, Brian T. Ratchford, Ali Rasouli, "The Effect of Promotions and Tariff Structure on Revenue and Churn: Evidence from an Online Field Experiment," To be submitted to *Journal of Marketing.*

Zihagh, Fereshteh, Brian T. Ratchford, "The Role of Motivations and Emotional Tone in Message Virality: Evidence from Twitter," Data analysis stage.
Target journal: *Journal of the Academy of Marketing Science (JAMS)*

Working Papers

Zihagh, Fereshteh, "Sales Dynamics for an Online Service Provider: the Role of Product Involvement and Purchase Timing." *Writing Stage*

Zihagh, Fereshteh, and Ram C. Rao, "Users' Fundraising Behavior on Online Platforms: the Role of Popularity Information." *Data Analysis Stage.*

Zihagh, Fereshteh, and Brian T. Ratchford, "When and Why Too Many Service Features Backfire: Evidence from an Online Service Provider." *Data Analysis Stage*

Presentations

Zihagh, Fereshteh**,** Brian T. Ratchford, and Xiaolin Li, "Structural Embeddedness and Business Partner Selection: A Network Perspective.

40th ISMS Marketing Science Conference, Temple University — *June 2018*
Marketing Seminar, Northwestern University — *October 2018*

Zihagh, Fereshteh, and Brian T. Ratchford, "The Effect of Promotions and Tariff Structure on Revenue and Churn: Evidence from an Online Field Experiment."

39th ISMS Marketing Science Conference, University of Southern California — *June 2017*

Awards & Honors

2017	ISMS Doctoral Consortium Fellow, University of Southern California
2017	Quantitative Marketing and Structural Economics Workshop Fellow, Olin Business School
2017	UH Marketing Doctoral Consortium Fellow
2016	Graduate Fellowship, University of Texas at Dallas
2014-present	Current Ph.D. Scholarship, University of Texas at Dallas
2011	Ranked 20th among more than 80,000 participants, Iranian National University Entrance Exam for graduate programs
2006	Ranked in top 0.1% among more than 1 million participants, Iranian National University Entrance Exam for undergraduate programs
2000-2006	Qualified to attend NODET (National Organization for Development of Exceptional Talents)

Teaching Experience

Instructor

Fall 2017	Principles of Marketing, Evaluation: **4.63/5** (Class Size=48)
Spring 2019	Principles of Marketing, Evaluation: In Progress (Class Size=60)

Teaching Assistant

2015-2018	Capstone Course in Marketing (Online Class,) Predictive Analytics using SAS, Marketing Management, Marketing Research, Consumer Behavior, Principles of Marketing

Computer Skills

R, SAS, MATLAB, MySQL, STATA, Maple, SPSS, phpMyAdmin

Related Coursework & Training

Marketing Analytics

Seminar in Consumer Search	*Brian T. Ratchford*
Seminar in Bayesian Dynamic Methods	*Norris Bruce*
Seminar in Applied Marketing Econometrics	*B.P.S. Murthi*
Seminar in Consumer Choice	*Brian T. Ratchford*
Seminar in Online Marketing	*Brian T. Ratchford*

Special Topics in Marketing

Seminar in Digital Marketing	*Ram C. Rao*
Seminar in Market Design	*Ernan Haruvy*
Seminar in Special Topics in Marketing Management	*Dmitri Kuksov*

Methodology & Statistics

Optimization	*Milind Dawande*
Advanced Managerial Economics	*Kyle Hyndman*
Statistical Inference	*Sam Efromovich*
Econometrics I	*Donggyu Sul*
Econometrics II	*Dong Li*
Industrial Organization	*Bernhard Ganglmair*
Game Theory	*Gary Bolton*

References

Brian T. Ratchford
Professor of Marketing
Charles and Nancy Davidson Professor
Naveen Jindal School of Management
University of Texas at Dallas
Email: btr051000@utdallas.edu
Phone: +1 (972) 883-5975

Xiaolin Li
Assistant Professor of Marketing
Naveen Jindal School of Management
University of Texas at Dallas
Email: Xiaolin.li@utdallas.edu
Phone: +1 (972) 883-5821

B.P.S. Murthi
Professor of Marketing
Director, Morris Hite Center
Naveen Jindal School of Management
University of Texas at Dallas
Email: murthi@utdallas.edu
Phone: +1 (972)-883-6355