January 2013

The Impact of Organizations' Collaboration Strategies and Alliance Network Positions on Invention Performance

Fethullah Caliskan
University of South Florida, fcaliska@mail.usf.edu

The Impact of Organizations' Collaboration Strategies and

Alliance Network Positions on Invention Performance

by

Fethullah Caliskan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Kingsley A. Reeves, Jr., Ph.D.
John Skvoretz, Ph.D.
Adriana Iamnitchi, Ph.D.
Tapas Das, Ph.D.
Alex Savachkin, Ph.D.

Date of Approval:
November 12, 2013

Keywords: Innovation networks, social network analysis, collaborative R&D, technological knowledge creation, network simulation

DEDICATION

I dedicate this dissertation to my family, particularly to my wonderful wife, Ezgi, who has supported and encouraged me throughout my endeavor, and to our precious daughter Isranur Mina, who is the joy of our lives.

I would like to thank my beloved and respected mother and father Muazzez Ayse and Suleyman Caliskan, the best sisters and brothers in the world, Handenur, Sinan, Ahmet and Aysun for their love. I send my special thanks to my mother-in-law Huriye, for her selfless dedication to Isranur, Ezgi and me.

ACKNOWLEDGMENTS

I cannot express the degree of my gratitude to my adviser, Dr. Kingsley A. Reeves, who has inspired, encouraged, guided and supported at the times that I most needed. He has been the true meaning of a teacher to me.

I would like to extend my appreciation to both Dr. Jose Zayas-Castro and Dr. Tapas Das for providing support and improvement opportunities during their chairmanship of our department.

I have been very lucky to enjoy having a dissertation committee that perfectly fits with my research needs. I would like to thank Dr. John Skvoretz for effectively teaching me social network analysis methods and putting his time and expertise to help, Dr. Adriana Iamnitchi for providing valuable feedback, especially in computer simulation stage of the dissertation, Dr. Aliaksei Savachkin for addressing the improvement areas and for his motivational support, and Dr. Tapas Das for his encouragement that kept my momentum consistently increasing.

I also wish to thank Dr. Ozan Ozcan, Omar Alrehaili, Irene Zubillaga Lopez, and Jennyfer Sanchez Cumare for their help at various stages of the research.

Finally, I would like to express my gratitude to Dr. Karen Liller for the USF Challenge Grant Program, which gave opportunities for my research to get promoted and supported, Dr. Huseyin Arslan for being the chair of my dissertation defense meeting and to Ms. Catherine Burton for her patience and corrections on my dissertation draft.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This research investigates the impact of organizations' collaboration strategies and network positional variables on invention performance. Organizations, particularly those pursuing a differentiation strategy, are motivated to introduce novel products and services in order to remain competitive. Thus, two questions of interest to such organizations regarding the network dynamics of the invention process are: 1) What kind of strategies allow them to attain superior invention results? 2) What is the most advantageous structural positioning in a collaborative network of innovators? Three independent studies attempt to find answers to these questions by using three complementary study approaches.

In the first study, in order to examine invention performance of organizations with different collaboration strategies, a simulation model is constructed and run at various levels of technological dynamism and with various types of invention tasks. The cognitive cooperation strategy, which pursues technological knowledge integration as a priority, is found to be the most effective strategy in most experiments. Success-driven cooperators, whose priority is to collaborate with the most effective performers in the network, provided the most consistent invention performance. Moreover, different strategies are shown to perform best at certain points of the industry environment space spanned by technological dynamism and invention type.

The second study investigates the impact of positional metrics in a collaboration network on the inventive performance of organizations (as measured by the number of patents issued) by using survey data. Twenty-eight high-tech companies and universities located in Florida are

surveyed to reveal their collaborative networking map. Network structural measures are obtained in order to test the hypotheses that high values in centrality metrics correspond with higher inventive performance. The regression analyses imply that degree and closeness centralities are predictive indicators of high inventive performance but the data does not support the significance of impact of local clustering.

The third study analyzes the impact of positional metrics on inventive performance by using a national database for the announced research and production joint ventures. From the most inventive organizations (in terms of patent counts) 63 of them are selected for analysis. 371 research and production joint ventures are analyzed to identify their relationship network every year from 1994 to 2012. The results indicate that the number of joint venture partners, being close to other members of the network through network connections and strong local connectivity (except for their interaction effect) is associated with higher invention performance.

All three studies bring new methodological contributions to the area of invention network research. The simulation study offers a new model in the area of collaborative invention networks. Furthermore, the ideas and practices developed during the construction of the agent based simulation model for the invention network can be adapted to similar areas of application. The survey study offers a holistic approach for the definition of connections in the development of invention network map and empirically tests it. The database study longitudinally analyzes the evolution of a highly accurate joint venture network over 19-year period while testing the impact of positional metrics with un-weighted and weighted calculation methods.

Solutions to our health problems, communication or transportation needs etc. are not usually found due to some series of fortunate events. They are the product of an effective recombination process of technological knowledge. Moreover, effective invention performance

is not only important for organizations individually, it is crucial for governments that are concerned with the problems of its citizens. Effective ways of facilitating the recombination of technological knowledge are addressed and presented to inform both companies and policy makers. Better understanding of the dynamics of the invention process will bring more solutions to existing problems.

CHAPTER 1: INTRODUCTION

One of the key drivers of technological advancement and sustainable growth is discovery that is followed by innovation. The technological knowledge needed for innovation, however, may lie beyond an organization's own capabilities. A solution to address this problem is to form alliances with other organizations. Forming alliances provides important benefits such as better financial and other resource capacity, the ability to pool the technological knowledge of partners, better infrastructure, reduction of risks, and market penetration.

Invention, defined as the first proof of possibility in industrial practice for something previously not demonstrated, is a product of a unique process that involves the novel recombination and reconfiguration of the ways in which the technological knowledge elements are linked together (2). Moreover, it is collective in nature (1). I adopt a network interaction perspective to explain the technological knowledge recombination and invention process. Organizations, embedded in the network of relationships, use those relationships as conduits of information and know-how. An illustration for the development of an inter-organizational relationship network is given in the successive pictures of Figure 1.1. In very simplified terms, a few organizations form an alliance (to solve an industrial problem, to make a purchase agreement, or otherwise allow sharing of technological knowledge) in Figure1.1 (a), then an external organization makes another alliance relationship with an existing member Figure1.1 (b) and they eventually weave a network of direct and indirect relationships.

Researchers have attempted to explain the invention process and have studied several aspects of invention through network relationships. Organizations are motivated to introduce novel products and services in order to establish or maintain a competitive advantage. Thus, two questions of interest to organizations regarding the dynamics of the invention process are: 1) What kind of strategies allow organizations to attain superior invention results? 2) What is the most advantageous structural positioning in a collaborative network of innovators?

Effective innovation performance is not only important for organizations individually; it is crucial for industries and even governments that are concerned with the urgent problems of society. It is probably not fully acknowledged but “waste” is not only in terms of the assets that we can see but it can well be in terms of the “lost” or “locked” technological knowledge or what may be called wasted opportunity. Solutions to simple problems may lie just beyond untapped collaboration opportunities.

Figure 1.2 illustrates a bigger picture that the studies of the following chapters address. The first research question relates to the impact of the collaboration strategies on network formation and consequently the invention performance. The second question is specifically about the impact of the network positional metrics on invention performance.

In order to address the research questions, three independent studies are conducted. Chapter 2 presents a simulation study that examines the impact of partner selection strategies on invention performance, given varying degrees of need for knowledge complementariness and motivation to form alliances. Chapter 3 provides the results of a survey administered in the State of Florida where the relationship network of 51 high-tech companies and universities were obtained and the impact of network positional variables were investigated. Chapter 4 also presents research results that focus on the impact of positional metrics on invention output, but

this time a national database is utilized to obtain panel data for 63 organizations participating in 371 research and production joint ventures over a period of 19 years. Chapter 5 concludes the work and emphasizes the collective contribution of the three independent studies.

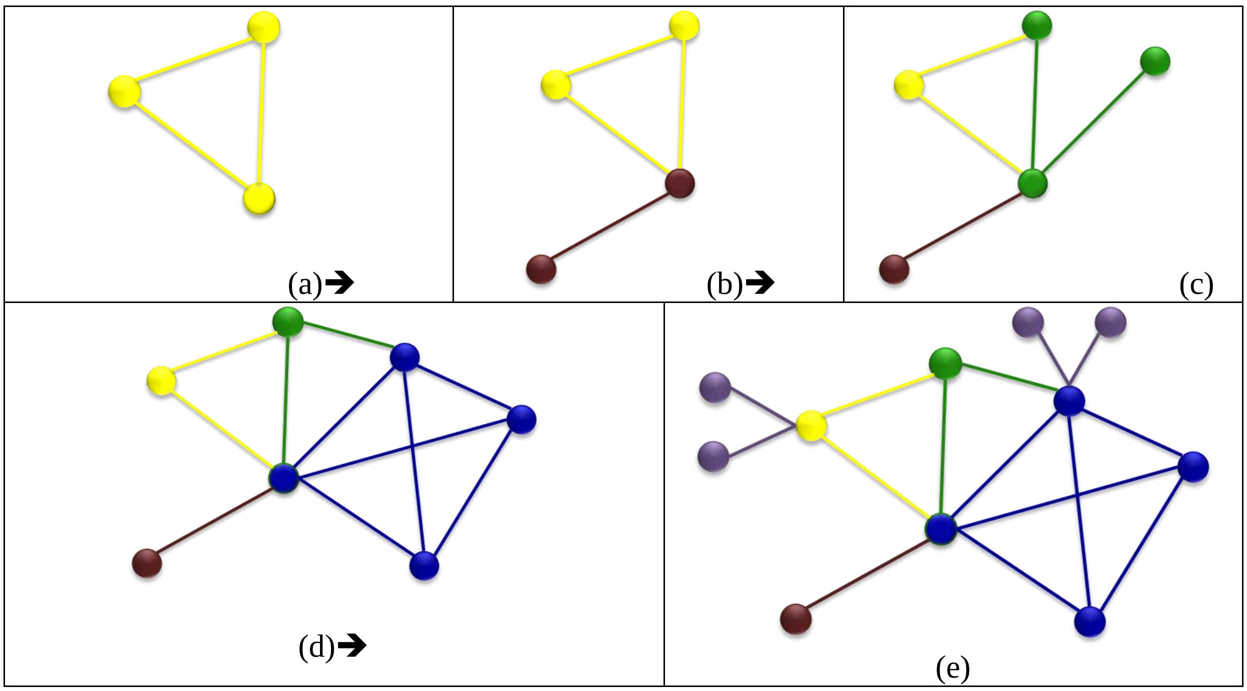


Figure 1.1. Inter-organizational network formation process

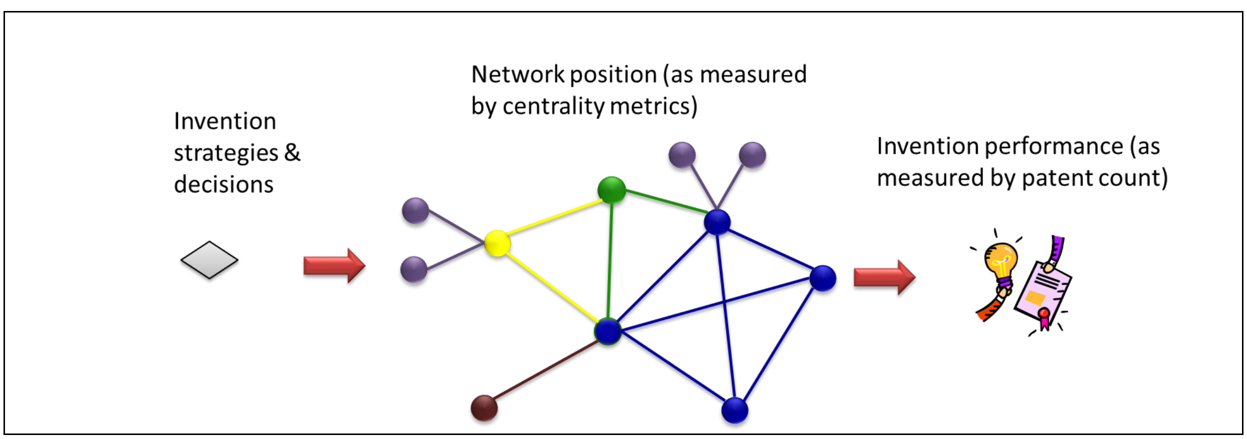


Figure 1.2. An illustration of the research questions addressed

1.1. References

1. Miettinen, R. Theories of Invention and an Industrial Innovation. Science Studies 9(2):34-48; 1996.

2. Schilling, M. A. Phelps C. C., Interfirm collaboration networks: The Impact of large-scale network structure on firm innovation. Management Science 53(7):1113-1126; 2007.

CHAPTER 2:

THE IMPACT OF COLLABORATION STRATEGIES ON INVENTION PERFORMANCE: A SIMULATION STUDY

2.1. Introduction

Over the last two decades, the rate of formation of alliances has increased and the motivation for forming alliances has shifted significantly. As knowledge exchange and technology transfer motivate alliance formation, alliances became widespread in technology-intensive industries (17). In parallel with the alliance formation trend, inventions have increasingly been the product of more than one inventor. For example, the data from a sample of 750,000 patents shows that about 65% of patents were registered by one inventor in 1970s, but only 30% in 2009 were registered by one inventor. On the other hand, patents registered by three and more inventors increased from around 10% to above 40% during the same period (8). Similarly, top innovative products borne from collaborative research and development (R&D) has increased from 53% in 1975 to 87% in 2006 (5).

In this research, organizations are considered as unit of analysis and assume that they aim to attain a high invention performance. Since invention is collective in nature (16) and the creation of inventions involves the novel recombination and reconfiguration of the ways in which the technological knowledge elements are linked together (22), an invention process can be modeled using a social network of organizations (15). Previous studies that have used a social network approach to model collaborative invention of organizations have examined: the impact

of the medium of knowledge exchange on knowledge growth (6), the impact of how knowledge is pooled and how firms derive information about potential partners on network formation (7), the role of complementary knowledge stocks and knowledge dynamics on the network formation (4), the effects of alternative configuration of the knowledge structure on the generation of new technological knowledge (3), how knowledge complementariness explain network formation (13), and how the knowledge sharing strategies effect invention network's success (11).

Firstly, to our knowledge, collaborative invention network research has not yet studied the impact of organizations' alliance strategies on their invention performances. What type of partner selection strategy would yield the best or the most consistent invention performance? For partner selection, must the emphasis be on technological capabilities, invention performance outcomes or trust?

Secondly, an answer indicating the best alliance strategy can be very inadequate if it does not take into account the necessities of various levels of knowledge overlap and technological intensities in industrial sectors. In related studies, the impact of knowledge complementariness (knowledge overlap or cognitive distance are also used) on network formation is examined but its impact on organizations' invention performances is yet to be addressed. Noteboom et al. (18) suggest that in order for cooperation to be effective, there is a certain level of knowledge complementariness necessary[1]. Sobrero and Roberts (24) also suggest that decomposability of tasks for certain inventions plays an important role on performance through affecting the level of knowledge complementariness. This raises the following question: What type of partner

[1] They use the term "cognitive distance" rather than knowledge complementariness. In several of their studies they suggest that the technological knowledge levels of the two partners must be not too close and also not too far away from each other for a successful invention process.

selection strategy would benefit most for varying degree of necessity of technological knowledge overlap?

Thirdly, economic environments or industries force organizations to form alliances at various degrees. For instance, Mowery et al. (17) state that there has been a significant increase in the rate of formation of alliances in semiconductors, computers, software, and commercial aircraft industries over the last two decades[2]. Rosenkopf and Schilling (11) calls the phenomenon technological dynamism. They show that the alliance participation rate (that is number of firms participating in alliances divided by the number of publicly held firms) varies a great deal with variances ranging from 0.05 to 2.60. Therefore, the technological dynamism in some industries force organizations to form alliances more so than in other industries. The following question arises: Given varying degrees of technological dynamism in different industries, what strategy would yield the highest and the most consistent invention performance?

To address the above questions, this research examines the impact of partner selection strategies on invention performance, given varying degrees of need for knowledge complementariness and motivation to form alliances.

This research brings two methodological contributions. All invention network simulation studies known to the author use an approach that allows only a dyadic (bilateral) partnership at any given time. This is an attempt to introduce a model that allows using multilateral partnerships at any given time (cycle). Organizations are allowed to form more than one partnership, that is, participate in more than one project or venture as in a real practice. Furthermore, simultaneously modeling the invention performances of organizations pursuing different partner selection strategies is a unique contribution.

[2] Two decades up to the time of their publication in 1996.

2.1.1. Embeddedness

Relationships among the network of organizations are defined as “embeddedness” (see Uzzi (25), and Iandoli et al. (13), for discussion of how embeddedness and network structure effect network behavior). Embeddedness implies that the structure and the quality of network connections affect and shape the economic behavior and productive performance of organizations. The literature proposes three categories of embeddedness: cognitive, relational, and structural. Cowan et al. (7) define cognitive embeddedness as organizations’ ability to effectively integrate their respective knowledge. In the model, I define one of the partner selection strategies as “Cognitive Cooperator” (CC), for which the partner selection decision is based on the most effective integration of technological knowledge. Relational embeddedness is associated with the accumulation of a trust capital (25). In the model, one of the partner selection strategies is defined as “Relational Cooperator” (RC), for which the partner selection decision is based on trust. Structural embeddedness is defined as a social control mechanism to coordinate and safeguard exchanges (Jones et al. 1997). For the purposes of the model, I adopt the concept in a slightly different way. Structural embeddedness is associated with more efficient information spread. Based on the assumption that the invention performance information is critical and available to all network members, some organizations’ partner selection decision is based on the invention performance aspect of the other members. In order to convey the best meaning, I use the term “Success-driven Cooperator”[3] (SC) instead of the term 'structural'.

2.1.2. Technological Dynamism and Uncertainty: The Alpha (α) Parameter

Industrial sectors are categorized due to various characteristics (22). An invention network study (11) that uses Thomson’s SDC Platinum database reports that, in a certain year,

[3] Alternatively, it can be called as a “Success-following Cooperator”.

there is only 1 firm that participates in a research joint venture from the carpets and rugs industry, whereas the number of firms from pharmaceuticals industry participating in a joint venture is 568. This does not necessarily mean that the carpets and rugs industry is not innovative. On the contrary, it may simply mean that for industries like carpets and rugs, organizations tend to make inventions in-house. On the other hand in pharmaceuticals industry, firms are forced to form alliances due to the high costs and potential risks. That is, my intuition is that due to the nature of their business, organizations are forced to form alliances at various levels.

When technology is changing rapidly, organizations make use of alliances more often. Furthermore, the degree of uncertainty about the direction of technological change can make alliances more attractive because they provide flexibility as compared to in-house development to solve an industrial problem (20). Altough I adopt the term "technological dynamism and uncertainty" for this phenomenon, as defined by Rosenkopf and Schilling (20), several studies define this characteristic from their perspectives. For example, Hagedoorn used the term "technology intensity of sectors" (12), Ang used "technology intensive industries" (2), Segarra-Cipres et al. used the term "technological intensity of the sector" (TIS), for similar phenomenon (23). Segarra-Cipres et al. take it one step further and parameterizes the level of TIS as 0 for less intense sectors and 1 for more intense sectors.

Thus, technological dynamism encourages alliance formation and create opportunities brings important opportunities for organizations to make inventions. It is therefore expected that technological dynamism will increase the invention performance of organizations that are motivated to invent, no matter what their partner selection strategy is. It is therefore hypothesized

in this study that all three strategies are viable and technological dynamism increases all strategy groups' invention performances.

One objective is to observe the performances of CCs, RCs and SCs over the spectrum of the industries with different levels of technological dynamism. I introduce the parameter Alpha (α) that takes real values from 0 to 1 in order to model the spectrum of technological dynamism levels in various industries. The α parameter provides a spectrum of various industries and characterizes the amount of force that is imposed on organizations to form alliances. The purpose is the identification of the α space where changes, if any, in performance of organizations pursuing different alliance strategies are observable.

2.1.3. Knowledge Complementariness: The Beta (β) Parameter

As mentioned in cognitive embeddedness, organizations seek effective integration with of technological knowledge elements with their partner. The literature suggests that cognitive distance in their technological knowledge elements must be a particular match. If firms are too close together in technological knowledge, they can understand each other well but there will be limited points to share. On the other hand, when they are too far apart, they have difficulty in understanding each other but there will be a good chance of information sharing.

Noteboom et al. (18) uses the term "cognitive distance" to define the difference between the technological knowledge levels of alliance partners. Noteboom et al. (18) and Wuyts et al. (26) provide support for their hypothesis that innovation performance is an inverted U-shaped function of the technological cognitive distance between partners. Gilsing et al. (10) adopt the same idea and use the term "technological distance" for the phenomenon.

Cowan et al. (7) use the terms knowledge pooling or knowledge complementarity and introduces the idea of "decomposability of innovation" adopted from Sobrero and Roberts (24).

If the process is composed of discrete tasks, can be done in isolation, and integrated easily, then each partner can be specialized in their area. As mentioned earlier, a definition for invention is given as recombination or reconfiguration of knowledge elements. In one extreme, the partners can be as far apart and specialized as possible for the invention process to be successful. On the other hand, if specialization is not possible, i.e., the process is composed of tasks that must be done all together, then, the weaker partner becomes a bottleneck. This is the other extreme where the technological knowledge levels of the partners must be as close as possible for the invention process to be successful. Iandoli et al. (15) uses the term "knowledge complementariness", which is the form adopted in the meaning of a mutual fit of technology levels that allows best performance.

Assuming that the best productive cooperation is ensured at different levels of knowledge complementariness, a parameter is defined where it characterizes the level at which the knowledge difference is the best fit. The Beta (β) parameter is defined from 0 to 1, where at minimum, the partners are most inventive when there is no difference between their technological knowledge levels. When Beta is at its maximum, it is a business environment in which the partners are most inventive with a maximum difference in between their technological knowledge levels. The purpose here is not to identify whether β increases the invention performance. It rather provides a spectrum of invention types in the two extremes. One purpose of this study is to show that strategies must be differentiated depending on the type of invention. The cognitive cooperation strategy is applicable for the type of invention where the tasks are decomposable and specialization is possible and desired. The relational cooperation, on the other hand, is applicable to the type of invention where the tasks are not decomposable and the partners' knowledge level are required to be similar. Success-driven cooperation is applicable to

the new entrants of the market. They need to partner up with the most successful network members. Normally, success-driven cooperators' performance is affected whether decomposability or specialization is desired.

2.1.4. Knowledge and Invention

In the context of the study, the term technological knowledge refers to a solution to a specific technology problem. Unlike many previous studies, I want to set a clear distinction between the term knowledge and invention. In this context, invention refers to the first proof of possibility in industrial practice for something previously not known or demonstrated (see Lane (14) for a discussion of states of knowledge).

As technological problems may pertain to various sectors and domains, technological knowledge is defined as categorized elements (domains). For example, an organization may have technological knowledge with five different elements, each one at varying levels. Therefore, the knowledge level for an organization can be defined as a five dimensional vector. Given the definition of invention, the demonstration is tangible (e.g., through patent registration). Therefore, invention is represented by one-dimensional scalar value; a higher the invention count means a better invention performance.

One result of partnership is that it makes partners' knowledge profile become more similar (7). In this study, when the partners collaborate, the peer who has a lower knowledge level must increase its knowledge level towards its partner's higher-level knowledge. The process of increase takes place separately for different knowledge elements. Furthermore, the rate of increase depends on the strength of their connection. If the two organizations cooperate in more than one project (or one joint venture, etc.), the rate of knowledge increase must be greater as compared to cooperation in only one project (in one venture).

Several factors play a role in a subsequent invention taking place including technological knowledge levels, resources, infrastructure, etc. One of the major decisions for finding a solution to a technological problem is whether to do it with internal resources or to form an alliance venture. Zahra (27) evaluates independent ventures and corporate ventures for their strategies and invention performances. He finds support for the hypotheses that the internal and external R&D resource usage differs significantly depending on their venturing strategy. He also shows that corporate ventures surpass independent ventures in patenting performance. In this study, the two categories of resources are identified as follows: Internal R&D capabilities are translated into individual knowledge levels whereas the external resources are translated into the amount of network connections both in number of alliances and in number of projects (or ventures) with each ally. Specifically, internal knowledge levels increase the chances of invention as long as they are high compared to the knowledge levels of the other members of the network. On the other hand, high numbers of connections increase the chances of invention too. As discussed above (the discussion of the Alpha parameter), the propensity of making alliances is adjusted by the Alpha parameter. I aim to examine the invention performance of CCs, RCs and SCs at various Alpha levels.

Another important factor that is proposed to affect invention success is the knowledge complementariness. It is my purpose to observe the performance of CCs, RCs and SCs at various conditions of best knowledge-fit. Beta is the adjusting factor for the partners' most inventive state, either when there is no difference, a maximum amount of difference or some moderate levels of difference between their technological knowledge levels.

2.1.5. Agent Based Simulation

Agent based simulation is defined as a collection of intelligent and interacting agents, which exist and operate in an environment made up of agents and their relationships. Agent based simulation is a powerful research method and is used in innovation networks in many research areas (1), including innovation networks and collaboration (9). In this study, organizations are considered as independent agents interacting with other members of network, based on a set of rules. Based on its specified category of strategy, each one goes through the stages of partnership selection, alliance formation, invention and knowledge increase as described in the next section.

2.2. Model

A narrative description along with the flowchart of the model is as follows. It is assumed that the cluster is a closed system where no external organization is allowed to enter and no member is allowed to exit. Organizations are assumed to be a part of a greater network in the economy. At each cycle, every organization attempts to make inventions, either by itself or in collaborations with other organizations. The simulation consists of five stages, the last four are repeated cycles (that may typically represent a quarter in practice) as depicted in Figure 2.1.

2.2.1. Initialization

One third of the organizations are assigned with one of the strategies (s_i): CC, RC or SC. The knowledge levels of an organization are identified by a t dimensional vector that specifies technological knowledge[4]:

[4] In the initial cycle only, knowledge element levels are produced randomly with normal distribution. The number of dimensions in K reduces the chances of disadvantage for starting the simulation with low knowledge levels. The model is tested for the effect of the randomization to

$$K_i = [k_{i1}, k_{i2}, .. k_{it}]$$

Each element k_{it} represents the technological knowledge level possessed by the i^{th} firm for the t^{th} technological element. The number of technological domains is initially selected to be five. However, the simulation has been run for several parameters including different values of t[5]. The organizations are also given a unique ID (d_i). The last element of each organizations is their invention (c_i) counts[6]. The state of each organization is, therefore, represented by the following S vector:

$$S_i = [d_i, s_{i1}, k_{i1}, k_{i2}, k_{i3}, k_{i4}, k_{i5}, c_i]$$

2.2.2. Partnership Proposals Stage

Partnership proposal is a crucial stage of the simulation where CCs, RCs and SCs differentiate in their selections strategies. When forming alliances, cognitive cooperators aim to for the most effective integration of technological knowledge for themselves. For this reason, CCs identify their two lowest-level knowledge elements:

$$L_i = [l_{i1}, l_{i2}]$$

Along with the assumption that all information regarding knowledge levels of network members is available to every member, a CC searches among the members for who has the highest level of knowledge at its lowest-level knowledge element (for its l_{i1}). Similarly, a second search is made for who among the members has the highest level of knowledge at its second lowest-level

---

see whether starting the simulation with low knowledge levels has an impact on the final invention counts. There is no significant impact of starting the simulation with one or two very-low knowledge elements on the final invention performance. After the first cycle, knowledge increase (and decrease) are calculated by the determined rules.

[5] The simulations are run at t=2 and t=20 levels.

[6] In the initial cycle only, invention counts are produced randomly with a uniform distribution. This is in order for SCs to make their initial selections. Right after (before invention stage of the first cycle), invention counts are set to ‘0’. Then, they are calculated according to the determined simulation rules, for the first and the remaining cycles.

knowledge element (l_{i2}). Finally, a CC identifies the index numbers of the two best partners for its two lowest-level knowledge elements:

$$B_i = [b_{j1}, b_{j2}]$$

In the best partners vector, b_{j1} represents that the index value of the 1st best match for i is j. The best partners for a CC can be any one of the CCs, RCs or SCs (except for itself). The 1st and the 2nd elements of B_i are defined as highly desirable and moderately desirable potential allies. The number of projects (or ventures) that a CC proposes in a cycle term is set to be 2 for highly desirable potential allies, and 1 for moderately desirable potential allies.

The project or the ventures proposed are translated into the number of potential links (e.g., connections, ties) in the network model. In this stage, these proposals are determined by all CCs and are sent to the potential allies.

Relational cooperators base their partner selection decision on trust. Based on the idea that continuing partnership enhances the relationship and reduces the risks, RCs aim to keep their current allies. For this reason, RCs produce their potential trusted partners list by identifying two allies with the highest number of connections.

$$T_i = [t_{j1}, t_{j2}]$$

In the trusted allies vector, t_{j1} represents that the index value of the 1st trusted (and also potential future) ally for i is j. Similarly, the best partners for an RC can be any one of CCs, RCs or SCs (except for itself). The 1st and the 2nd elements of T_i, are defined as highly desirable and moderately desirable potential allies respectively. The relative numbers of projects (connections) proposed are 2 and 1, respectively. The proposals are determined by all RCs and are sent to the potential allies.

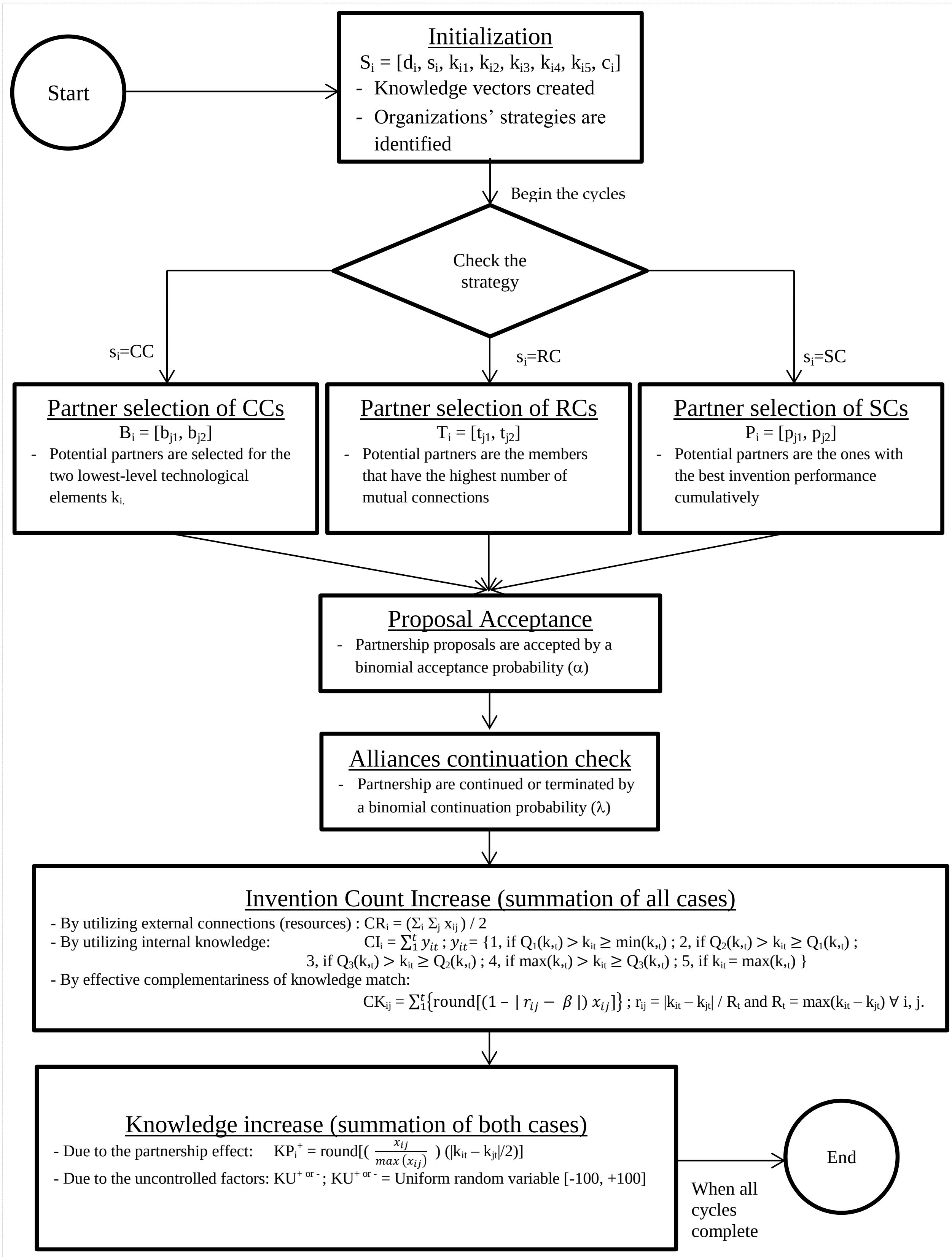


Figure 2.1. Invention network simulation stages

Success-driven cooperators make their selection decision based on the cumulative invention performance of other members. SCs want to know who has the highest number of invention counts currently. Therefore, each SC makes a search among c_j's and identifies the members with the best and the next best invention performance:

$$P_i = [p_{j1}, p_{j2}]$$

The first and the second best performers are proposed 2 and 1 projects respectively. The proposals are determined by all SCs and are sent to the potential allies.

2.2.3. Responding to Proposals Stage

Responding is basically a three-step process. Firstly, partnership proposals are evaluated and responded to as i) fully accepted, ii) partially (some of proposed projects) accepted, or iii) fully denied. To operationalize this process, a random binomial variable is produced over the number of proposals during the current cycle with a probability of α. As discussed in previous sections, alpha is the level of technological dynamism and determines the inclination to make (more at high α or less at low α) alliances. It can also be interpreted as the acceptance rate. The acceptance rate is the probability of success for a binomial random variable of the accepted proposals. It is a global parameter that is applied to all set of proposals at each cycle. So, for example, suppose that organization i proposes 2 connections to organization j. The number of accepted connections will be within {0, 1, 2} where P(0) = P(2) < P(1) where P(a) is the probability of accepting 'a' number of proposals. The accepted proposals are scheduled to be effective during the next cycle.

Secondly, at each cycle, each project has a potential to come to an end. Current connections (before accepted project proposals are put into practice) are evaluated for continuation and termination. Similarly, a random binomial variable is produced over the number

of current connections with a probability λ that is called the continuation rate. The continuation rate is a global parameter, which is applied to all current connections at each cycle.

Lastly, responses to connection proposals are obtained and added to the current connections state. That is if, say, i has 2 connections with j (e.g., i is working with j over two different projects, or as an alternative interpretation, within two different ventures) and j accepted 1 of the newly proposed connections from i, then, the number of connections between i and j is increased to 3.

2.2.4. The Stage of Invention Count Increase

There are three different ways to increase invention counts increase: i) utilize resources that are made available with connections, ii) utilize internal technological knowledge capacity, and iii) utilize the effective complementariness of the technological knowledge match. Because various industrial conditions are characterized by α (the level of need to form alliances) and by β (the knowledge difference level at which the partners are productive), the invention count increase due to (i) and (ii) will depend on the α level, and count increase due to (iii) depend on the β level.

The utilization of the external resources for (i) must be thought of in terms of financial, infrastructural resources. It is separated from the cognitive side of an invention. The increase in the invention count for i due to connections is determined by the following formula:

$$CR_i = (\Sigma_i \Sigma_j x_{ij}) / 2 \tag{2.1}$$

CR_i represents the amount of i's increase in inventions by utilizing the resources that are made available through connections and x_{ij} is i's current connections with j (CR: Count increase due to the resources). x_{ij} values are higher when the value of α is high and lower when the value of α is low. The summation over the rows and columns is divided by 2 due to technical reasons.

The adjacency matrix **X** is symmetrical (representing an undirected network) and the number of alliances at x_{ij} is similarly represented at x_{ji}. In order to get the correct number of alliances, half of the summation over the rows and columns must be taken.

In (ii), higher knowledge capacity means better chances of invention increase. The following operation is applied to realize this process. For each knowledge element (k_{i1} through k_{it}), the technological knowledge levels of the members are sorted. This sorted dataset's minimum, 1st, 2nd, 3rd quartile and max values are identified[7]. A member's invention increase due to the internal technological knowledge stock depends on the quartile bracket in which the member's technological knowledge is located for the particular $k_{,t}$ values set. Namely, the higher the quartile brackets in which a member's knowledge falls, the greater invention increase it gets. Therefore, in (ii), the increase in the invention counts is determined by the following:

$$\mathrm{CI_i} = \sum_1^t y_{it}\ ;\ y_{it} = \begin{cases} 1, \text{if } \mathrm{Q}_1(\mathrm{k}_{,t}) > k_{it} \geq \min(\mathrm{k}_{,t}) \\ 2, \text{if } \mathrm{Q}_2(\mathrm{k}_{,t}) > k_{it} \geq \mathrm{Q}_1(\mathrm{k}_{,t}) \\ 3, \text{if } \mathrm{Q}_3(\mathrm{k}_{,t}) > k_{it} \geq \mathrm{Q}_2(\mathrm{k}_{,t}) \\ 4, \text{if } \mathrm{Q}_4(\mathrm{k}_{,t}) > k_{it} \geq \mathrm{Q}_3(\mathrm{k}_{,t}) \\ 5, \text{if } k_{it} = \max(\mathrm{k}_{,t}) \end{cases} \tag{2.2}$$

CI_i represents the amount of i's increase in inventions by utilizing the internal technological knowledge stock (CI: Count increase due to internal knowledge stock) and y_i is invention increase determined at knowledge elements (1 through t). Q_1, Q_2 and Q_3 are 1st, 2nd and 3rd quartiles of the k values at tth technological knowledge domain. Those members who have knowledge levels in between the minimum level and Q_1 are given one count (because they still have some technological knowledge) but it is five times less than those who have the maximum level of technological knowledge. Obviously, when α is close to zero (that is, technological

[7] Basically, the minimum, the 25th percentile, the median, the 75th percentile and the maximum k_{it} value are determined.

inertia does not force organizations to form alliances) then i's invention productivity becomes solely from utilizing the internal technological knowledge.

The increase on invention counts due to knowledge complementariness is the cognitive side of an invention. To operationalize (iii), the minimum and the maximum knowledge levels are identified for each knowledge element (k_{i1} through k_{it}). The difference between the maximum and the minimum levels are defined as the technological range in that particular area. Now, for any partnership between i and j, the absolute value of the difference between the knowledge levels of k_{it} and k_{jt} is divided by the technological range. The yielding percentage amount is called paired-percentage that takes a real value (0, 1). When this percentage is small, we understand that there is not much discrepancy in between their knowledge levels. This process is repeated for each pair at each knowledge element. Each paired-percentage is evaluated against the β value in order to find the amount of invention count increase due to effective knowledge complementariness. The final increase due to (iii) is found by multiplying with the impact of the number of connections between i and j.

$$CK_i = CK_j = \sum_1^t \{\text{round}[(1 - |\, r_{ij} - \beta\,|)\, x_{ij}]\} \text{ ; } r_{ij} = |k_{it} - k_{jt}| / R_t \text{ and } R_t = \max(k_{it} - k_{jt}) \ \forall \ i, j \quad (2.3)$$

CK_i represents the amount of i’s increase in inventions due to effective knowledge discrepancies. r_{ij} is the paired-percentage and R_t is technological range at t^{th} technology domain in a particular cycle. The β, the invention type that defines the best knowledge discrepancy levels) is defined in (0, 1). When β is close to one (that is, partners are productive at high knowledge discrepancy) then i’s and j’s invention productivity is good if they have high discrepancy. When β is around 0.5 (that is, partners are productive at a moderate discrepancy) then i’s and j’s invention productivity is good if they have a moderate discrepancy.

2.2.5. Technological Knowledge Increase Stage

There are two ways of technological knowledge increase: i) It is possible through the inflow of knowledge due to the alliances, ii) there are also a number of other factors that we cannot control. In (i), the partner who has less knowledge increases its knowledge level towards its partner's higher-level knowledge. Just like in the process of finding the knowledge discrepancies (to find CK_i's), for any partnership between i and j, the absolute value of the difference between the knowledge levels of k_i and k_j are identified at each knowledge element. The knowledge levels' difference is multiplied by a current collaboration factor that is found by scaling the number of current connections between i and j by the maximum number of connections made between any i and j.

$$KP_i^+ = \text{round}\left[\left(\frac{x_{ij}}{max\,(x_{ij})}\right)\left(|k_{it} - k_{jt}| / 2\right)\right] \tag{2.4}$$

KP_i^+ represents the amount of increase in the knowledge levels of i.

For the factors that we cannot control, I assume a random knowledge increase or decrease. The increase is basically explained by the potential knowledge enhancement through several possible ways of the organization increasing knowledge internally. The decrease is due to the fact that an organization's loss of human capital in the knowledge area or due to the enhancement of other members of the network. Increase (or decrease) is determined by a uniform random variable with a mean value of 0. Knowledge increase (decrease) due to uncontrolled factors is operationalized by the following formula:

$$KU^{+\text{ or }-} = \text{Uniform random variable } [-100, +100] \tag{2.5}$$

2.3. Validation Tests

The code is written in R (19) using R Studio (R core team, 2013)[8]. Before running simulations of the model, a series of validation tests were applied to make sure that the code performs as intended to. A few of the important tests are listed below.

Firstly, from the initialization to the end of the first cycle of the run, all matrices and parameter values produced are observed step by step.

Secondly, the network map is constructed and visually observed for the parameter effects at each cycle. Figure 2.2. is a typical screen-shot for a 45-member network at the end of 8 cycles that shows both the network map and the invention counts of individual members. The effect of some critical parameters is also visible through visual inspection. For a simple example, alpha (proposal acceptance rate) is expected and observed to affect the density of the network. Similarly, the lower lambda (continuation rate) levels give a less dense network picture.

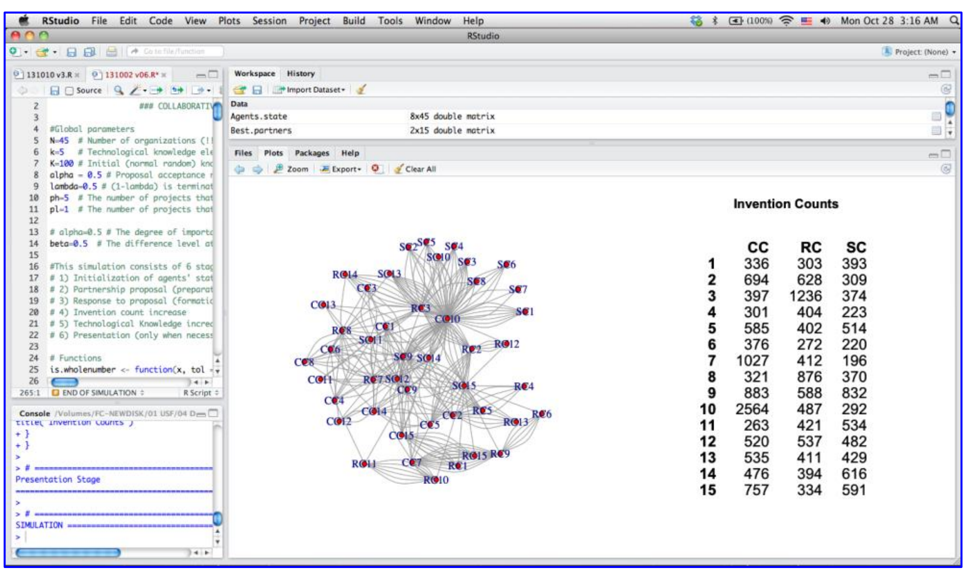


Figure 2.2. A screen-shot of the network map and invention counts of each member

---

[8] A version of the code is provided in the Appendix A. The soft copy can also be obtained from the author on request.

Finally, invention counts are also inspected for balanced contributions of three possible method of increase. For example, if the invention count increases due to internal knowledge stocks becomes overwhelmingly larger than the increase due to knowledge complementariness, then we possibly miss the observed impact of the β parameter (that is, the spectrum of invention types identified by knowledge distances). We would largely observe the impact of internal knowledge differences.

Similarly, if the invention count increase due to (i), utilization of external connections, becomes overwhelmingly larger than the increase due to the (ii), internal knowledge stocks, then we possibly miss the observed impact of internal knowledge levels (or indirectly, knowledge increases due to the complementariness effect). We would largely observe the impact of network connectivity. The cases are so set that all cases provide balanced amounts of contributions to invention count increase.

2.4. Expected Parameter Effects

2.4.1. The Summary of the Model Parameters

Performance of the strategy groups is expected to be affected by the technology parameters α (dynamism in the industry) and β (the type of invention). However, there are other important parameters of the model. The simulation will be run over different values of an important few of them:

- N: The number of network members
- k : The number of technological knowledge domains
- cyc : The number of cycles to be run

The strategy groups in all experiments above are set to constitute $1/3^{rd}$ of the networks' members. As the final experiment, the combination of the number of members in each strategy groups is changed.

2.4.2. Expected Effects

The technological dynamism parameter, α, is directly related to the connection density because it is formulated as the acceptance rate for proposed projects. At the lowest extreme ($\alpha = 0$), no member will be able to establish future connections (except for the ones randomly provided in the initialization). In this case, the only increase can be gained through the internal knowledge levels. Unless there are connections, internal knowledge levels are only changed by uncontrolled (random) increases or decreases. Therefore, in the long run, $\alpha = 0$ must provide a completely random picture of invention performances.

Increase in α is expected to benefit all strategy groups but presumably in different ways. CCs will always look for the members who have best knowledge levels. As α is increased, the higher number of connections provides higher amount of increases in CCs' knowledge levels. The higher internal knowledge levels lead to higher invention counts. Therefore, an increase in α is expected to have an indirect effect on the invention performance of CCs. RCs always propose partnerships to their current partners. As the new proposals are accepted, the current number of alliances is incremented by the number of new alliances. As α is increased, the higher number of connections accumulates, that is expected to bring invention counts from the category (i), utilization of external connections. SCs always follow those members who made the highest invention counts. The advantage of SCs is that, because many SCs will offer/get accepted by the member of the highest invention count, there exists an accumulation of members. The SCs will benefit from the number of connections (in practice, shared resources) this accumulation

provides. Note that although there is an increase in the formula of CR_i (increase due to the external resources) the hypothesized increase in the invention counts is not guaranteed. The members' knowledge levels and knowledge complemetariness will probably be different at each cycle and at each run of α. Therefore, maybe not for every member but as a general trend, increase in α (technological dynamism) is expected to bring more invention counts for each strategy group.

H1a: Invention counts in CCs are proportional to α level.

H1b: Invention counts in RCs are proportional to α level.

H1c: Invention counts in SCs are proportional to α level.

The impact of α on the invention counts of each strategy group can be observed, however, if the impact on each group is at similar levels, the success rates of the groups against each other may not be observable. At this point, it may be critical to clarify what is meant by "success rate". In this study, especially for the following hypotheses, the success rate is defined as obtaining the higher rates of invention counts against the other groups. The focus of the study is to give a comparison of strategies rather than solely observing the parameter effects. If α is observed to impact each group at similar levels, the success rates of the strategy groups are not expected to change. In such a case, no group gains advantage over another by changing values of α.

The parameter β is expected to provide advantage over the (high or low) knowledge level differences. This should noticeably impact the performance of CCs. Because CCs make alliance proposals purely based on their lowest knowledge (to match with the highest in the sector) β is expected to give a competitive advantage to CCs. Because other groups lack this strategy, an increase in β is expected to impact the success rate of CCs against RCs and SCs.

H2a: As β increases the success rate of CCs increases.

Since RCs always keep the partnership with the same members, the differences in between the partner's technological levels are reduced due to longer terms of partnership. Therefore, it is expected that lower levels of β would benefit the RCs strategy group. The interpretation is that as RCs and their partners (they don't have to be RCs) work closely over so many cycles, they can understand each other very well. These partners expectedly attain higher success levels for inventions with less decomposable invention tasks. Such tasks are represented by small β and small differences in between the technological levels of partners are rewarded through the invention count increase due to knowledge complementariness ($CK_i = CK_j$) formula. Therefore,

H2b: As β decreases the success rate of RCs increases.

SCs follow the members that make the highest invention counts. Those who make highest invention counts are probably from among the members with higher internal knowledge levels. Assuming the distribution of the knowledge levels is uniform in SC members, changes in the β are not expected to bring more advantage or disadvantage to SCs.

H2c: The β levels do not affect the success rate of SCs.

The hypothesized affects are expected to intensify or weaken by some other parameters. Consider the number of members in the network. If it is small, some potentially extreme points in the knowledge levels or excessive connectivity may prevent the observation of the suggested effects. Higher number of members in the network, however, can be expected to intensify the abovementioned effects. Therefore, it is hypothesized that N has a positive effect on the phenomena of H2a and H2b.

H3a: Increase in N intensifies the impact of β for the success rate of CCs.

H3b: Increase in N intensifies the impact of low-β for the success rate of RCs.

Like the number of members, the number of cycles is also expected to diminish the potential extremity effects. Every other parameter held constant, the higher number of cycles can give an idea of the long-term effect of the other parameters on the invention counts. Like in N, cyc is also expected to have a positive impact on the phenomena of H2a and H2b.

H4a: Increase in cyc intensifies the impact of β for the success rate of CCs.

H4b: Increase in cyc intensifies the impact of low-β for the success rate of RCs.

Like in a real practice, the number of technological knowledge domains, k, must increase the invention counts in the model. As CCs primarily decide on the knowledge levels, it can be expected that any change in k mostly impacts CCs. Like all others, since CCs send proposals to the most desirable and the moderately desirable members only, the impact of k must be felt at the peak when k is at the smallest level (k=2). So any increase in k is expected to decrease the impact of this cognitive match advantage and result in the reduction of the success rate of CCs. No increase or decrease is suggested for the success rate of RCs and SCs due to a change in the level of k.

H5: Increase in k has a negative effect on the success rate of CCs.

The members in the strategy groups are initially designed to be equal in number (N/3, N/3 and N/3). As the final analysis, the impact of the change in their numbers is to be observed. For example, instead of a balanced number (N/3) of CCs, RCs and SCs in the network, CCs can be set to have the majority, say 2N/3, then RCs or SCs are left at N/6 each. The same unbalance can be applied in turns. In one of the extreme points, say, when all members are CCs, all members will have the same advantages and disadvantages. No significant differences must be expected between the success rates of members. A similar effect can be stated for all RCs and

SCs. In case of an increase in the members of one group, the case of the two groups with a smaller number of members must be considered for they may either gain benefit out of being in small numbers or disadvantaged due to that. However, there seems to be no reason to affect the performance in case of a change in the combination of the number of members. Therefore, changing the combination of the number of strategy members is not expected to change the success rates of any strategy groups.

H6: Unbalanced number of members does not affect the suggested β impact on CCs, RCs and SCs.

The phenomena in the twelve hypotheses (listed as *H1a, H1b, H1c, H2a, H2b, H2c, H3a, H3b, H4a, H4b, H5,* and *H6*) are analyzed in the results section. Table 2.1 provides a summary of which parameter is related to the phenomena represented by hypotheses.

Table 2.1. The summary of parameters and related hypotheses

Parameter	**Related hypotheses**
α	H1a, H1b, H1c
β	H2a, H2b, H2c
N	H3a, H3b
cyc	H4a, H4b
k	H5
Unbalanced number of members	H6

2.5. Results

In the sections that follow, one can observe how the invention performances of the strategy groups change with the expected impact of the parameters. Firstly, a visual impression of the observable effects of some parameters is provided in Figure 2.3. The network is run with N=45 members. Each node shows the name of its strategy group and its identification number.

The changes in density (the ratio of the number of edges and the number of possible edges) must give an idea for the effect of α. Dense connectivity (α = 0.8) represent technologically dynamic industries. As the number of cycles increased, not only does the shape of the network change but also members are repositioned. Some members occupy very central positions over time.

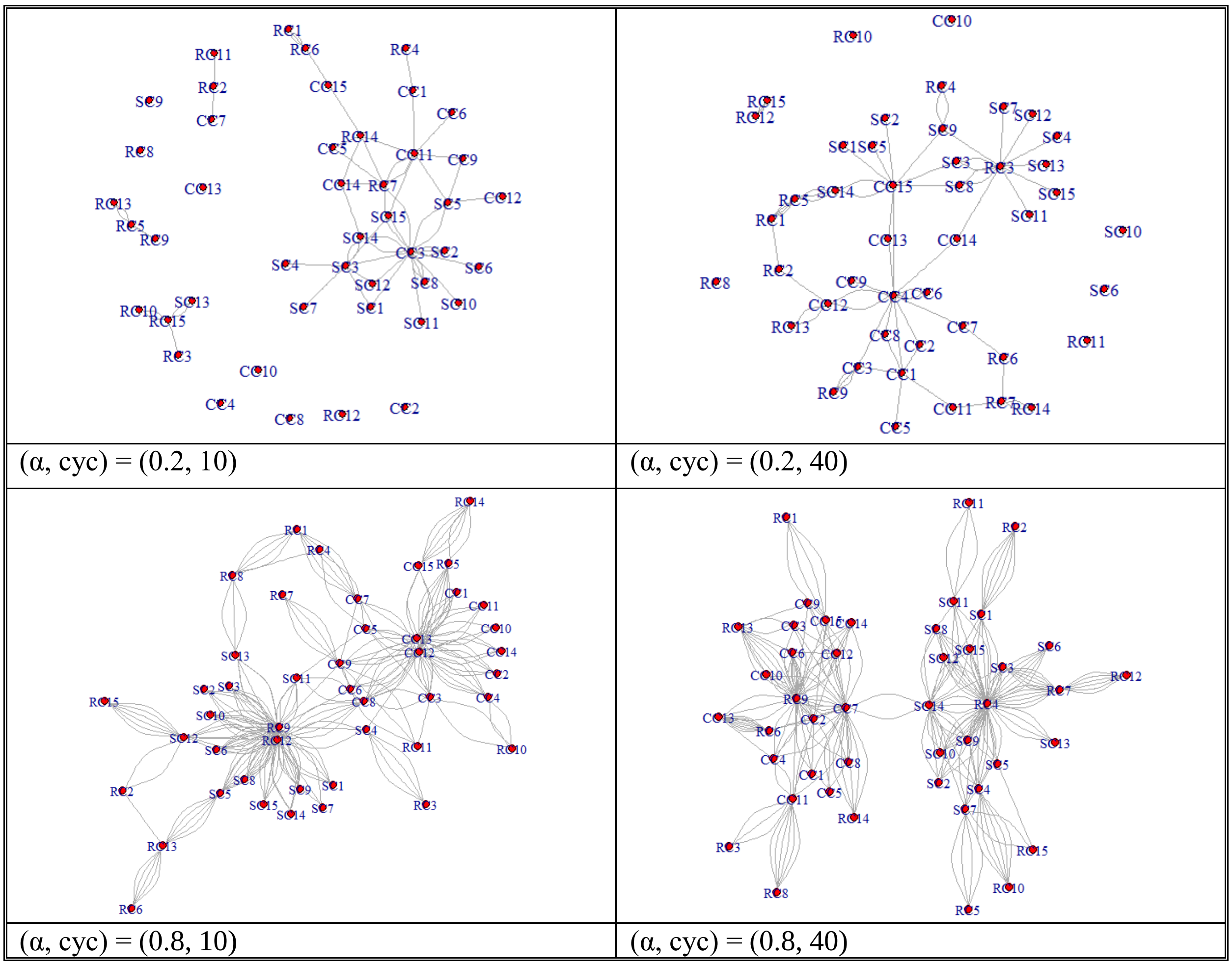


Figure 2.3. The effects of α and cyc on the network density and the connection structure.

Note that the groups (CCs, RCs and SCs) are more separated from each other over 40 cycles. The reason of separation may be explained differently for each group. Because CCs only go after the highest k_t's, they probably find it at another CC in the long run because knowledge

increases due to the partnership provides knowledge advantage to CCs at every cycle. They do not need RCs and SCs anymore unless they somehow have high k_t values. RCs are not separated completely because they value their historical connections and continue connected to their CC, RC and SC partners.

SCs usually create an accumulation on one or a few members. In fact, when they are connected to a historically successful member, they make it more successful due to the number of connections (resources) they provide, like they did it to RC4 and SC14 in the right bottom picture of the Figure 2.3. In fact, the invention count performances (accumulated over 40 cycles) clearly indicated that RC4 and SC14 are leading in the invention counts (not shown here).

2.5.1. The Effect of α on the Strategy Groups

It was claimed that the invention counts are proportionate to α levels (*H1a, H1b* and *H1c*). A quick experiment shows the comparison of the performance levels of members in each group against α levels. Keeping everything else constant[9], α is set to the values of 0, 0.2, 0.4, 0.6, 0.8 and 1. The corresponding impact can be observed in Figure 2.4. In general, the invention counts increase with an increase in *α* for all strategy groups. The correlation values for each member are provided in Table 2.2. Note that the magnitude of the invention counts must be taken as comparative values rather than exact representation of a real practice. Invention counts in an empirical observation depend on several factors, like size of the organization, invention policy etc., which I do not consider in this research. Similar increases can be observed in invention counts with an increase in *α*, for all strategy groups at various parameter values of *N, k* and *cyc*. Therefore, I find support for my hypotheses *H1a, H1b* and *H1c*.

[9] *N*: Number of organizations, *k*: number of technological domains and *cyc*: number of cycles are set to (12, 2, 10)

3.5.2. Performance Plots

Before evaluation of the next set of hypotheses, it is useful to view some more visual impressions of the simulation. Figure 2.5 shows an average performance plot of the three groups against the counting cycles. The experiment in Figure 2.5 is run with N=12 where there are 4 members in each group. The lines represent the strategy groups and each data point is the average invention count of the 4 members. Because the invention counts is a cumulative measure, they are non-decreasing plots.

The performance at the end of 40^{th} cycle is the average of the total counts for each group over 40 cycles. Figure 2.5 is a typical picture of one experiment. Mean performances do not exhibit perfectly linear increases and the differences between groups may increase or decrease even after several cycles. For this reason, the experiments are applied in *cyc*=10, 20 and 40 cycles to observe changes and test the hypotheses *H4a and H4b*.

Figure 2.6 is the representation of the same plot with the results of 10 experiments (replications) shown on the right hand side. Note that the current plot only shows the last experiment's result where CC has been the winner at average invention performance counts. The plots related to the initial 9 experiments are not shown here.

For the results in Figure 2.6, the experiments are run in the following settings: N=12, k=2, *cyc*=10, α=0.0 and β=0.5.

2.5.3. Performance Table of Experiments

As discussed in the model specification, the primary representation of the results will be for the changing values of α (technological dynamism) and β (invention type). α and β are both set to assume the values (0, 0.2, 0.4, 0.6, 0.8, 1.0) and 10 replications of the experiment is run for each pair of setting, like (α, β) = (0, 0) ; (α, β) = (0, 0.2), and so on.

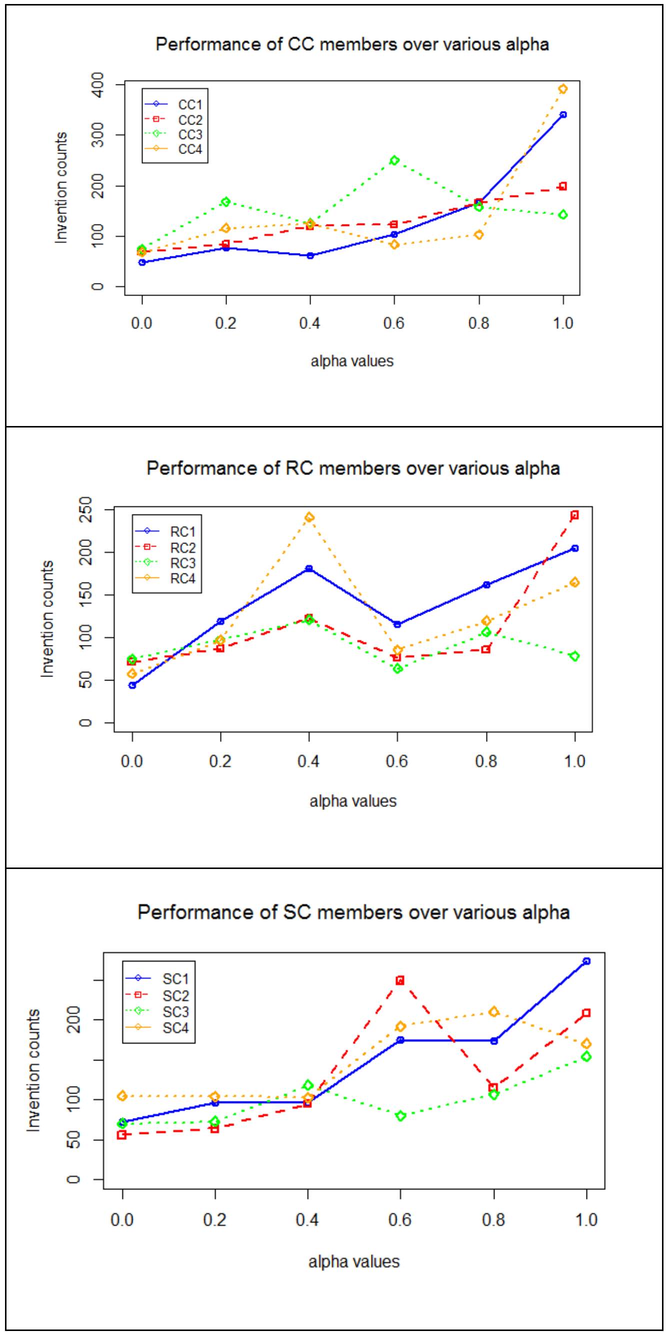


Figure 2.4. Invention performances at α levels

Table 2.2. Pearson-correlation values for α against invention counts

Members	**CC1**	**CC2**	**CC3**	**CC4**	**RC1**	**RC2**	**RC3**	**RC4**	**SC1**	**SC2**	**SC3**	**SC4**
Correlation values	0.86	0.98	0.40	0.68	0.81	0.66	-0.02	0.36	0.94	0.72	0.79	0.80

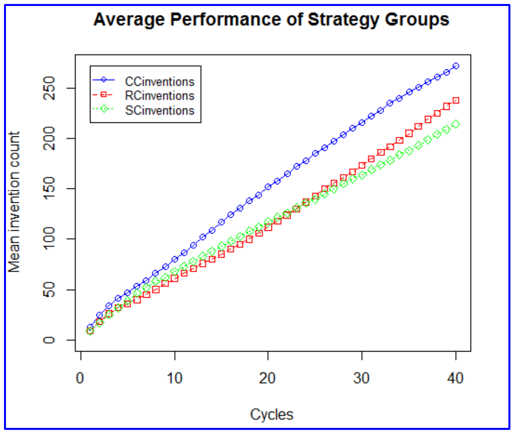


Figure 2.5. Average performance plot for one experiment

In order to present the results at different pairs of *(α, β)*, a 6x6 table is constructed with each cell showing the results of 10 experiment runs.

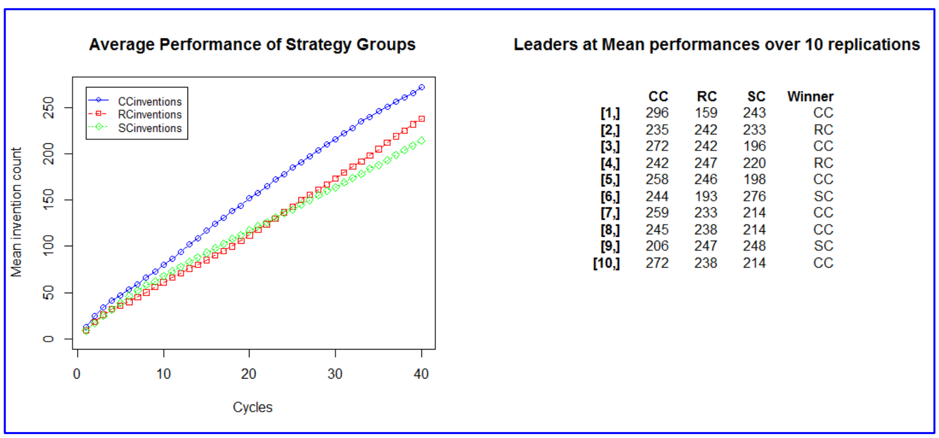


Figure 2.6. The results of 10 replications

Table 2.3 is the representation of such a table where other parameter values are at $N=24$, $k=5$ and $cyc=20$. The α levels are increased from left to right and the β values are increased from bottom to top. In each cell, the numbers across strategy groups indicate the number of "wins", that is the number of times that the average invention counts of the members of the strategy groups become the highest among 10 replications of the same experiment. The outcomes with a

member winning more than 5 experiments are colored and bolded. This is because the outcomes that groups receive up to 5 wins can be considered out of chance, however, the probability of getting outcome sets above 5 wins by chance quickly diminishes towards 10 wins. For example, the probability of obtaining the outcome of CC=7, RC=3, SC=0 wins by chance, P(7, 3, 0), is 0.002. The probability of obtaining the family of the outcome set for 7, 3, 0 values, that is P(7, 3, 0)+ P(7, 0, 3)+ P(3, 7, 0)+P(3, 0, 7)+P(0, 7, 3)+P(0, 3, 7), is 0.012. The multinomial probabilities of 14 outcome sets for three outcomes, and p=1/3 for each, are provided in Table 2.4.

Table 2.3. An example table for performances over various *(α, β)* values

β =	**1.0**	CC	2	CC	5	CC	4	**CC**	**6**	**CC**	**7**	**CC**	**7**
		RC	4	RC	1	RC	4	**RC**	**3**	**RC**	**0**	**RC**	**3**
		SC	4	SC	4	SC	2	**SC**	**1**	**SC**	**3**	**SC**	**0**
	0.8	CC	5	CC	3	**CC**	**6**	CC	4	CC	5	**CC**	**6**
		RC	3	RC	3	**RC**	**3**	RC	3	RC	4	**RC**	**3**
		SC	2	SC	4	**SC**	**1**	SC	3	SC	1	**SC**	**1**
	0.6	CC	3	**CC**	**6**	CC	3	CC	3	**CC**	**7**	CC	3
		RC	4	**RC**	**3**	RC	3	RC	4	**RC**	**3**	RC	3
		SC	3	**SC**	**1**	SC	4	SC	3	**SC**	**0**	SC	4
	0.4	CC	5	CC	3	CC	4	**CC**	**10**	CC	4	**CC**	**3**
		RC	1	RC	3	RC	2	**RC**	**0**	RC	4	**RC**	**6**
		SC	4	SC	4	SC	4	**SC**	**0**	SC	2	**SC**	**1**
	0.2	**CC**	**6**	CC	4	CC	5	CC	5	CC	5	CC	4
		RC	**2**	RC	4	RC	3	RC	2	RC	4	RC	4
		SC	**2**	SC	2	SC	2	SC	3	SC	1	SC	2
	0.0	CC	2	CC	4	**CC**	**7**	CC	3	CC	5	CC	5
		RC	3	RC	4	**RC**	**3**	RC	4	RC	2	RC	2
		SC	5	SC	2	**SC**	**0**	SC	3	SC	3	SC	3
α =		**0.0**		**2.0**		**4.0**		**6.0**		**8.0**		**1.0**	

Note that the sum of the probabilities of obtaining outcomes by chance up to 5 wins is 0.77. The sum of the probabilities of obtaining outcomes with 6 by chance is 0.17 and with 7 by chance is only 0.05. When you consider the probability of 6 or 7 wins of only one specific group,

CC, RC or SC, by chance becomes one third of those values. In other words, the probability of obtaining 6 wins by chance, for CCs for example, becomes 0.06 (and 0.017 for 7 wins).

In summary, those cells with results containing 6 or more wins are considered to require attention, especially, when a couple of them are observed side-by-side in the table of performances.

Table 2.4. The multinomial probabilities of outcome sets

Family of outcomes*			**Multinomial Coefficient**	**Probability of obtaining the outcome by chance**	**Number of combinations**	**The probability of obtaining the outcome family**
4	3	3	4200	0.071	3	0.213
4	4	2	3150	0.053	3	0.160
5	3	2	2520	0.043	6	0.256
5	4	1	1260	0.021	6	0.128
5	5	0	252	0.004	3	0.013
6	3	1	840	0.014	6	0.085
6	2	2	1260	0.021	3	0.064
6	4	0	210	0.004	6	0.021
7	2	1	360	0.006	6	0.037
7	3	0	120	0.002	6	0.012
8	1	1	90	0.002	3	0.005
8	2	0	45	0.001	6	0.005
9	1	0	10	0.000	6	0.001
10	0	0	1	0.000	3	0.000

*In the family of, say (5, 4, 1), there are 6 different combinations of outcomes. Therefore, the probability of obtaining exactly (5, 4, 1), is multiplied by 6 in order to find the probability of the family of outcomes.

2.5.4. The Parameter Space of N, k and cyc

The performance results in Table 2.2 are obtained at parameter values N=24, k=5, cyc=20. The series of hypotheses, *H2a*, *H2b*, *H2c*, *H3a*, *H3b*, *H4a*, *H4b*, and *H5*, are related with the success rates of the strategies against each other with changing parameters. In order to proceed with the observations related to these hypotheses, each parameter value is extended to a

smaller and a higher point. Therefore, the experiments are designed to run at N=12, 24 and 45; k=2, 5 and 20; cyc=10, 20 and 40. This creates an experiment space of 27 combinations. One can, then, report whether the phenomena explained in hypotheses are observed towards increasing or decreasing values of N, k and cyc.

2.5.5. The Effect of β on the Strategy Groups

The hypotheses *H2a*, *H2b* and *H2c* are related to the impact of β on the success rates of strategies against each other. A visual observation helps to see the support for each hypotheses at the performance table at $N=45$, $k=2$ and $cyc=10$ (Table 2.5). In order to understand the impact better, correlation values are provided at Table 2.6.

The positive correlation values for CCs indicate that an increase in β level increases the success rate of CCs against the other strategies. The negative correlation values for RCs provide support for *H2b* that claimed lower β gives higher RC success rates against others. SCs success rates do not exhibit a consistent increase or a decrease. Therefore, *H2c* cannot be rejected. Similar results are obtained in the majority of performance tables created at various parameter values of N, k and cyc. In some performance tables, especially in low N and low cyc, β impact is not observed, which will be presented in the next section.

Note that the correlation values at $\alpha = 0.0$ is not provided. The success rates at $\alpha = 0.0$ must not be used because it is the level when no venture proposals are accepted. Because no additional connection can be done, the success rates are partly the results of initial (random) endowments of the connections and mostly due to the randomly increased/decreased knowledge levels (increase/decrease because of the unknown factors, $KU^{+ \text{ or } -}$).

One advantage it provides, though, is that one can observe the complete randomness in the first column of the performance tables as compared to the various phenomena observed in the

next five columns. It serves like a control group to other columns but must be disregarded in the analyses.

Table 2.5. The β impact on success rates of CCs and RCs

β =	**1.0**	CC	4	**CC**	**6**	**CC**	**6**	**CC**	**6**	**CC**	**8**	CC	3
		RC	5	**RC**	**2**	**RC**	**2**	**RC**	**2**	**RC**	**1**	RC	4
		SC	1	**SC**	**2**	**SC**	**2**	**SC**	**2**	**SC**	**1**	SC	3
	0.8	**CC**	**3**	**CC**	**6**	CC	4	CC	3	**CC**	**6**	**CC**	**6**
		RC	**6**	**RC**	**1**	RC	3	RC	5	**RC**	**3**	**RC**	**1**
		SC	**1**	**SC**	**3**	SC	3	SC	2	**SC**	**1**	**SC**	**3**
	0.6	**CC**	**2**	CC	3	**CC**	**6**	CC	5	CC	5	**CC**	**6**
		RC	**1**	RC	2	**RC**	**3**	RC	4	RC	1	**RC**	**1**
		SC	**7**	SC	5	**SC**	**1**	SC	1	SC	4	**SC**	**3**
	0.4	CC	3	**CC**	**7**	CC	5	**CC**	**6**	CC	5	CC	4
		RC	3	**RC**	**3**	RC	3	**RC**	**0**	RC	1	RC	2
		SC	4	**SC**	**0**	SC	2	**SC**	**4**	SC	4	SC	4
	0.2	CC	3	**CC**	**2**	CC	5	**CC**	**7**	CC	4	CC	5
		RC	2	**RC**	**7**	RC	3	**RC**	**1**	RC	3	RC	4
		SC	5	**SC**	**1**	SC	2	**SC**	**2**	SC	3	SC	1
	0.0	**CC**	**1**	CC	3	CC	4	**CC**	**1**	CC	4	**CC**	**6**
		RC	**7**	RC	4	RC	3	**RC**	**6**	RC	2	**RC**	**2**
		SC	**2**	SC	3	SC	3	**SC**	**3**	SC	4	**SC**	**2**
α =		**0.0**		**2.0**		**4.0**		**6.0**		**8.0**		**1.0**	

Table 2.6. Pearson-correlation values for β against the success rates

Correlation values at	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
CCs	0.59	0.48	0.28	0.92	-0.42
RCs	-0.73	-0.65	-0.09	-0.27	0.00
SCs	0.18	-0.21	-0.41	-0.76	0.52

2.5.6. The Effect of the Size of the Network, *N*

The hypotheses *H3a* and *H3b* are related to the *N*'s intensification effect for the impact of β on CCs and RCs. The performance table, at *N=12, k=2* and *cyc=40*, provides support for *H3a* and *H3b* (Table 2.7). This example, however, must be observed in a reverse manner. In fact, the intensification can be observe at Table 2.5, whereas Table 2.7 shows a diminishing effect of lower *N* for the impact of β. Because N=12 in this case, it can be observed that correlation values for CCs are lowered (Table 2.8).

As mentioned in the previous section, the performance table at *N=12, k=2* and *cyc=40* (Table 2.7) is not an example of support for the negative the impact of β on RCs (H2b). However, as already shown in Table 2.5, higher *N* intensified the impact of β on RCs, and made it observable. Therefore Table 2.7 does not provide support for H2b or H3b but the results in Table 2.5 of the previous section actually provides support for both H2b and H3b.

2.5.7. The Effect of the Number of Cycles, *cyc*

The hypotheses *H4a* and *H4b* are related to the intensification effect of *cyc* for the β impact on CCs and RCs' success rates. Table 2.5 provided the performance table at *N=45, k=2* and *cyc=10*. Now, we can make a comparison if we look at the performance table at *N=45, k=2* and *cyc=40,* which is provided in Table 2.9.

The correlation values at Table 2.10 clearly show that the β impact on CCs' success rates intensified as compared to a lower *cyc* experiment (given in Table 2.5). However it does not seem to make the same impact on the RCs' success rates. Therefore, in the performance table at *N=45, k=2* and *cyc=40,* I find support for *H4a* but not for *H4b.*

Table 2.7. The diminishing β impact towards lower N

β =	**1.0**	CC	4	**CC**	**6**	**CC**	**8**	**CC**	**9**	**CC**	**6**	**CC**	**6**
		RC	3	**RC**	**2**	**RC**	**0**	**RC**	**1**	**RC**	**2**	**RC**	**4**
		SC	3	**SC**	**2**	**SC**	**2**	**SC**	**0**	**SC**	**2**	**SC**	**0**
	0.8	CC	5	**CC**	**8**	CC	4	**CC**	**6**	CC	5	**CC**	**2**
		RC	3	**RC**	**1**	RC	3	**RC**	**3**	RC	5	**RC**	**6**
		SC	2	**SC**	**1**	SC	3	**SC**	**1**	SC	0	**SC**	**2**
	0.6	CC	2	CC	5	**CC**	**8**	**CC**	**8**	**CC**	**7**	CC	4
		RC	3	RC	2	**RC**	**2**	**RC**	**1**	**RC**	**2**	RC	2
		SC	5	SC	3	**SC**	**0**	**SC**	**1**	**SC**	**1**	SC	4
	0.4	CC	2	**CC**	**6**	CC	5	**CC**	**7**	**CC**	**7**	CC	5
		RC	4	**RC**	**1**	RC	3	**RC**	**2**	**RC**	**2**	RC	4
		SC	4	**SC**	**3**	SC	2	**SC**	**1**	**SC**	**1**	SC	1
	0.2	**CC**	**7**	**CC**	**7**	CC	5	**CC**	**7**	**CC**	**6**	**CC**	**6**
		RC	**2**	**RC**	**3**	RC	3	**RC**	**2**	**RC**	**4**	**RC**	**0**
		SC	**1**	**SC**	**0**	SC	2	**SC**	**1**	**SC**	**0**	**SC**	**4**
	0.0	CC	3	**CC**	**7**	**CC**	**7**	**CC**	**6**	CC	5	CC	3
		RC	2	**RC**	**1**	**RC**	**1**	**RC**	**1**	RC	3	RC	4
		SC	5	**SC**	**2**	**SC**	**2**	**SC**	**3**	SC	2	SC	3
α =		**0.0**		**2.0**		**4.0**		**6.0**		**8.0**		**1.0**	

Table 2.8. Pearson-correlation values for β against the success rates: The effect of N

Correlation values at	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
CCs	-0.15	0.16	0.59	0.12	0.07
RCs	0.00	-0.25	0.13	-0.08	0.41
SCs	0.14	0.05	-0.82	0.00	-0.59

2.5.8. The Effect of the Technological Knowledge Domains, *k*

H5 is related to the negative effect of k for the β impact on CCs' success rates. Table 2.9 provided the performance table at N=45, k=2 and cyc=40. Now, we can make a comparison if we look at the performance table at N=45, k=20 and cyc=40, which is provided in Table 2.11.

Table 2.9. The intensifying β impact is towards higher *cyc*

β =	**1.0**	CC	1	**CC**	**10**	**CC**	**10**	**CC**	**9**	**CC**	**9**	**CC**	**9**
		RC	5	**RC**	**0**	**RC**	**0**	**RC**	**1**	**RC**	**1**	**RC**	**1**
		SC	4	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**
	0.8	**CC**	**6**	**CC**	**9**	**CC**	**10**	**CC**	**9**	**CC**	**10**	**CC**	**10**
		RC	**3**	**RC**	**1**	**RC**	**0**	**RC**	**1**	**RC**	**0**	**RC**	**0**
		SC	**1**	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**
	0.6	CC	1	**CC**	**9**	**CC**	**8**	**CC**	**9**	**CC**	**10**	**CC**	**8**
		RC	4	**RC**	**0**	**RC**	**1**	**RC**	**1**	**RC**	**0**	**RC**	**1**
		SC	5	**SC**	**1**	**SC**	**1**	**SC**	**0**	**SC**	**0**	**SC**	**1**
	0.4	CC	4	**CC**	**10**	**CC**	**8**	**CC**	**10**	**CC**	**8**	**CC**	**10**
		RC	3	**RC**	**0**	**RC**	**2**	**RC**	**0**	**RC**	**2**	**RC**	**0**
		SC	3	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**	**SC**	**0**
	0.2	CC	2	**CC**	**8**	**CC**	**7**	**CC**	**6**	**CC**	**7**	**CC**	**9**
		RC	4	**RC**	**1**	**RC**	**3**	**RC**	**3**	**RC**	**1**	**RC**	**1**
		SC	4	**SC**	**1**	**SC**	**0**	**SC**	**1**	**SC**	**2**	**SC**	**0**
	0.0	**CC**	**2**	**CC**	**8**	**CC**	**8**	**CC**	**8**	**CC**	**7**	**CC**	**8**
		RC	**6**	**RC**	**1**	**RC**	**1**	**RC**	**2**	**RC**	**1**	**RC**	**0**
		SC	**2**	**SC**	**1**	**SC**	**1**	**SC**	**0**	**SC**	**2**	**SC**	**2**
α =		**0.0**		**2.0**		**4.0**		**6.0**		**8.0**		**1.0**	

Table 2.10. Pearson-correlation values for β against the success rates: The effect of *cyc*

Correlation values at	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
CCs	0.72	0.83	0.50	0.81	0.36
RCs	-0.49	-0.69	-0.52	-0.36	0.29
SCs	-0.68	-0.41	-0.39	-0.83	-0.57

The correlation values at Table 2.12 clearly show that the β impact on CCs' success rates weakened as compared to a lower k experiment (given in Table 2.10). It gives a strong support to H5, which claims that as k increases the β impact on CCs' success rates diminish.

Table 2.11. The diminishing β impact towards higher k

$\beta =$						
1.0	CC 3 RC 4 SC 3	**CC 8** **RC 1** **SC 1**	CC 5 RC 2 SC 3	CC 4 RC 3 SC 3	**CC 3** **RC 6** **SC 1**	CC 3 RC 5 SC 2
0.8	**CC 3** **RC 6** **SC 1**	CC 5 RC 4 SC 1	CC 4 RC 4 SC 2	CC 3 RC 5 SC 2	**CC 6** **RC 1** **SC 3**	CC 4 RC 3 SC 3
0.6	CC 2 RC 3 SC 5	**CC 6** **RC 3** **SC 1**	CC 3 RC 4 SC 3	CC 5 RC 4 SC 1	CC 5 RC 0 SC 5	CC 3 RC 5 SC 2
0.4	CC 2 RC 3 SC 5	CC 3 RC 5 SC 2	**CC 6** **RC 2** **SC 2**	CC 2 RC 4 SC 4	CC 4 RC 3 SC 3	**CC 7** **RC 2** **SC 1**
0.2	CC 3 RC 2 SC 5	CC 5 RC 3 SC 2	**CC 7** **RC 2** **SC 1**	CC 1 RC 5 SC 4	CC 2 RC 5 SC 3	**CC 1** **RC 7** **SC 2**
0.0	CC 3 RC 3 SC 4	CC 5 RC 3 SC 2	CC 4 RC 3 SC 3	**CC 3** **RC 1** **SC 6**	CC 3 RC 2 SC 5	CC 3 RC 2 SC 5
$\alpha =$	**0.0**	**2.0**	**4.0**	**6.0**	**8.0**	**1.0**

Table 2.12. Pearson-correlation values for β against the success rates: The effect of k

Correlation values at	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
CCs	0.59	-0.25	0.53	0.47	0.14
RCs	-0.36	0.16	0.36	0.12	0.16
SCs	-0.88	0.26	-0.73	-0.64	-0.43

Note that the performance table at $N=45$, $k=20$ and $cyc=40$ is another example where β impact on RCs' success rates cannot find support. It is true for all performance tables at various combinations of N, k and cyc that they present support for some hypothesis but not for some others. The next section provides the major picture for the supported hypotheses at each combination of N, k and cyc.

2.5.9. The Main Table: A Summary for the Representation of the Effects

As mentioned earlier, for all 27 combinations of *N*, *k* and *cyc* values at (12, 24, 45); (2, 5, 20) and (10, 20, 40) consecutively, experiments are run and the performance tables are created. Not all performance tables are shown here, but the observed phenomena related to hypotheses H2a, H2b, H2c can be listed for all combinations. The Main Table, Table 2.13, lists the supported hypotheses for combinations of *N*, *k* and *cyc* values. The table also gives us the chance to see whether H3a, H3b, H4a, H4b and H5 are supported by observation through higher *N*, *k* and *cyc* values.

Table 2.13. The main table: Supported hypotheses for *N*, *k* and *cyc* values

		Number of members in the invention network		
Number of cycles	Number of technological domains	**N=12**	**N=24**	**N=45**
cyc=10	**k=2**	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=5	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=20	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
cyc=20	**k=2**	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=5	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=20	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
cyc=40	**k=2**	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=5	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c
	k=20	H2a, H2b, H2c	H2a, H2b, H2c	H2a, H2b, H2c

To ensure the clarity of the results, the following representation method is applied. When these phenomena related to the hypotheses are strongly observed, the name of the hypotheses are indicated with normal text (100% black font color) in the corresponding cell. When very weakly observed or not observed at all, they are written with light grey text (25% darkness), which

means the phenomena is barely there or not there. For example, if there is a cell with no normal text (all grey), it means no proof for hypotheses are observed and the average invention count performance wins for given combination of the parameters, *N*, *k* and *cyc*, are similar to a complete random multinomial experiment.

H2a and H2c find support from the majority of the invention count performance tables. Although H2b does not from support the majority of the tables, it finds support from many of them and there were no tables that disprove the phenomenon (that an increase in β to lead higher success rates in RC).

I opt to not represent the group of hypotheses H3a, H3b, H4a and H4b on the Main Table. Based on the trends towards higher *N*, *k* and *cyc* on the table, it can be observed whether they are supported or not. For example, for H3a and H3b, we can observe whether there is an increasing evidence in support of H2a and H2b for increasing *N* (from the left to the right hand side of the Table 2.13). Over the nine observations (the nine rows) for the intensity of H2a towards the right hand side of the table, one cannot see a clear support for H3a. Similarly, H3b cannot find support from many lines of the table.

H4a and H4b are related to *cyc*, so, one has to check each column to see if there is an increasing evidence in support of H2a for increasing *cyc*. The nine observations here must be by checking each H2a at every three rows. In fact, not many columns support the increase the in visibility of H2a. Therefore, H4a is not supported. Similarly, H3b cannot find support from many columns of the table.

H5 is related to *k*, so, one has to check each column to see if there is a trend for more H2a for <u>decreasing</u> *k*. Indeed, there is an increase in the evidence in support of H2a as *k* decreases. Therefore, I find support for H5.

2.5.10. The Effect of Unbalanced Number of Members

Three experiments were conducted to observe the effect of an unbalanced number of members. The experiments are prepared for *N=24, k=5* and *cyc=20* and in each case, one of the strategy groups had 2N/3 members and the other two had N/6 members each. The results are compared with the previously applied balanced experiments at *N=24, k=5* and *cyc=20.* The reason that the experiments are done at (24, 5, 20) is because it is the performance table where the three of the phenomena are observable explained in H2a, H2b and H2c. The purpose is to test whether the unbalanced member configuration changes the results on the observed phenomena.

Table 2.14 provides the performance table for the experiment where CCs are with 4, RCs are with 16 and SCs are with 4 members at *N=24, k=5* and *cyc=20*, which was one of the experiments designed. This table shows that the impact of β on the success rate of CCs still observable. A similar observation is possible for the experiment that was run for table for the configuration where CCs are with 4, RCs are with 4 and SCs are with 16 members. However, it was not observable at the experiment where CCs with 4, RCs with 4 and SCs with 16 members. Table 2.15, actually, provides a complete picture of correlation values for β against the success rates in all unbalanced experiments applied.

2.5.11. Success Rates of Strategy Groups

One of the primarily expected results of the developed model was the identification of the most effective collaboration strategy. This section is provided after several steps of analyses in order to give a better idea of what parameters are influencing the success rates of the strategies. It is also provided along with the consistencies in the next section to make a better evaluation. The results from the 27 performance tables indicate that cognitive cooperation is the most effective

Table 2.14. The success rates of an experiment for unbalanced members

β =						
1.0	CC 5 RC 2 SC 3	**CC 6** **RC 1** **SC 3**	CC 5 RC 1 SC 4	CC 5 RC 2 SC 3	CC 5 RC 3 SC 2	**CC 7** **RC 2** **SC 1**
0.8	**CC 6** **RC 1** **SC 3**	**CC 2** **RC 1** **SC 7**	CC 4 RC 2 SC 4	CC 5 RC 1 SC 4	**CC 6** **RC 2** **SC 2**	CC 5 RC 0 SC 5
0.6	**CC 3** **RC 1** **SC 6**	CC 5 RC 3 SC 2	**CC 6** **RC 3** **SC 1**	CC 3 RC 5 SC 2	CC 5 RC 2 SC 3	CC 5 RC 4 SC 1
0.4	CC 2 RC 4 SC 4	CC 2 RC 4 SC 4	CC 5 RC 2 SC 3	**CC 6** **RC 3** **SC 1**	**CC 2** **RC 6** **SC 2**	CC 5 RC 1 SC 4
0.2	**CC 1** **RC 3** **SC 6**	CC 4 RC 2 SC 4	CC 3 RC 5 SC 2	CC 3 RC 3 SC 4	CC 3 RC 4 SC 3	**CC 6** **RC 1** **SC 3**
0.0	CC 1 RC 4 SC 5	CC 4 RC 4 SC 2	CC 1 RC 4 SC 5	**CC 6** **RC 3** **SC 1**	CC 5 RC 2 SC 3	**CC 3** **RC 1** **SC 6**
α =	**0.0**	**2.0**	**4.0**	**6.0**	**8.0**	**1.0**

Table 2.15. Pearson-correlation values for β against the success rates: The effect of unbalance

Experiment number	Configuration of the members	Correlation values at =>	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1.0$
1	**16**	**CCs**	0.00	-0.38	0.00	-0.11	-0.25
	4	**RCs**	-0.31	0.08	-0.45	-0.12	0.40
	4	**SCs**	0.40	0.34	0.36	0.26	-0.20
2	**4**	**CCs**	0.23	0.72	-0.08	0.43	0.68
	16	**RCs**	-0.74	-0.84	-0.36	-0.17	0.19
	4	**SCs**	0.34	-0.04	0.43	-0.68	-0.57
3	**4**	**CCs**	0.69	0.07	0.76	0.66	-0.72
	4	**RCs**	-0.51	0.41	-0.31	0.00	-0.06
	16	**SCs**	0.00	-0.38	-0.85	-0.55	0.66

strategy in most areas of the parameter space of (*α, β).* In a given performance table, if a strategy group has the largest number of cells (i.e., an *αxβ* combination, except for the *α=0* column) with "6 and above" experiment wins, that strategy can be considered as a successful strategy. If the strategy group has more than half of the cells with "6 and above" wins in a table, that strategy can be considered as the predominantly successful strategy. Cognitive cooperators have been predominantly successful in 4, and successful in 18 (including the 4 predominant successes) of the 27 performance tables, especially where *k* is small. Relational cooperators are found to be successful in 3 and success-driven cooperators in only 1 (at *N=24, k=20, cyc=10*) of the tables. Neither RCs nor SCs are observed to be predominantly successful in a table.

2.5.12. Consistency: Standard Deviation of Invention Performances

The consistency of the invention performance is no less important than the magnitude of their average counts. If a strategy group is performing best on average, one must check if their performance is consistent across its members or if it only creates one or a few star performers. Along with the average performance, the standard deviations of the invention counts are obtained. Eight standard deviation performance tables are produced at the extreme values of *N, k* and *cyc (*i.e., *N=12,45; k=5,20* and *cyc=10, 40).*

Interesting results are observable in the standard deviation performance tables. The following table for the standard deviation of the performances provides an example for three of those interesting phenomena (Table 2.16). Note that the winners in this table are the ones who produce the smallest standard deviation values among the members of the strategy groups.

First of all, the CCs who exhibit high success rates against RCs and SCs in most performance tables of the previous section perform poorly in the standard deviation tables (i.e., produced high standard deviation values).

Table 2.16. Standard deviation of the performances at *N=45, k=2* and *cyc=40*

β =							
	1.0	CC 4	CC 4	**CC 2**	CC 2	CC 5	CC 4
		RC 1	RC 3	**RC 6**	RC 5	RC 2	RC 4
		SC 5	SC 3	**SC 2**	SC 3	SC 3	SC 2
	0.8	CC 1	**CC 3**	CC 2	CC 4	CC 2	CC 4
		RC 5	**RC 1**	RC 5	RC 3	RC 5	RC 1
		SC 4	**SC 6**	SC 3	SC 3	SC 3	SC 5
	0.6	CC 3	CC 1	CC 3	**CC 1**	CC 3	CC 3
		RC 4	RC 4	RC 2	**RC 3**	RC 5	RC 3
		SC 3	SC 5	SC 5	**SC 6**	SC 2	SC 4
	0.4	CC 5	**CC 4**	**CC 6**	CC 3	CC 2	CC 3
		RC 3	**RC 6**	**RC 2**	RC 3	RC 3	RC 3
		SC 2	**SC 0**	**SC 2**	SC 4	SC 5	SC 4
	0.2	CC 3	CC 1	CC 4	**CC 1**	CC 3	**CC 1**
		RC 4	RC 3	RC 3	**RC 3**	RC 2	**RC 1**
		SC 3	SC 6	SC 3	**SC 6**	SC 5	**SC 8**
	0.0	CC 3	**CC 1**	CC 2	CC 5	**CC 2**	**CC 0**
		RC 5	**RC 6**	RC 4	RC 3	**RC 1**	**RC 3**
		SC 2	**SC 3**	SC 4	SC 2	**SC 7**	**SC 7**
α =		**0.0**	**2.0**	**4.0**	**6.0**	**8.0**	**1.0**

When the actual invention counts of CCs are analyzed, the larger standard deviations are found to be mostly due to the fact that their group produces one or more leaders in counts. One can easily deduce that once a member becomes a star, it usually stays so because SC members suddenly create the richness of connectivity. This result, of course, implies that the success rates of CCs are not always consistent because they don't have a high number of wins in standard deviation performance tables. This phenomenon will be called "High SD of CCs" to indicate that CCs produced high standard deviation values. Note that "High SD of CCs" is represented by lack of CCs in a standard deviation performance table. Therefore, Table 2.16 is an example of "High SD of CCs".

Another easily observable phenomenon is the high winning rates of SCs. They do not have high success rates in the invention count performance tables but it is understood that they produce smaller standard deviations and perform more consistently across members. It is also easily observable that the higher wins for SCs (i.e., more experiments that SCs produced the smallest standard deviation results) accumulate at the right bottom of the table. In many standard deviation performance tables, higher wins for SCs seemed to be correlated with a decrease in β although it is not always the case. In some standard deviation performance tables, higher wins for SCs correlates with increasing β and sometimes does not correlate at all. These phenomena will be called as "β^+ on SCs" (higher SCs correlate with increasing β), "β on SCs" (when β effect is not certain, that is SC wins accumulate in both ends of β levels or central levels) and "β^- on SCs" (higher SCs correlates with decreasing β). The underlying reason for why β might be affecting is probably the abovementioned phenomenon related to CCs. Because all SCs produce connectivity around the most successful member, they probably create invention count results very similar among members. Note that only one of the β effects must be observable in a standard deviation performance table and Table 2.16 is an example of "β^- on SCs".

The last phenomenon in this table is that, although not very strong, some wins by RCs are observable. The interesting thing with RC wins is that they usually appear on the left hand side of the tables, which may be an indication of a lower α impact on RCs' standard deviations. Lower α impact can be explained in the following way. RCs stick on the same members and always bring proposals to them. Because increasing α brings more and more connectivity, the higher α can bring RCs more varied connectivity results because the number of connections is not bounded at the upper end. On the other hand, lower α brings less and less connectivity results and due to the fact that lesser connectivity is bounded by 0 at the lower end, lower α can bring

less variability, thus, higher number of RC wins at lower α levels. This phenomenon is called "α^- on RCs" and Table 2.16 is an example of the phenomenon.

Table 2.17 provides the summary of observed phenomena in the standard deviation performance tables run for eight different combinations of *N*, *k* and *cyc*. When these phenomena are strongly observed, they are indicated with normal text (100% black font color) in the corresponding cell. When very weakly observed or not observed at all, they are written with light grey text (25% darkness), which means the phenomena is barely there or not there. For example, if there is a cell with no normal text (all grey), it means no abovementioned phenomena are observed and the standard deviation performance wins for given combination of the parameters, *N*, *k* and *cyc*, are similar to a complete random multinomial experiment.

Table 2.17. The main table for the impact on SD performances: Observed phenomena

		Number of members in the invention network	
Number of cycles	Number of technological domains	**N=12**	**N=45**
cyc=10	**k=2**	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs
	k=20	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs
cyc=40	**k=2**	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs
	k=20	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs	High SD of CCs α^- on RCs β^+ on SCs β on SCs β^- on SCs

2.6. Conclusions and Discussion

This study develops a model for a collaborating network of organizations that are motivated to make inventions but use different alliance strategies: cognitive, relational and success-driven cooperators. Their invention performance and success rates relating to each other are analyzed over an array of parameters. One of the two critical parameters is the level of technological dynamism, which forces the organizations to enter a higher number of alliances.

The other critical parameter is related to the type of invention that determines whether the allies must be at equal levels in their technological knowledge or specialized in their own areas and have differences in between their technological knowledge levels. All other parameters are related to the system's operation, like number of members, length of the analysis period and number of technological domains.

Intuitively, the technological dynamism leads to increased numbers of invention counts through higher number of partners, learning and knowledge complementariness effect. The simulated model produced increased numbers of invention counts as expected. In addition, increases in the technological dynamism did not give a clear competitive advantage to any of the strategies. That is, when everything else is kept constant and the technological dynamism is increased from minimum to maximum levels, no strategy group gained a superior position over the others.

The type of invention, however, affected all strategy groups differently. It was initially proposed that a cognitive cooperator would perform better relative to others when the invention task is decomposable and specialization is allowed. However, most simulations run with various system parameters proved that cognitive cooperators gain competitive advantage at decomposable tasks. Because cognitive cooperators always find their partners based on the

strategy that aims to get rid of inadequacies in their technology levels, they become successful by making partnerships with others specialized in their inadequate areas. On the other hand, the relational cooperators are suggested to operate best when decomposition of the invention task is not possible. RCs base their partnership decisions purely on trust and lose advantage when specialization is required for inventions. In many simulation runs applied with different system parameters RCs gain competitive advantage when the invention has to be done collectively and does not allow specialization of the partners. Although not fully, the model gives some support for the initial proposition. Lastly, success-driven cooperators are not expected to suffer from the decomposability requirements of the invention task. Success driven cooperators always select their partners from the most successful members in invention counts. SCs can be considered as the market entrants who actually do not have experience and follow the successful examples to make partnership. Making alliances with successful partners allows them to learn, reduce risks and, in the meantime, make inventions at moderate levels. The change of invention type neither increases nor decreases the success rate of SCs against other cooperators. The simulated model strongly proved that increases or decreases in decomposability of the invention task does not affect the success rate of SCs against other cooperators.

It was initially considered that when the number of organizations in the network increased, and similarly, when the number of cycles increased, the expected effects of technological dynamism and task decomposability would be strengthened. It turned out that the results were affected neither by any change in the number of organizations nor by the number of cycles. It can be told that the model is found to be robust, at least for less than or equal to 45 members of the network and less than or equal to 40 cycles.

The model is also tested for unbalanced number of members in each strategy group in the network. It was suggested that even though one strategy group constitutes the majority of the members, it would not affect the results regarding the suggested impacts of the technological dynamism and the task decomposability. The experiments indicated partial support for this hypothesis. Although one of the three experiments did not indicate support, it is still consistent with the general results of the experiments with balanced number of members because not all 27 experiments applied with balanced number of members indicated the impact at the same levels. It can be seen that the model operates robustly with an unbalanced number of members among the strategy groups.

In conclusion, several simulations of the developed model indicated that the most effective collaboration strategy is cognitive collaboration. A purely cognitive collaborator considers the knowledge integration and learning as priority. The knowledge increase and complementariness effect puts the CCs in a competitive position. One of the limitations of the model is lack of applicable mixed strategies. It would be interesting to test the success rates of some viable mixed strategies. It would also be an interesting extension to do an empirical analysis about the various alliance strategies of organizations along with an evaluation of their success.

The consistency tests indicated some interesting results. In most average invention count performance tables, cognitive cooperators dominated in the success rates against the relational and success-driven cooperators, especially at higher number of cycles and lower number of technological knowledge domains. Success-driven cooperators on the other hand, received the lowest success rates among the three strategies. In the eight experiments, designed to determine the standard deviation values across the members of the strategy groups, SCs were the ones who

produced the lowest values in general. Although success-driven cooperation does not offer the high rates of success in terms of invention performances, it suggests consistent results among its members. Success-driven cooperation can be identified as the least risky option both because it is not affected by the type of invention and because of the consistency of the invention count performances.

Another interesting observation is regarding the technological dynamism impact on the consistency results or relational cooperators. When there is less dynamism, RCs are found to produce less standard deviation values across its members' invention count values. Although the invention counts of RCs increase as do those for CCs and SCs, the level of inconsistency also increases and apparently faster than other members.

This study has a number of limitations. Like all simulation models, this study attempts to imitate the real practice and experiment of critical parameters of industry and product types. Many factors have been taken into account but obviously many other practical factors were not. One of the major assumptions is that the members of the network are treated as though each one is equally motivated to make inventions, all the time. Practically, organizations have their unique approach towards being "inventive" and it is probably not constant over time. Another practical issue in this aspect is that, sometimes inventions produced due to alliances are not equally counted among the partners. In practice, there are "hub" organizations that make higher investments and although some peripheral allies do specific invention activities, it is the hub organization that integrates those specific inventive activities. Secondly, although important, the number of invention is probably not a major concern for organizations in practice. It is a successful new product or service based on the inventions what matters for most organizations. Furthermore, not all inventions have the same value, but in this study, they were treated as if they

are equally valuable. Thirdly, one of the model's assumptions is that members' information is perfectly available to all members of the network, which is not practical even though it is easy to access information with recently available technologies. As the last practical issue, the number of technological domains has to be kept constant for all members of the network, which is rarely true between the collaborating partners.

Besides practical issues, there are also technical inadequacies worth mentioning. A typical simulation run to produce an invention counts performance table takes from several minutes to several hours. Considering that tens of tables had to be produced for the study, the parameter values had to be constrained at certain levels. Furthermore, there are more system parameters mentioned in the study but included in the analyses. Although a dramatic impact is not expected, the impact of lambda (the continuation rate), the number of proposals to the highly desired potential partner and the number of proposals to the moderately desired potential partner is not considered.

Although each of the existing limitations represents an exciting area for future research, this study could be extended in several other ways. In the model used, the strategies are defined as the sole way of action all throughout the cycles. As mentioned above, a future study may define mixed strategies that incorporate some best parts of the existing strategies to better mimic the practice and better inform the organizations. Another useful extension can be the learning capabilities of the members regarding what strategy to use. Once they observe the system to identify what strategy fits best their needs, the members might be given the ability to switch across different strategies.

2.7. References

1. Albino, V.; Carbonara, N.; Giannoccarro, I. Innovation in industrial districts: An agent-based simulation model. Int. J. Production Economics 104:30-45; 2006.

2. Ang S. H. Competitive intensity and collaboration: Impact on firm growth across technological environments. Stategic Management Journal, (29): 1057-1075; 2008

3. Antonelli, C.; Ferraris, G. Innovation as an emerging system property: An agent based simulation model. Journal of Artificial Societies and Social Simulation 14(2):1; 2011.

4. Baum, J. A. C.; Cowan, R.; Jonard, N. Network-independent partner selection and the evolution of innovation networks. Management Science 56(11):2094-2110; 2010.

5. Block, F. Swimming Against the Current: The Rise of a Hidden Developmental State in the United States. Politics Society, 36:169; 2008.

6. Cowan, R.; Jonard, N.; Ozman, M. Knowledge dynamics in a network industry. Technological Forecasting & Social Change 71:469-484; 2004.

7. Cowan, R.; Jonard, N.; Zimmermann, J. Bilateral collaboration and the emergence of innovation networks. Management Science 53(7):1051-1067; 2007.

8. Crouch D. D. The Changing Nature Inventing: Collaborative Inventing, http://www.patentlyo.com/patent/2013/01/inventors.html; 2009. Retrieved on 06/28/2013.

9. Garcia, R. Uses of agent-based modeling in innovation/new product development research. J. Prod. Innovation Management 22:380-398; 2005.

10. Gilsing V., Noteboom B., Vanhaverbeke Wim., Duysters G., Oord A. Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. Research Policy 37; 1717–1731; 2008.

11. Guoxiang, R.; Pingnan, R. Knowledge sharing within innovation network members with netlogo simulation platform. Information Technology, Computer Engineering and Management Sciences (ICM), 2011 International Conference 2:210-213; 2011

12. Hagedoorn J. Understanding the Rationale of Strategic Technology Partnering: Interorganizational Modesof Cooperation and Sectoral Differences. Strategic Management Journal 14(5): 371-385; 1993.

13. Iandoli, L.; Marchione. E.; Ponsiglione, C.; Zollo, G. Knowledge exchange processes in industrial districts and the emergence of networks. Cejor 20:231-250; 2012.

14. Lane, J. P.; Flagg, J. L. Translating three states of knowledge-discovery, invention, and innovation. Implementation Science 5:9; 2010.

15. Lovejoy, W. S.; Sinha, A. Efficient structures for innovative social networks. Management Science 1-9; 2010

16. Miettinen, R. Theories of Invention and an Industrial Innovation. Science Studies 9(2):34-48; 1996.

17. Mowery, D. C.; Oxley, J. E. Silverman B. S., Strategic Alliances and Interfirm Knowledge Transfer. Strategic Management Journal 17:77-91; 1996.

18. Noteboom, B., Haverbeke W.V., Duysters G., Gilsing V., Oord A. Optimal cognitive distance and absorptive capacity. Research Policy 36 1016–1034; 2007.

19. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

20. Rosenkopf, L.; Schilling, M. A. Comparing alliance network structure across industries: Observations and explanations. Strat. Entrepreneurship J. 1:191-209; 2007.

21. Schilling, M. A. Phelps C. C., Interfirm collaboration networks: The Impact of large-scale network structure on firm innovation. Management Science 53(7):1113-1126; 2007.

22. Schilling, M. A. Understanding the alliance data. Strategic Management Journal 30: 233-260; 2009.

23. Segarra-Cipres M., Bou Llusar J. C., Roca-Puig V. Exploring and exploiting external knowledge: The effect and sector and firm technological intensity. Innovation: Management, Policy and Practice. 14(2) 203-217.

24. Sobrero M. Roberts E.B. The trade-off between efficiency and learning in inter-organizational relationships for product development. Management Science 47 493-511; 2001.

25. Uzzi B. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. American Sociological Review 61(4):674-698; 1996.

26. Wuyts S., Colombo M. G., Dutta S., Noteboom B. Empirical tests of optimal cognitive distance. Journal of Economic Behavior and Organization (58): 277-302; 2005.

27. Zahra, S. A. Technology strategy and new venture performance: A study of corporate-sponsored and independent biotechnology ventures. Journal of Business Venturing 11:289-321; 1996.

CHAPTER 3:

THE IMPACT OF COLLABORATION NETWORK POSITIONS ON INVENTION PERFORMANCE: A SURVEY STUDY IN FLORIDA[1]

3.1. Introduction

An organization's surrounding environment provides channels for flow of information, knowledge or know-how, all necessary for invention. The relationship network of firms, which may be composed of supply chain ties, strategic alliance ties, and social ties constitute conduits that allow such an in-and-out flow. Effective collaboration in supply chains (6, 16, 2) and in alliances (28, 29, 20, 26) is found to have an important impact on the dynamics of knowledge flow. Social network methods are employed to assess the impact of network position on the performance outcomes. Although there is limited research, existing studies indicate that an individual firm's position in its own inter-organizational network influences the firm's innovative performance (30, 3, 31, 25, 28, 14, 15, 23). Also, the structure of an entire network influences the innovative performance of the entire network (5, 4, 18).

In social network studies, several metrics have been devised that are applicable to a network of individuals. For example, "Degree Centrality" is measured by the number of ties connected to an actor in a network. In a research and development (R&D) team of individuals, a larger number of networking ties implies more knowledge sharing and thus inventiveness (24).

[1] Portions of this chapter have been previously published in *Technology and Innov*ation, 2012, 14: 351–363, and have been reproduced with permission from Cognizant Communication Corp. **DOI: http://dx.doi.org/10.3727/194982412X13500042169171.** Image of the written permission can be found at Appendix B.

However, one cannot be confident enough to say that the impact of the metric 'degree centrality' applies similarly to organizational networks. An organization also seeks to increase the number and enhance the quality of ties (collaborations) but maintaining large numbers of ties is costly. Therefore, a firm has to find a balanced strategy that would increase knowledge inflow opportunities with the least cost. One of this study's objectives is to see if centrality measures vs. performance of actors similarly apply at the organizational level.

3.1.1. Metrics to Assess the Network Position and Approaches to Find the Impact

The degree centrality is defined as $DC(a_i) = \sum_j x_{ij}$ where a_i is i^{th} actor in the network and x_{ij} is any existing connection from actor i to j (33). Although not commonly used, it has been utilized as an independent variable in Ahuja's (3) study to assess the effects of a firm's network of relations on innovation. Similarly, Tsai (31) uses it as an independent variable in his study of intra-organizational networks in order to assess the effects of network position on business unit innovation and performance.

Closeness centrality is based on closeness or distance to other actors. In network analysis, a shortest path between two actors is referred to as geodesic (33). The closeness centrality index for actor i is $CC(a_i) = \left[\sum_{j=1}^{g} d(a_i, a_j)\right]^{-1}$, $i \neq j$; where g is the total number of actors in a connected network, $d(n_i, n_j)$ is the number of connections traversed to link actors i and j (27). Abbasi et al. (1) uses it by hypothesizing that normalized closeness centrality of a scholar impacts her research performance in their study of identifying the effects of co-authorship networks on performance of scholars. Uzzi and Spirro (32) use it as a control variable to control for the production team's ability to reach talent in the network of artists in their study on small world network of Broadway musicals.

A dense connectivity in the neighborhood of the organization can be a critical element of an invention process (17). Intense and frequent interaction with other organizations, collective problem solving and trust are facilitated by locally clustered networks. The local clustering coefficient (*LCC*) is a measure of locally dense connectivity. It can be calculated as the proportion of the partners that are themselves directly linked to each other (28).

$$LCC_i = \frac{2|(e_{jk}: v_j, v_k \in N_i , e_{jk} \in E)|}{k_i\,(k_i - 1)} \tag{3.1}$$

E is the representation for a set of edges and e_{ij} is the edge that connects node v_i with node v_j. N_i is the immediate neighbors set for organization i, that is defined as $N_i = \{v_j: e_{ij} \in E \wedge e_{ji} \in E\}$.

A counter argument posits that the cohesion in the cluster can cause the knowledge shared to become homogenous and redundant, which limits invention performance, as opposed to the positive impact view of LCC (13). So there are competing arguments for the impact of *LCC*, which could either result in a positive or negative impact on the invention performance.

Using patents as the dependent variable has challenges that may result in a potential source of bias. For several reasons, the propensity to patent may vary from organization to organization. To address this potential bias, *Patenting Pattern* (PP_i) is used as a control variable. PP_i is calculated as the average number of patents by the organization over the available years prior to the beginning of the analysis window (8, 28). PP_i is a constant value for each organization. It is assumed that it controls the effects of predictors by absorbing uncontrolled variation in order to obtain unbiased prediction results.

For each organization in a network, the cited structural measures can be calculated to see their relative positional differences. The main objective of this study is to assess the impact of centrality (*DC, CC)* and clustering (*LCC)* measures on the innovative performance of

organizations (as measured by the number of patents issued) and the resulting practical implication for making collaboration decisions.

Based on the established theoretical constructs, one can expect that organizations have a high potential to have good innovation performance, if they have more collaborations opportunities, are the closest organizations to all other organizations, are on many shortest (geodesic) paths between other pairs of organizations, and are connected with other centrally located organizations. That means high performing innovators are in the center of a collaboration network. Therefore, the following hypotheses are derived:

H1: Everything else being equal, degree centrality (DC) of an organization impacts their innovation performance.

H2: Everything else being equal, closeness centrality (CC) of an organization impacts their innovation performance.

H3: Everything else being equal, local clustering coefficient (LCC) of an organization impacts their innovation performance.

The tests of the cited hypotheses regarding the metrics used and their predictive power (based on the sample) on innovative performances are discussed in the results section.

3.1.2. Approaches to the Definition of Network and Connection Types

One approach to assess the impact of connectivity on performance has been egocentric analysis where information is obtained only around an immediate locality of a given organization. An egocentric approach does not require a priori enumeration of organizations in a network (21). It also does not take into account the complete map of a network, which may lead to loss of valuable information regarding interactions among neighbors of egos (29). We targeted capturing the map of a whole network with all existing interrelationships among organizations.

The boundary of the network is defined as "The inventive organizations located in the State of Florida".

Another type of approach is to take a single collaboration type into account, for example only a specific set of supply chain ties (12) or alliance ties (28) are considered in a network. In this study, I approached the notion of collaboration holistically by taking into account all possible connections that may allow flow of technological knowledge. Therefore, all types of connections below are considered as 'collaboration' towards knowledge sharing, when organizations report that they have one or more of connections with another organization.

- Alliance tie: Organization has been in the same trade association or consortia, have shared a contractual agreement with another organization.
- Supplier tie: Organization buys products or services from another organization.
- Customer tie: Organization sells products or services to another organization.
- Common ownership: Organization has you has inter-firm cross-holding of equities or property rights.
- Social tie: Organization's managers/inventors have relatives or friends in another organization, excluding the corporate relationships mentioned above.

Furthermore, collaborative network studies that include universities as inventive entities are very limited. Although companies and universities are not pursuing the same exact purposes, they are operating in the same environment, constitute a collaboration network and produce inventions. I included universities in the study along with the companies, keeping in mind that both groups may not be comparable in every aspect.

3.1.3. The Outstanding Problem with Collaborations

Collaborations among organizations are becoming more and more prevalent, but the return on the investment of establishing collaboration is very unpredictable. For example, the number of top ranked inventions borne from collaborative R&D (rather than in-house R&D) increased from 53% in 1975 to 87% in 2006 (8) in the US. While these data show the importance of pursuing such relationships, a large percentage of inter-organizational collaborations fail to live up to expectations. Most studies estimate that 30 to 70% of all collaborations end up failing (23, 7). Therefore, from the standpoint of an organization in the pursuit of innovation, collaboration is increasingly necessary, yet difficult. The challenge is not only to determine whether or not to collaborate, but *with whom* and *how to* collaborate so that innovation is increased. This research also aims to contribute to the enhancement of organizations' collaboration decisions.

3.2. Methodology and Data

3.2.1. Selection of Inventive Organizations

The boundary of the network is defined as those companies and universities located in Florida that have registered at least one patent from 2006 to 2010. The United States Patent and Trademark Office's (USPTO) search facility can pull the list of organizations that have registered patents during certain time periods (34). The stated criteria for the boundary returned 502 organizations. When searching for the list of organizations that have registered patents, the USPTO's search facility returns the results whenever the inventor's residence is Florida. These results, however, included non-Florida based organizations, too. Since Florida based companies and organizations are targeted, I had to check all organizations to make sure whether they have

an operating business in Florida. The search facility also returns individuals who are not affiliated to any organization. These results are excluded from the scope of the study, too, since I only focused on the data at the organizational level.

After exclusions, 298 inventive organizations are left with varying numbers of patents, which were called “focal organizations”. For example, the company with the highest number (Harris Corp.) registered 315 patents during the stated period and the list goes down to the organizations that have 1 registered patent. The high variability in the dependent variable allows a more robust regression analysis. From among 298 focal organizations, 270 of them are companies from various industries, 14 of them are universities and 19 of them are other public institutions.

3.2.2. Survey Design

A web-based survey was developed (IRB Pro00002567) to be sent to focal organizations to inquire about their collaborative ties among each other. Information regarding collaboration ties in the five different categories mentioned above is collected. In the pilot stage, the list of all focal organizations are presented to the respondents and they were each asked to indicate if they have any relationship with any of the organizations from the list. If respondents cannot identify any category of collaboration with any organization in the given list, they are given the opportunity to enter the names of organizations with whom they are collaborating. (After the pilot stage, I presented a smaller list of random organizations to make it more user friendly for respondents.) Newly reported names, of course, constitute a new list of organizations, separate from the focal list. I called them no-patent organization in the analysis.

The survey includes another section, which is related to the innovative characteristics of the respondent organizations. They are asked what percentage of their inventions they have been

able to commercialize. This would help us to see at what levels patenting is representative of innovativeness, that is, new product and service development. Respondents are also asked for what percentage of their new products and processes they have registered a patent. There are some new products and process that inventor-company or organization wants to keep confidential to the extent that they even avoid patent disclosures. The rate that they apply for patents for new products and processes would reveal how representative the registered patents are for the innovations in general. Finally, they are asked to identify what type of connection (supplier, alliance, social, etc.) would impact the new product development efforts most. The results of these can give implications for policy makers in order to decide where to invest in the development of ties within industrial networks to spur innovation.

3.2.3. Survey Responses and Sample Representativeness

A rigorous search was needed to find the contact information of the persons that would be able to answer the questions of the survey. The USPTO gives the names of the inventors, however, e-mail contacts are not provided. Persons were reached from about half of the focal organizations from their websites. Twenty-eight of the focal organizations (about 20% of those reached) responded and 21 of them reported their collaboration ties. Figure 3.1 indicates the representativeness of the sample by comparing them with the whole set of focal organizations.

3.2.4. Strength of Connections

Twenty-one respondents reported 211 connections (any of the five collaboration categories listed above) with other focal organizations or with “no-patent” organizations. For each reported connection, the respondents are further asked to indicate the significance of the connection on the development of the new products and services. The reported significance levels of 211 connections are shown in Table 3.1.

When a connection is reported at a higher level in its significance on the development of the new products and services, it shows that the particular connection has a good potential as a conduit for inflow of technological knowledge. Therefore, I was able to validate the approach regarding the connection types and network definition by observing the percentage of reported significance levels.

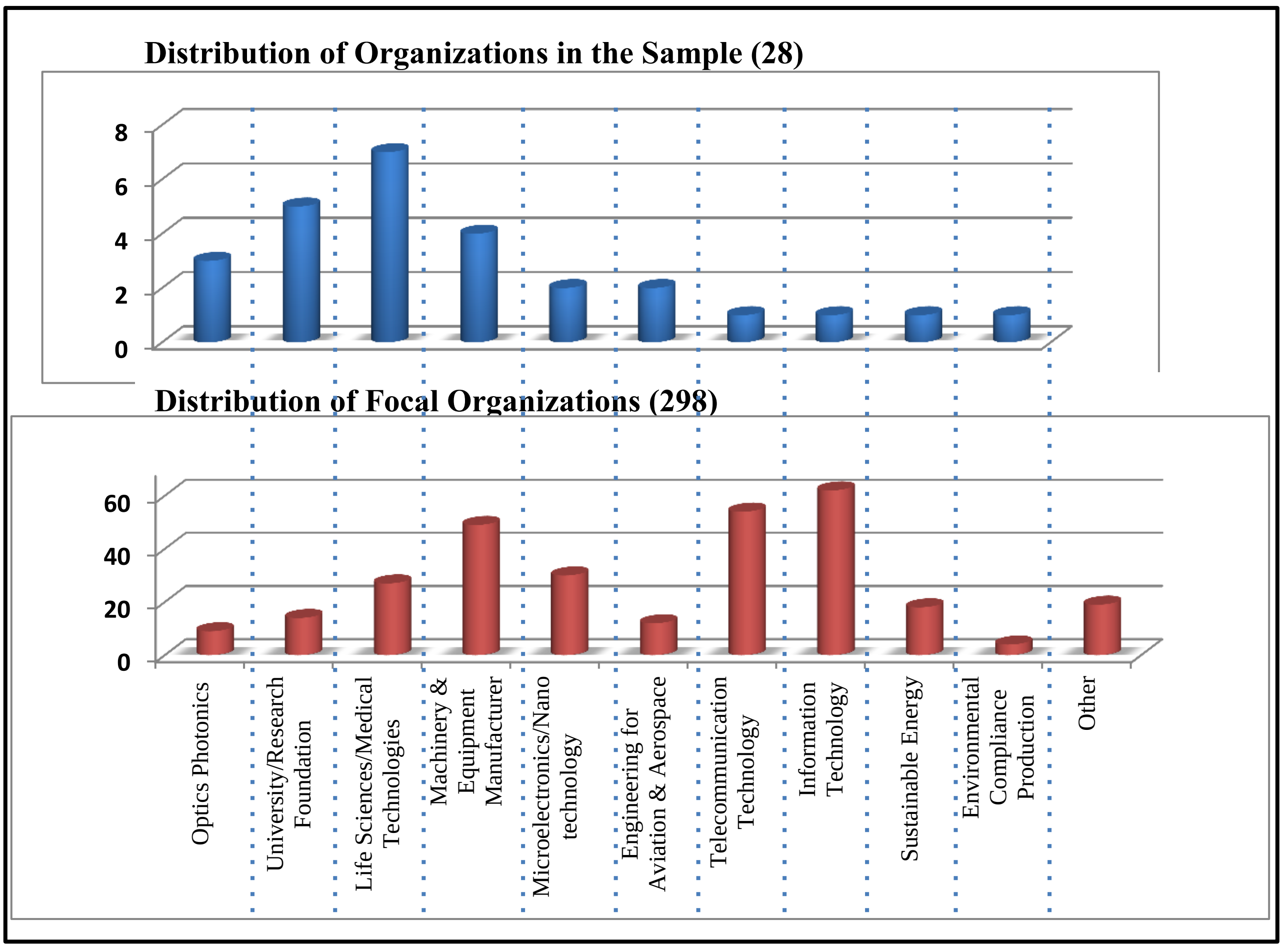


Figure 3.1. The industrial distribution of the sample compared to all focal organizations

'Very Significant' or 'Significant' connections constitute 54% of all connections. Together with 'Neutral' category, they constitute 85% of all connections. On the other hand,

‘Very significant’ and ‘Insignificant’ categories are not excluded from the study, as they may also have the potential for knowledge inflow one would not want to ignore.

Table 3.1. Significance distribution of reported connections

Reported significance	**Number of Connections**
Very significant	59
Significant	55
Neutral	66
Insignificant	13
Very insignificant	18

3.3. Results

3.3.1. Network Map

Organizations are enumerated starting with the focal organizations and continued with the no-patent group. Although 21 of the respondents reported connections, there became 51 focal organizations in the analysis. This is because the respondents reported connections with some other focal organizations that are not respondents to the survey. The respondents reported connections with 83 “no-patent” organizations, so in total, 134 organizations are included in the network. Figure 2.2 shows the collaboration (as identified by five connection types) network map for the sample from Florida’s inventive organizations. Those enumerated up to 51 and indicated by circles are focal organizations and those enumerated from 52 to 134 and indicated by rounded squares are no-patent organizations.

3.3.2. Dependent and Independent Variables

In most studies, the dependent variable, patent counts of organizations ($PatentC_i$), presents overdispersion and this study is no exception. That is, the response variable has a

greater variability than would be assumed in a statistical model. One suggested solution to this problem is to use log of raw count values. Using log-counts provides another advantage in understanding of any regression model result as well. Because the response and the predictors are at similar levels, the regression coefficients would also be easier to interpret.

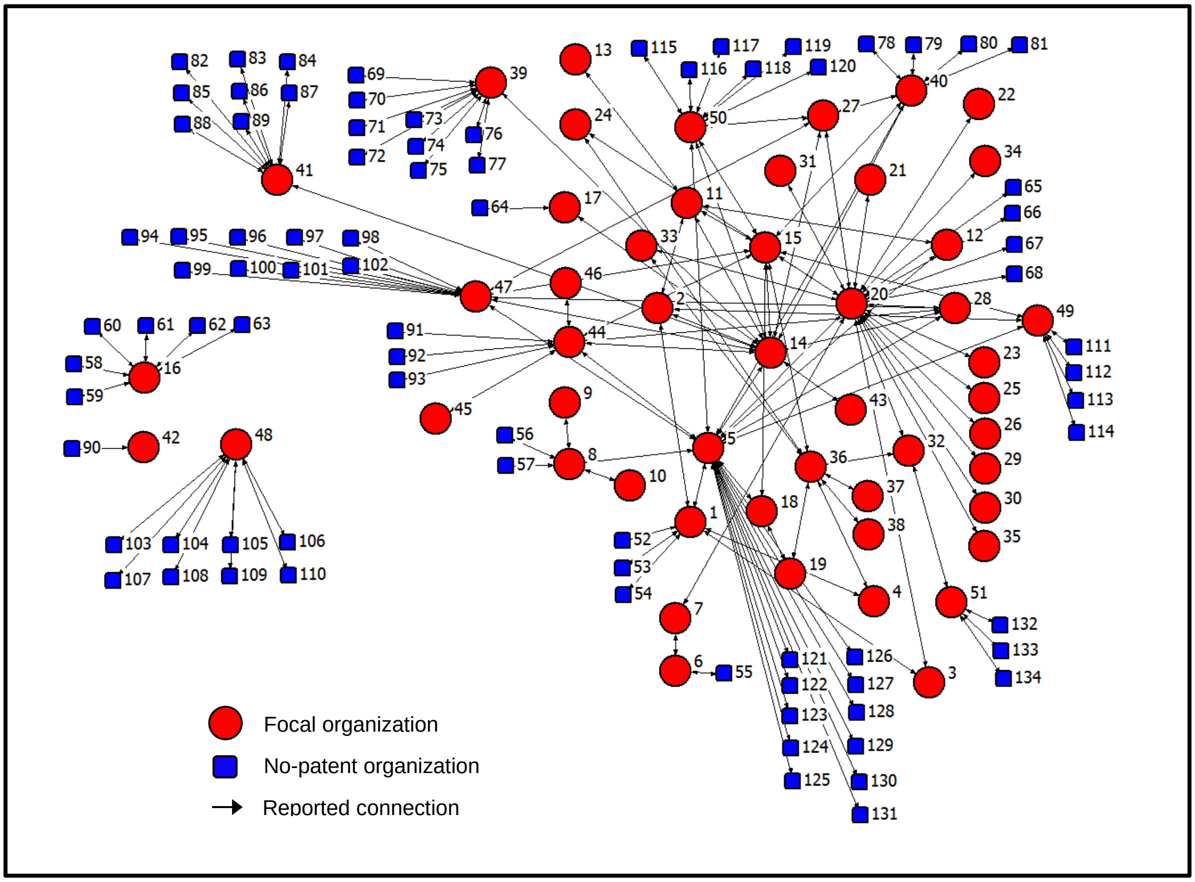


Figure 3.2. Collaboration network of Florida's inventive organizations

Although the values for *CC* and *LCC* are obtained as normalized values, *DC* values are obtained as raw values. In order to avoid the same issues mentioned for *PatentC*$_i$'s, *DC* values are also not used as they are, but log- values are used instead.

3.3.3. Testing Hypotheses

Table 3.2 shows the R output for cross-correlations of all variables. The correlation coefficients for *DC*, *CC* and *LCC* are at about similar levels and all at high significance levels. The hypotheses seemed to be supported by individual correlation values of the sample data. These finding seems contradictory to Tsai's (35) results. On the other hand, the results are similar to Abbasi et al.'s (1) study, where their study finds support of the significant correlation for especially degree centrality impact on research performance. They also are consistent with Ahuja's (3) results where he works on organizations and uses the same dependent variable as in this study. Although the correlations may indicate high association with the patent counts, the true impact can be observed only through a regression model where the predictor's pure impact is visible while controlled by all other variables.

Table 3.2. Cross-correlations among variables and their significance values

```
> rcorr(as.matrix(CS_uw[,-1]))
        PatentC   DC   CC  LCC   PP
PatentC    1.00 0.49 0.42 0.47 0.52
DC         0.49 1.00 0.54 0.41 0.19
CC         0.42 0.54 1.00 0.33 0.21
LCC        0.47 0.41 0.33 1.00 0.58
PP         0.52 0.19 0.21 0.58 1.00

n= 134


P
        PatentC DC     CC     LCC    PP    
PatentC         0.0000 0.0000 0.0000 0.0000
DC      0.0000         0.0000 0.0000 0.0289
CC      0.0000  0.0000        0.0001 0.0162
LCC     0.0000  0.0000 0.0001        0.0000
PP      0.0000  0.0289 0.0162 0.0000       
```

Closeness Centrality (CC) also has a relatively high correlation with Patent Counts (PC). Uzzi and Spirro (32) also find a high correlation (based on different models, from 0.527 to

0.591) for Closeness Centrality while studying its impact on creativity. On the other hand, Abbasi et al. (1) do not find support for Closeness Centrality.

At this moment, a plot of patent counts versus the predictor variables may give a further insight regarding their interactions. Figure 3.3 shows the three predictor variables versus a log of patent counts of 51 focal organizations. Although some association is observable, a regression model would indicate more accurate impact and its significance. Table 3.3 provides the results for the general linear model constructed by the dependent, independent and control variables.

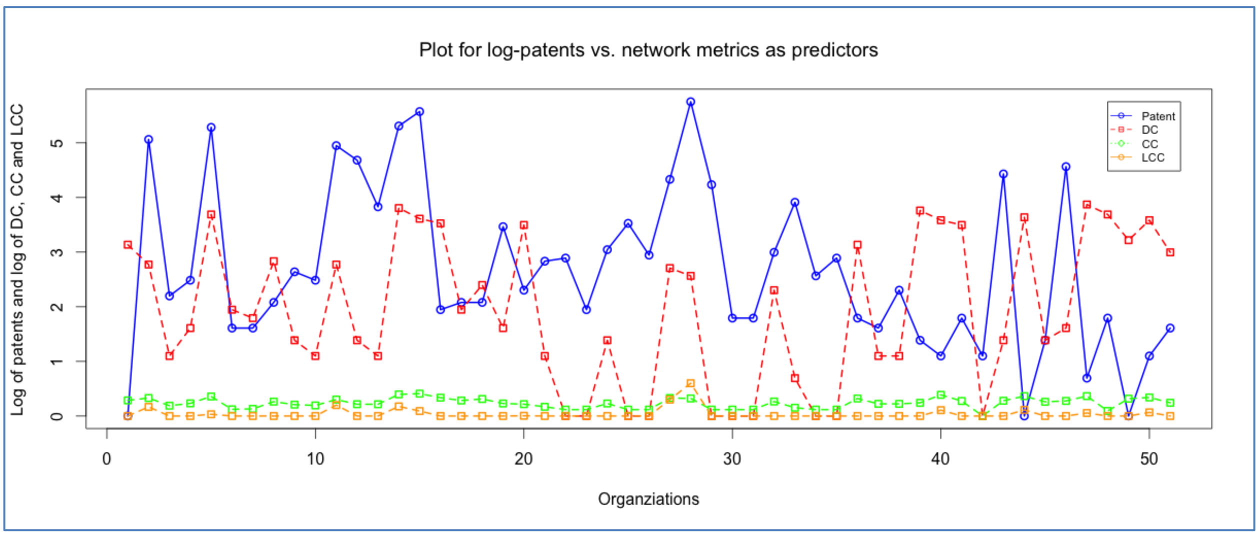


Figure 3.3. A plot of (log of) patent counts and predictors

3.3.4. Regression Analysis

Although the residual distribution is slightly skewed (median -0.33), model significance is at a very satisfactory level ($2.2e^{-16}$). Not that the explanatory power of the regression model is at moderate levels (Adj-R^2=0.45). Given the model parameters, *DC* and *CC* are found to have a significant impact on the patent counts. Although close to an accepted significance level, *LCC* is not found significantly predicting the patent counts.

The analysis results lend support for H1 and H2 but not for H3. Although not very strongly significant, *DC* and *CC* are found to impact the invention count performance.

Table 3.3. Regression analysis results

```
> LM_uw <- lm(PatentC~DC+CC+LCC+PP, data=CS_uw)
> summary(LM_uw)

Call:
lm(formula = PatentC ~ DC + CC + LCC + PP, data = CS_uw)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.1944 -0.4079 -0.3279  0.1510  2.9370 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.77010    0.27038   2.848  0.00512 ** 
DC           0.31047    0.12110   2.564  0.01150 *  
CC           2.33036    1.07784   2.162  0.03246 *  
LCC          2.64083    1.57678   1.675  0.09639 .  
PP           0.10373    0.02015   5.149 9.55e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8899 on 129 degrees of freedom
Multiple R-squared:  0.4643, Adjusted R-squared:  0.4477 
F-statistic: 27.96 on 4 and 129 DF,  p-value: < 2.2e-16
```

3.3.5. Innovative Characteristics Analysis

The responses to invention commercialization, invention patenting and collaboration impact on innovation are provided in Figure 3.4., 3.5., and 3.6., respectively.

3.4. Discussion and Conclusion

I investigated the impact of position in a collaboration network on the innovative performance of organizations. Innovative performance is measured by the number of patents registered by the organizations and used as response variable in the analyses. An inventive organization would seek ways to maximize its technological knowledge inflow opportunities using collaborations. Alliance, supplier and customer relationship, common ownership and social

connections are specified as a type of collaboration useful for knowledge inflow. A survey administered to Florida's innovating organizations (companies, universities and institutions) that

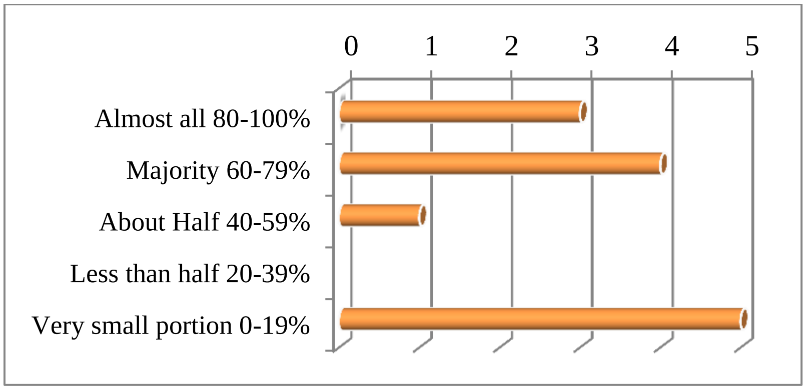


Figure 3.4. What percentage of your patented inventions have you been able to commercialize?

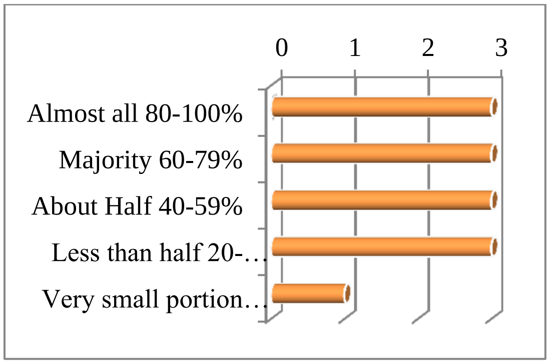


Figure 3.5. For what percentage of your new products and processes have you registered a patent?

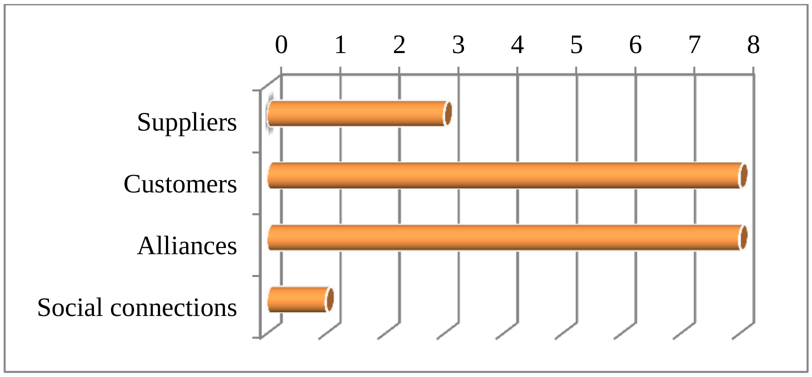


Figure 3.6. The new product development efforts would be positively affected if you could establish better/stronger connections with your … ?

asked them to report their above-listed connections with other Florida based organizations. A collaboration networking map is obtained based on the sample of 28 respondents with the help of UCINET and NetDraw. In order to validate the assumption regarding the knowledge flow potential of reported connections, a further question is asked regarding every reported connection's significance on innovations. 85% of the connections being reported Very Significantly, Significantly or Neutrally effecting their new products and services development efforts.

It was proposed that centrality measures indicating central network position in a collaboration network—Degree Centrality, Closeness Centrality, and Local Clustering Coefficient—play an important role in innovative performance. They are hypothesized to impact the invention performance when everything else is kept constant. I obtained the network structural measures using R in order to test the hypotheses that high values in centrality and clustering metrics indicate higher innovative performance. It suggests that any organization in the pursuit of inventive performance must consider that there is a strong impact of being in such a position that allows for a closer indirect interaction with other members of the collaboration network, besides making high numbers of direct connections.

The results did not give support for the hypothesis that, local connectivity, as measured by local clustering coefficient, has much influence on knowledge recombination and thus inventiveness.

The survey also included a section called "innovative characteristics. The motivation in this section was to learn how inventions are translated into innovations and how good the patent counts are as the dependent variable. The clearest idea out of the inquiry is that the customer and alliance connections have a stronger impact on the efforts of new product and service

development. It is further observable that, manufacturing oriented enterprises tend to see customer connections as more important in new product and process development. The results also indicate that as opposed to the other collaboration ties, customer and alliance connections have a stronger impact on the efforts of new product and service development.

The sample size of the study is one of the limitations of this study. In this regard, I was not able to do industry specific analyses. Since a local industry is studied, the generalizability of the results is likely to be limited. With a further study that could capture the networking map of a larger set of organizations, industry specific analysis could be employed. The contribution of this study is to show the impact of network position in a holistic collaboration network. The results from the sample pointed to the influence of two of the network measures.

3.5. References

1. Abbasi, A.; Hossain, L.; Uddin, S.; Rasmussen K.J.R. Evolutionary dynamics of scientific collaboration networks: multi-levels and cross-time analysis. Scientometrics 89(2), 687-710; 2011.

2. Agarwal, R.; Selen, W. Dynamic Capability Building in Service Value Networks for Achieving Service Innovation. Decision Sciences 40(3): 431-475; 2009.

3. Ahuja G. Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. Administrative Science Quarterly 45(3): 425-455. 2000.

4. Albino V.; Carbonara N.; Giannoccaro I. Innovation in industrial districts: An agent-based simulation model, Int. J. Production Economics 104: 30–45; 2006.

5. Arndt O.; Sternberg R. Do Manufacturing Firms Profit from Intraregional Innovation Linkages? An Empirical Based Answer. European Planning Studies 8(4); 2000.

6. Bessant, J.; R. Kaplinsky; R. Lamming. Putting supply chain learning into practice. International Journal of Operations and Production Management 23(2): 167-184; 2003.

7. Bleeke, J.; Ernst, D. Is Your Strategic Alliance Really a Sale? Harvard Business Review 73(1): 97–105; 1995.

8. Block F. Swimming Against the Current: The Rise of a Hidden Developmental State in the United States. Politics & Society 36-2: 169-206; 2008.

9. Blundell, R. R., R. Griffith, J. Van Reenen. Dynamic count data models of technological innovation. Econom. J. 105 333–344; 1995.

10. Bonacich, P. Factoring and weighting approaches to status scores and clique identification. Journal of Mathematical Sociology 2(1): 113–120; 1972.

11. Borgatti, S. P.; Everett M. G.; Freeman L. C. UCINET for Windows: Software for social network analysis. Harvard, Analytic Technologies; 2002.

12. Boroush M. NSF Releases New Statistics on Business Innovation, NSF 11-300. National Science Foundation, Division of Science Resources Statistics. Retrieved October 2010 from http://www.nsf.gov/statistics/infbrief/nsf11300.

13. Burt R. S. Structural Holes: The Social Structure of Competition; Cambridge, MA: Harvard University Press; 1992.

14. Chi L.; Holsapple C. W.; Srinivasan C. Competitive Dynamics in Electronic Networks: A Model and the Case of Interorganizational Systems. International Journal of Electronic Commerce 11(3): 7–49; 2007.

15. Capaldo A. Network structure and innovation: the leveraging of a dual network as a Distinctive relational capability. Strategic Management Journal 28: 585–608; 2007.

16. Choi Y. T.; Dooley K. J.; Rungtusanatham, M. Supply networks and complex adaptive systems: control versus emergence. Journal of Operations Management 19: 351-366; 2001.

17. Coleman, J. S. Social capital in the creation of human capital. Amer. J. Sociol. 94 S95–S120; 1988.

18. Fleming L.; King III. C.; Juda A.I. Small Worlds and Regional Innovation. Organization Science 18(6): 938–954; 2007.

19. Gibbons, Deborah. E. Inter-organizational network structures and diffusion of information through a health system. American Journal of Public Health 97: 1684-1692; 2007.

20. Gulati, R. Alliances and networks. Strategic Management Journal 19: 293–317; 1998.

21. Hallikas, J.; Varis, J.; Sissonen, H.; Virolainen V.M. The evolution of the network structure in the ICT sector. International Journal of Production Economics, 115(2): 296-304; 2008.

22. Knudsen M.P. The Relative Importance of Interfirm Relationships and Knowledge Transfer for New Product Development Success. Journal of Product Innovation Management 24:117–138; 2007.

23. Kogut, B. The Stability of Joint Ventures: Reciprocity and Competitive Rivalry. Journal of Industrial Economics 38(2): 183–198; 1989.

24. Lovejoy, W. S.; Sinha, A. Efficient Structures for Innovative Social Networks. Management Science 56(7): 1127-1145; 2010.

25. Pittaway L.; Robertson M.; Munir Kamal.; Deyner David.; Neely Andy. Networking and innovation: a systematic review of the evidence. International Journal of Management Reviews 5/6(3&4): 137–168; 2004.

26. Powell W. W.; L. Smith-Doerr. Networks and economic life. N. J. Smelser, R. Swedberg, ed. The Handbook of Economic Sociology. Princeton, NJ: Princeton University Press; 1994: 368–402.

27. Sabidussi, G. The centrality index of a graph. Psychometrica 31: 581-603; 1966.

28. Schilling, M. A.; Phelps, C. C. Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation. Management Science 53: 1113-1126; 2007.

29. Schilling, M.A. Understanding the alliance data. Strategic Management Journal 30: 233-260; 2009.

30. Shan W.; Walker G.; Kogut B. Interfirm Cooperation and Startup Innovation in the Biotechnology Industry. Strategic Management Journal 15(5): 387-394; 1994.

31. Tsai W. Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance. The Academy of Management Journal 44(5): 996-1004; 2001.

32. Uzzi, B; Spirro J. "Collaboration and Creativity: The Small World Problem," American Journal of Sociology 111: 447-504; 2005.

33. Wasserman, S.; Faust K. Social Network Analysis: Methods and Applications. New York, NY: Cambridge University Press; 1994.

34. USPTO Patent Full Text and Image Database. Retrieved July 2011 from http://patft.uspto.gov/netahtml/PTO/search-adv.htm.

CHAPTER 4:

THE IMPACT OF RESEARCH AND DEVELOPMENT (R&D) ALLIANCE NETWORK POSITIONS ON INVENTION PERFORMANCE: A LONGITUDINAL STUDY

4.1. Introduction

Empirical research widely confirmed that the positional variables in inter-organizational alliance networks influence the invention performance outcomes (30, 15, 34, 1, 37, 28, 33, 29, 10, 14, 17, 4, 9). However, the combination of the positional variables used in such research, and the direction and magnitude of their impact, are as varied as the number of studies. Recent studies are diversified in their attempts to explore the impact of specific elements of network impact. For example, whether the knowledge that is fundamental to the invention is tacit or codified (12, 2), whether the cognitive distance between partners is large or small (13, 26, 19), and whether the positional distance to other members of the network and clustering levels are high or low (33, 29, 14, 1) are each found to play an important role in inventiveness.

This research studies the impact of the positional distance measures and the local clustering by using longitudinal data of a large-scale alliance network (see 33 for discussion) in the form of multifold alliances (as opposed to a binary assumption). Research endeavors in this sub-area mostly examine a single industry (1, 29, 15). One exception is Schilling and Phelps (33) who analyzed 11 high-technology manufacturing industries. Due to the nature of their data, however, the researchers usually had to make a strong assumption regarding the estimation of the alliance life. In contrast, this research uses Research and Production Joint Venture (RJV) alliance

data based on the notices of National Cooperative Research and Production Act of 1993 (NCRPA) (35, 25), which provides information regarding both the beginning and ending time of any organization's membership to a RJV since 1994. Considering the membership as connectivity among the members of the RJV in NCRPA notices, it is possible to relieve the usual assumption regarding when an alliance exists and ceases to exist.

Another restriction the previous research encountered was that the alliances had to be treated as binary (i.e., un-weighted) ties. That is, when at least one alliance is announced the tie weight is counted as 1, and otherwise it is counted as 0. This approach, however, ignores the potential multiple connections in-between organizations. It is likely that two organizations may be involved in more than one RJV simultaneously, which is not the same as being in one single RJV simultaneously. The number of common alliances for two organizations is translated as tie weights in network analysis. Therefore, it becomes a matter of taking the "weight" value of the connection as 1, or the actual number of alliances. A recent social network study from Opsahl et al. (27) proposes calculation methods for network centrality and local clustering measures in weighted networks. The empirical research has not yet examined the potential differences in calculations of network positional variables that take tie weights into account, which may well affect the direction and the magnitude of the impact on the invention performance. Using the membership information for RJV in the NCRPA notices, it is possible to relieve the binary assumption and take multifold alliances into account by using this proposed calculation method (27).

To address the abovementioned gap, the impact of two key large-scale network centrality properties (degree centrality and closeness) and local clustering (the term crowding is also used by Stuart, 1998 (34)) is examined on invention output.

4.1.1. Inter-organizational Networks and Patents as Invention Output

One of the definitions for invention is the novel recombination or the reconfiguration of the ways in which knowledge elements are linked. In explaining the process of invention, I adopt the perspective of recombination and selection of ideas via communication through a network of interconnected organizations (33, 24). To solve an industrial problem, two or more organizations decide to come together and form alliances in a venture or project. As more organizations are added to them, or the incumbent members of the alliance decide to form other unique alliances, they weave a network of such alliances. From the network science perspective, organizations are represented as nodes and their communication channels due to being in the same venture/project are represented as connections (i.e., ties). Note that if there are two members in a RJV, it means one (undirected) connection is created between the two organizations. When there are three members, three connections are created between the three members. For four members, six connections are created. So, the number of connections created for a given RJV is obtained by the following formula, where *n* is the number of members in the RJV:

$$\textit{Number of Connections} = \frac{n(n-1)}{2} \tag{4.1}$$

Patent counts are shown to correlate with invention counts (3) and provide a measure of novel invention (19). Numerous studies elect to use patent counts as a proxy for invention performance. Using patents, though, has a few challenges that may result in a potential source of bias. Firstly, for several reasons, the propensity to patent may vary from organization to organization. To address this potential bias, I control for the *Patenting Pattern (PP)* variable that is the average number of patents by the organization over the available years prior to the beginning of the analysis window (6, 33).

Secondly, organizations that enter the analysis time window with different knowledge stocks may be expected to perform differently, which constitutes another source of bias. Another interpretation could be the following: An organization that has successfully registered a patent can be expected to have a tendency to register further patents in the same area. To control for this phenomenon, a *Knowledge Stock Effect (KSE)* variable is used. *KSE* is calculated as the depreciated sum of past inventions and it assumes a different value every year in longitudinal analysis (6).

The US Patent and Trademark Office (USPTO) provides the counts of yearly utility patent grants from 1969 to 2012 by organizations (39). This source provides the granted patents for the organizations that are the sole owner or the first-named assignee. The dependent variable, $Patents_{it}$, is the number of registered patents for organization *i* at year *t*.

Since the analysis window start at 1994 and the patent data is available from 1969, the *PP* value can be calculated for each member on the network over 26 years as:

$$PP_i = \frac{\sum_{1969}^{1993} Patents_i}{26} \tag{4.2}$$

Similarly, the Knowledge Stock Effect for each member at year t is calculated as:

$$KSE_{it} = Patents_{it} + (1 - \delta)KSE_{it} \tag{4.3}$$

δ is called the depreciation rate and the value of 30% is found to be appropriate (6).

4.1.2. Centrality and Clustering Measures

In a network of alliances, the greater the access an organization has to novel knowledge, the better its chances are for making an invention. Access to other members is measured by centrality metrics. One of the primary metrics is the degree centrality (DC), which counts the number of connections from (or to, as connections are undirected) an organization:

$$DC_i = \sum_j x_{ij} \tag{4.4}$$

DC_i is the degree centrality value for i[th] organization, and x_{ij} is any existing connection from actor i to j (40). DC_i is a measure of direct contacts but it is known that the diffusion of knowledge also takes place via indirect contacts (27, 4, 7). Closeness centrality (CC) measures the inverse network distance to all other members of the network. Before moving further into the calculation of CC, another important network concept, geodesic, must be introduced. Geodesic refers to the shortest path between the two nodes in a network (40). The CC, then is the inverse of the summation of the geodesics from an organization *i* to any other organization *j* in the network. The geodesic distance is represented as $\mathrm{d}(\mathrm{i}, \mathrm{j})$ (28).

$$CC_i = \left[\sum_j d(i,j)\right]^{-1}, \ i \neq j \ ; \qquad (4.5)$$

Some critical elements of an invention process, like intense and frequent interaction with other organizations, collective problem solving and trust are facilitated by a dense connectivity in the neighborhood of the organization (11). The local clustering coefficient (*LCC*) is a measure of locally dense connectivity. It can be calculated as the proportion of the partners that are themselves directly linked to each other (33).

$$LCC_i = \frac{2|(e_{jk}: v_j, v_k \in N_i\,, e_{jk} \in E)|}{k_i\,(k_i - 1)} \qquad (4.6)$$

E is the representation for a set of edges and e_{ij} is the edge that connects node v_i with node v_j. N_i is the immediate neighbors set for organization *i*, that is defined as $N_i = \{v_j : e_{ij} \in E \ \wedge e_{ji} \in E\}$.

As opposed to the positive impact view of LCC, a counter argument posits that the cohesion in the cluster can cause the knowledge shared to become homogenous and redundant, which limits invention performance (8). So there are competing arguments for the impact of *LCC*, which could either result in a positive or negative impact on the invention performance.

The hypotheses from H1 to H3b follow:

H1: Higher standardized degree centrality (DC) is associated with higher invention performance.

H2: Higher standardized closeness centrality (CC) is associated with higher invention performance.

H3a: Higher local clustering coefficient (LCC) is associated with higher invention performance.

H3b: Higher local clustering coefficient (LCC) is associated with lower invention performance.

4.1.3. Calculation of Variables in Weighted Networks

In social network analysis, the strength (weight) of a tie is explained as a function of factors such as the tie's intensity or the amount of services exchanged (18). The fact that the strength of ties is not taken into account is a major limitation in studying large-scale networks (27). It is quite common organizations become members in more than one RJV simultaneously. The intensity and the amount of technological knowledge exchange between two organizations may differ if they share one RJV in common versus twenty-one[1]. Opsahl et al. (27) proposes a tuning parameter, say α (from 0 to 1), to set the relative importance of one tie (i.e., un-weighted) as compared to multiple ties (i.e., full weight is considered). When the parameter value is 0, the multiple connections between the two members are taken as just one. This is the binary assumption and completely ignores the tie strengths. When the parameter value is set to 1, the measure is based on the tie weights completely. That is, each common membership in between two organizations' multiple connections must be treated like all other connections in the

[1] It is observed that there are up to 21 common RJVs in NCRPA-1993 data.

network. Figure 4.1. illustrates an example of the function of the tuning parameter in a simplified network of two RJVs. Assume that firms A, B and C come together to solve an industrial problem and establishes an RJV. Similarly, B, C and other two firms E and D decide to establish another RJV. Now, all connections can be considered to have strength 1, except for the one in between B and C. It could be treated as 1, 2 or somewhere in between when adjusted by α parameter.

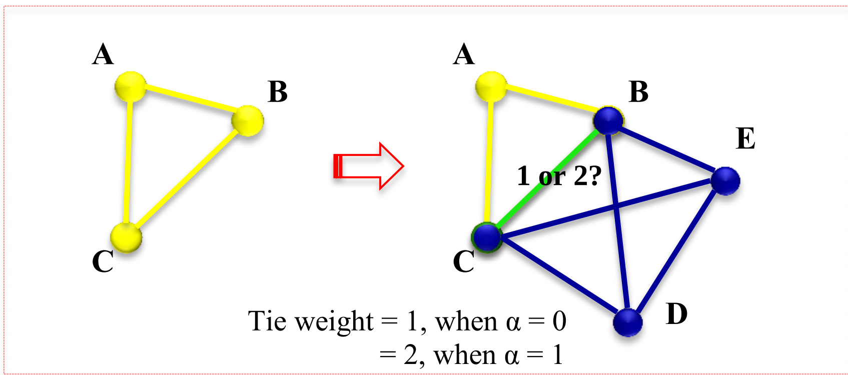


Figure 4.1. An example of the function of the tuning parameter α

Although the tuning parameter here is discussed on the measurement of degree centrality, the same approach applies to the measurement of closeness centrality and the local clustering coefficient as well. As Opsahl et al. (27) find significant changes in the metric values in various example networks, the magnitude of the impact of the network variables in this study is expected to be moderated by the tuning parameter. The three hypotheses follow:

H4a: The tuning parameter for calculation of standardized degree centrality (DC) moderates its impact on invention performance.

H4b: The tuning parameter for calculation of standardized closeness centrality (CC) moderates its impact on invention performance.

H4c: The tuning parameter for calculation of local clustering coefficient (LCC) moderates its impact on invention performance.

4.2. Data

4.2.1. Selection of the Organizations

In order to test the hypotheses, a balanced fixed-panel data set is constructed for 63 U.S. organizations[2] from 1994 to 2012. A prior enumeration of organizations is made to be able to observe each individual organization's network and patent data in a longitudinal analysis. Panel data models are more efficient than cross-sections or pooling cross sections data, since the observations of the same individual organizations over several time points reduce the variance with respect to random selections of individuals over the same period.

The choice of organizations was particularly important for two reasons: I intended to capture the greatest possible variability in the dependent variable, ($Patents_{it}$), across the organizations. Also, the intention is to select such a group so that I can include as large an amount of network activity as possible in order to capture the highest technological knowledge flow potential through the alliances. The patent count data available from USPTO provides several grouping options, including by organization for every year (39). Choosing the organizations with the highest number of patent counts serves both purposes, because there is a good variability in patent counts and the organizations at the top tend to make alliances among

[2] The list is primarily composed of companies (58 of them) although there are a few public institutions (3 of them) and universities (2 of them). I opt to keep the public institutions and universities as the members of the network. The purpose of their existence in the network fits with the previously given definitions and the assumptions.

themselves mostly[3]. Since the analysis period is defined from 1994 to 2012 (due to the NCRP Act of 1993), the organizations are sorted based on which were granted the most number of patents from 1994 to 2012. USPTO makes no attempt to combine data based on subsidiary relationships. I opted to combine all branches or subsidiaries into one parent company that has a distinct headquarters[4]. Their listings are merged and the number of patents are summed up accordingly[5]. After the combining process, the organizations that were granted a total of 2000+ patents from 1995[6] to 2012 were selected and enumerated for analysis (63 of them)[7]. Therefore I obtained a balanced panel-data of the dependent variable, ($Patents_{it}$), for all organizations, from 1995 to 2012.

4.2.2. Alliance Network Data

National Cooperative Research and Production Act of 1993 (NCRPA) is a U.S. federal law that establishes a rule of reason for evaluating the antitrust implications of Research Joint

---

[3] The alliance data confirm this statement. For example, in 2012, the number of alliances (i.e., common memberships detected, see 2.1.1. for details) that top-20 organizations (organizations that were granted the most number of patens in 2012) make among themselves is 1,351. The number of alliances that top-20 organizations make with the next-20 organizations (that are organizations from 21 to 40) is 583. That means, a sample of organizations at the top make around 2.5 times more alliances among themselves than they make alliances with the sample of the same size nearest to the top. The same phenomenon is observed in all other years (e.g., the ratio is 5.0 in 1994).

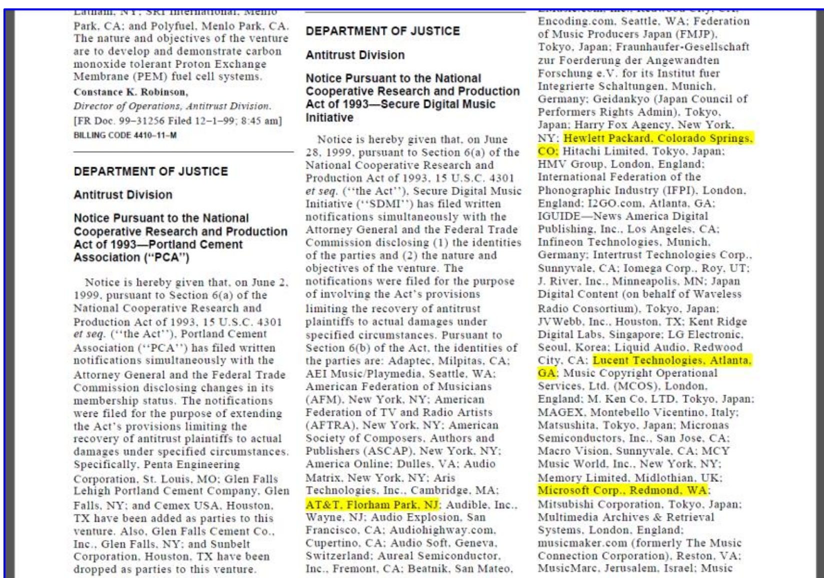
[4] For example, the AT&T Corporation's patents were listed under its name "AT&T Corporation" until 2007. Starting 2008, they were mostly listed under another name "AT&T Intellectual Property, L.P.". Some companies had several subsidiaries and branches (14 lines merged for Siemens).

[5] Naturally, some of the companies make mergers, cease to exist or newly emerge over the 18-year analysis period. I tried to include the organizations that represent a network membership at best. For example, Sun Microsystems has been acquired by Oracle in 2010. I included both Oracle and Sun in the analysis because Sun has been a member in the majority of the analysis period. However, as this is an obvious limitation the individual results like Sun must be evaluated by caution.

[6] The counts in 1994 are excluded for the reasons of causality. Simply, the (earliest) alliance network structure of 1994 is not expected to impact the patents registered in 1994.

[7] The complete list can be found in Appendix C.

Ventures (RJVs) on an individual case basis and reduces potential antitrust liabilities (25), and is designed to promote innovation, facilitate trade, and strengthen the competitiveness of U.S. firms and institutions. The establishment of an RJV and updated membership listings are publicly announced in the Federal Register (FR). Through the FR search facility NCRPA notices can be searched specifically (16). After collecting all NCRPA notices in a specific year, I combined all the documents and highlighted whenever one of the 63 organizations was cited for any reason. Figure 4.2. is an example of a highlighted NCRPA notice that announces the membership listings on December 2, 1999, in which I capture information for four of the 63 organizations.



Park, CA; and Polyfuel, Menlo Park, CA. The nature and objectives of the venture are to develop and demonstrate carbon monoxide tolerant Proton Exchange Membrane (PEM) fuel cell systems.

Constance K. Robinson,

Director of Operations, Antitrust Division.

[FR Doc. 99–31256 Filed 12–1–99; 8:45 am]

BILLING CODE 4410–11–M

DEPARTMENT OF JUSTICE

Antitrust Division

Notice Pursuant to the National Cooperative Research and Production Act of 1993—Portland Cement Association ("PCA")

Notice is hereby given that, on June 2, 1999, pursuant to Section 6(a) of the National Cooperative Research and Production Act of 1993, 15 U.S.C. 4301 *et seq.* ("the Act"), Portland Cement Association ("PCA") has filed written notifications simultaneously with the Attorney General and the Federal Trade Commission disclosing changes in its membership status. The notifications were filed for the purpose of extending the Act's provisions limiting the recovery of antitrust plaintiffs to actual damages under specified circumstances. Specifically, Penta Engineering Corporation, St. Louis, MO; Glen Falls Lehigh Portland Cement Company, Glen Falls, NY; and Cemex USA, Houston, TX have been added as parties to this venture. Also, Glen Falls Cement Co., Inc., Glen Falls, NY; and Sunbelt Corporation, Houston, TX have been dropped as parties to this venture.

DEPARTMENT OF JUSTICE

Antitrust Division

Notice Pursuant to the National Cooperative Research and Production Act of 1993—Secure Digital Music Initiative

Notice is hereby given that, on June 28, 1999, pursuant to Section 6(a) of the National Cooperative Research and Production Act of 1993, 15 U.S.C. 4301 *et seq.* ("the Act"), Secure Digital Music Initiative ("SDMI") has filed written notifications simultaneously with the Attorney General and the Federal Trade Commission disclosing (1) the identities of the parties and (2) the nature and objectives of the venture. The notifications were filed for the purpose of involving the Act's provisions limiting the recovery of antitrust plaintiffs to actual damages under specified circumstances. Pursuant to Section 6(b) of the Act, the identities of the parties are: Adaptec, Milpitas, CA; AEI Music/Playmedia, Seattle, WA; American Federation of Musicians (AFM), New York, NY; American Federation of TV and Radio Artists (AFTRA), New York, NY; American Society of Composers, Authors and Publishers (ASCAP), New York, NY; America Online; Dulles, VA; Audio Matrix, New York, NY; Aris Technologies, Inc., Cambridge, MA; AT&T, Florham Park, NJ; Audible, Inc., Wayne, NJ; Audio Explosion, San Francisco, CA; Audiohighway.com, Cupertino, CA; Audio Soft, Geneva, Switzerland; Aureal Semiconductor, Inc., Fremont, CA; Beatnik, San Mateo,

Encoding.com, Seattle, WA; Federation of Music Producers Japan (FMJP), Tokyo, Japan; Fraunhaufer-Gesellschaft zur Foerderung der Angewandten Forschung e.V. for its Institut fuer Integrierte Schaltungen, Munich, Germany; Geidankyo (Japan Council of Performers Rights Admin), Tokyo, Japan; Harry Fox Agency, New York, NY; Hewlett Packard, Colorado Springs, CO; Hitachi Limited, Tokyo, Japan; HMV Group, London, England; International Federation of the Phonographic Industry (IFPI), London, England; I2GO.com, Atlanta, GA; IGUIDE—News America Digital Publishing, Inc., Los Angeles, CA; Infineon Technologies, Munich, Germany; Intertrust Technologies Corp., Sunnyvale, CA; Iomega Corp., Roy, UT; J. River, Inc., Minneapolis, MN; Japan Digital Content (on behalf of Waveless Radio Consortium), Tokyo, Japan; JVWebb, Inc., Houston, TX; Kent Ridge Digital Labs, Singapore; LG Electronic, Seoul, Korea; Liquid Audio, Redwood City, CA; Lucent Technologies, Atlanta, GA; Music Copyright Operational Services, Ltd. (MCOS), London, England; M. Ken Co. LTD, Tokyo, Japan; MAGEX, Montebello Vicentino, Italy; Matsushita, Tokyo, Japan; Micronas Semiconductors, Inc., San Jose, CA; Macro Vision, Sunnyvale, CA; MCY Music World, Inc., New York, NY; Memory Limited, Midlothian, UK; Microsoft Corp., Redmond, WA; Mitsubishi Corporation, Tokyo, Japan; Multimedia Archives & Retrieval Systems, London, England; musicmaker.com (formerly The Music Connection Corporation), Reston, VA; MusicMarc, Jerusalem, Israel; Music

Figure 4.2. An example of an NCRPA notice.

Given the information from Figure 4.2. only, the four organizations (i.e., AT&T, Hewlett Packard, Lucent and Microsoft) are considered completely connected. Therefore, I define n(n-1)/2 = 6 single undirected connections among them. If any of the two from the four happens to be announced in another RJV in 1999, then I define another connection in between the two. So,

there becomes two undirected alliance connections between them. For every year, up to 150 notices are merged, lines highlighted when one of the 63 organizations is cited and analyzed for the common membership. They are not only cited for membership announcements, but also for withdrawals, mergers or any other actions of interest. In a withdrawal case, the connections defined for a particular RJV are omitted.

The connections are handled by forming adjacency matrices in MS Excel[8].

After the formation of adjacency matrices for every year, I obtained the balanced panel-data for the independent variables and control variables.

4.3. Model Specification

Panel data allows control for variables that cannot be observed or measured. In this study, they account for individual heterogeneity across organizations and time effects. A number of strategies are employed to control for unobserved heterogeneity. *Patenting Pattern (PP)* is an example to control for unobserved individual firm effects. *PP* is employed to control for unobserved, temporally constant differences in patenting across organizations. Furthermore, the *Knowledge Stock Effect (KSE)* is introduced to serve control for unobserved differences in organizations' knowledge stocks. It is also important to consider appropriate lag structure of the independent variables against the dependent variable. Apparently, today's network positional variables is very unlikely to impact the invention outputs of the same year. For this reason, I

[8] At first, I created columns for all RJVs and rows for all organizations. Whenever a common membership is detected of four organizations in an RJV, say in 1999, I wrote "99" in the corresponding cells in the intersections of the particular RJV and the organizations. COUNTIFS function with the criterion "99" collects the number of common memberships in between any two organizations, which, in fact, becomes the value of the corresponding cell in the adjacency matrix.

select to estimate models using one-year, two year and three year lags (33). I will also have a chance to explore the robustness of the regression model.

There are two prominent regression models suggested in the literature for panel data structure. Based on the different approaches to control for unobserved heterogeneity, the Fixed *Effect (FE) Model* is applied whenever the analysis of the impact of variables that vary over time are of interest. The underlying assumption in *FE* model is correlation between the observed entity's error term and the predictor variables (36). The *Random Effect (RE) Model*, however, requires that the error term is uncorrelated with the predictor variables and variation across entities are assumed random. This assumption allows for time-invariant variables to play a role as explanatory variables, which are already specified above. As I do not have any prior belief on what assumption to be made, both *FE* and *RE* models are to be applied. Since the hypothesized predictors are degree centrality (*DC*), closeness centrality (*CC*) and local clustering coefficient (*LCC*) add the control variables are *PP* and *KSE*, the model can be specified as follows:

$$Patent_{it+1}\, ,\, Patent_{it+2}\, ,\, Patent_{it+3} = f(DC_{it}\, ,\, CC_{it},\, LCC_{it},\, FE_{i},\, KSE_{it}\,) \qquad (4.7)$$

4.4. Results

4.4.1. Descriptive Results

In order to start with an accurate description of the networks and the data, a series of descriptive results are presented before the test of hypotheses. From 1994 to 2012, the names of 63 organizations were cited at least one time in one of the 371 RJVs analyzed. Each year's network map is obtained using network adjacency matrices, which are constructed based on the common membership information in RJVs. Figure 4.4 presents an example map from the most recent (2012) alliance network, obtained from UCINET and NetDraw (7). Both pictures are

different representation of the same network. The top figure lists the four organizations without any connections in the top-left corner. The bottom figure on the other hand, lists all organizations with less than 5 connections. Therefore, the organizations constituting the circle at the bottom have 5 or more connections. As the Federal Register is a publicly open source, I felt no restriction to give the names of the companies and institutions but they are mostly abbreviated for the purposes of analyses.

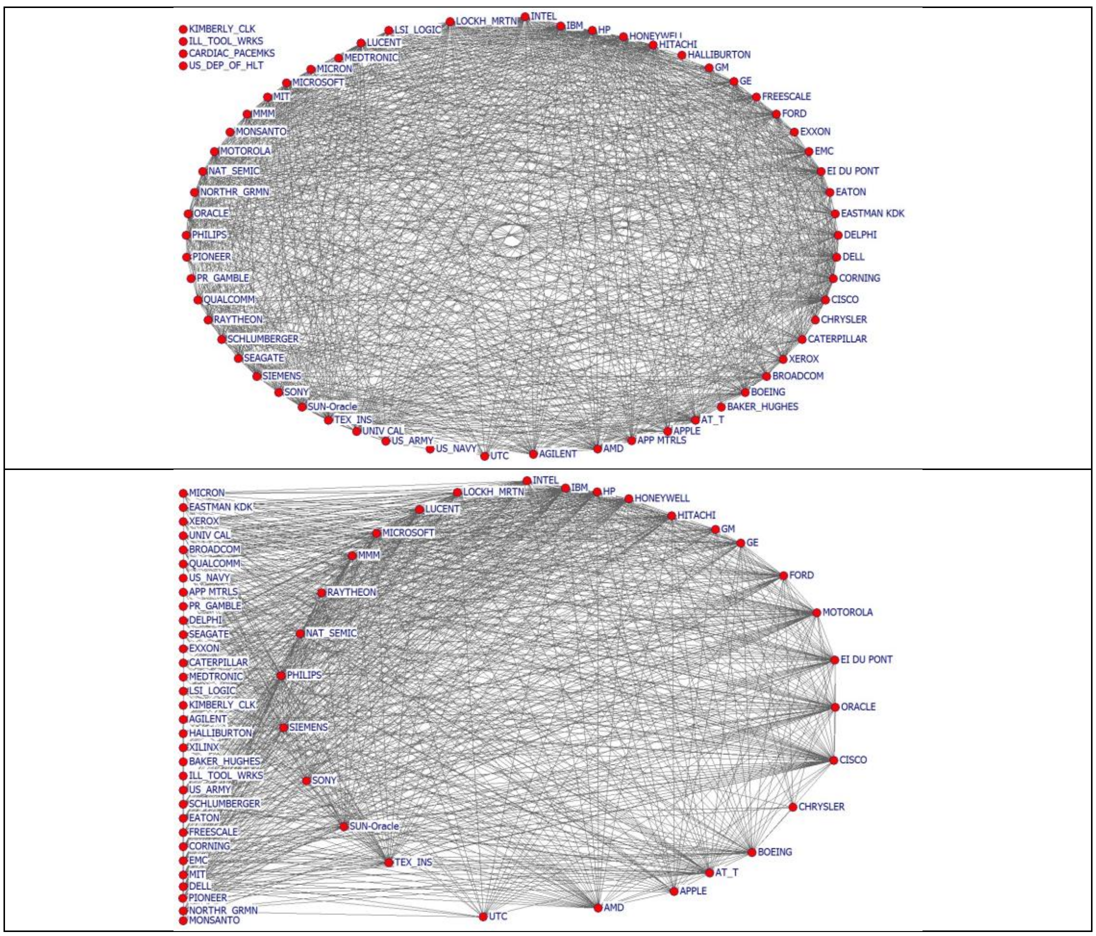


Figure 4.3. The alliance network map from 2012

The connection data obtained from notifications between 1994 to 2012 is accumulated into a 63x371 matrix. Each cell contains information regarding when the organization entered the corresponding venture, if ever it did, when ceased membership, when re-joined, etc. over 19 year

period. Although most cells in the matrix is empty, the information in the cells provide the membership history over 19 years. Therefore, the matrix is extended towards another dimension that indicate years. Consequently, a three-dimensional array is obtained, which allow for several interesting analyses regarding the RJVs and the organizations.

Table 4.1 provides some descriptive characteristics of the networks across the years. For the 63 organizations analyzed, the average number of alliances peaked in 2005, then slightly decreased. Interestingly, the percentage of the members in the main network component assumes its smallest values in the late 90s and then enters a never-decreasing trend.

Table 4.1. Descriptive characteristics of the alliance networks

Years	**Avg. number of alliances per organization**	**Percentage in main network component**	**Number of members (among 63) in the largest RJV**	
1994	7.0	0.73	16	Open Software Foundation
1995	11.2	0.81	18	Open Software Foundation
1996	13.6	0.84	18	Open Software Foundation
1997	17.2	0.71	18	Open Software Foundation
1998	21.7	0.64	18	Asynchronous Transfer Mode (ATM) Forum
1999	25.4	0.60	18	Open Software Foundation
2000	35.1	0.78	24	Infiniband Trade Association
2001	36.0	0.78	24	Infiniband Trade Association
2002	36.5	0.79	24	Infiniband Trade Association
2003	37.1	0.92	24	Infiniband Trade Association
2004	38.1	0.92	24	Infiniband Trade Association
2005	45.0	0.94	24	Infiniband Trade Association
2006	40.2	0.94	24	Infiniband Trade Association
2007	39.5	0.94	24	Infiniband Trade Association
2008	39.4	0.94	24	Infiniband Trade Association
2009	38.7	0.94	24	Infiniband Trade Association
2010	37.5	0.94	24	Infiniband Trade Association
2011	39.3	0.94	24	Infiniband Trade Association
2012	39.0	0.94	24	Infiniband Trade Association

Another interesting results of the membership information is the analysis average length of affiliation with an RJV. The organizations vs. RJV matrix (63x371) has 19 layers in the 3rd dimension and for each cell of (63x371) matrix, a calculation through the 3rd dimension can be made to find an organization's years of affiliation to an RJV. Using this method, a (63x371) matrix of a number of affiliation years is obtained. In fact, this important information was one of the motivation for this study where previous literature had to make assumption regarding the duration of alliances. Average number of affiliation years is obtained both averaging across organizations, which finds average number of years an RJV keeps an organization (a vector of 371 elements), and averaging across RJVs, which also finds average number of years an organization stays with an RJV (a vector of 63 elements). Essentially, either one has the same meaning. Since their histogram plots give similar pictures, one of them is provided in Figure 4.4.

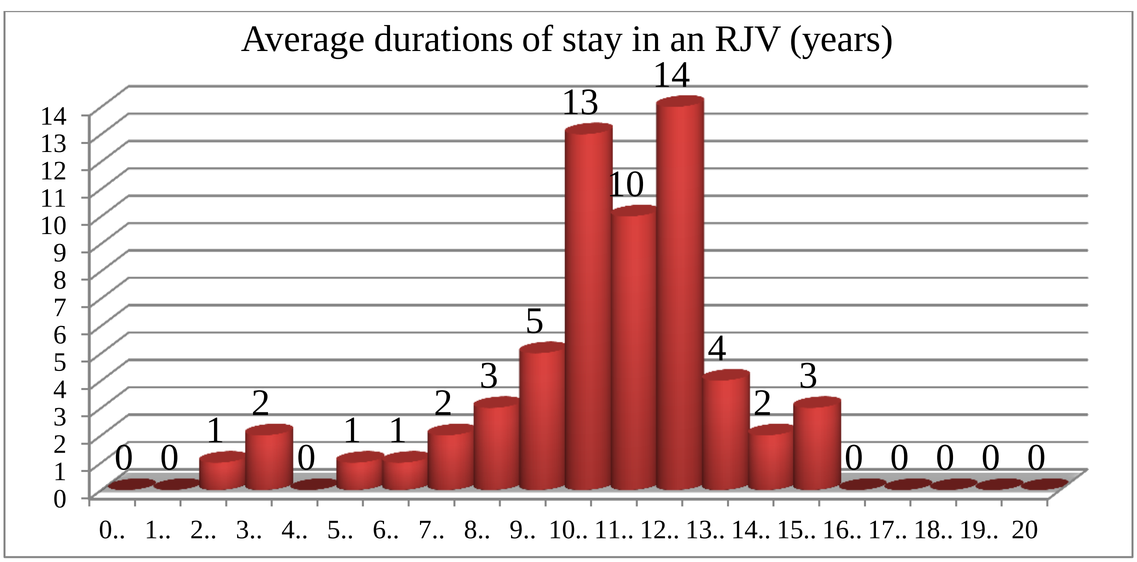


Figure 4.4. Frequencies for average number of years an organization stays with an RJV

The horizontal axis shows the number of years and the vertical axis shows the frequencies of the average durations of stay. Note that, these results are obtained from a sample

of highly successful set of organizations. Although studies estimate that 30 to 70% of all alliances end up failing to serve the purpose (23, 5), the above results are more likely to be representative of effective RJVs.

4.4.2. Correlational Analyses

Based on the 19 membership matrices, an adjacency matrix is obtained for every year using MsExcel's COUNTIFS function[9]. The adjacency matrices (63x63) are transferred to R-Studio (40) environment for the following analyses. For each year, besides $Patent_{it}$ data, *DC, CC and LCC* data are also calculated. *DC, CC and LCC* data are obtained as normalized values. In order to reduce overdispersion in $Patent_{it}$ data, literature suggest making analyses using log-values. It also brings dependent variable values to similar digit levels with predictors, which is an advantage in interpretations of regression coefficients that is discussed in the following section.

In order to allow a better insight regarding how predictor variables correlates with the patent counts, a longitudinal plot is obtained for each predictor variable. Since there are 63 organizations and it is impractical to include all in one plot, a representative set of 4 companies are shown in Figure 4.5. Note that the degree centralities (the plot at the top) make peak values at 2005 and drops afterwards, consistently with the information in the Table 4.1. It is also noticeable that there is very small variation in LCC, which does not seem to highly correlate with *Patents.*

Instead of pooling the data over the years, the correlation values for all organizations are obtained on yearly basis due to a potentially unobserved variability over the years.

[9] For every cell of each adjacency matrix, the COUNTIFS function checks the two corresponding rows (of the two organization that is represented in the cell) of the (63x371) membership matrix.

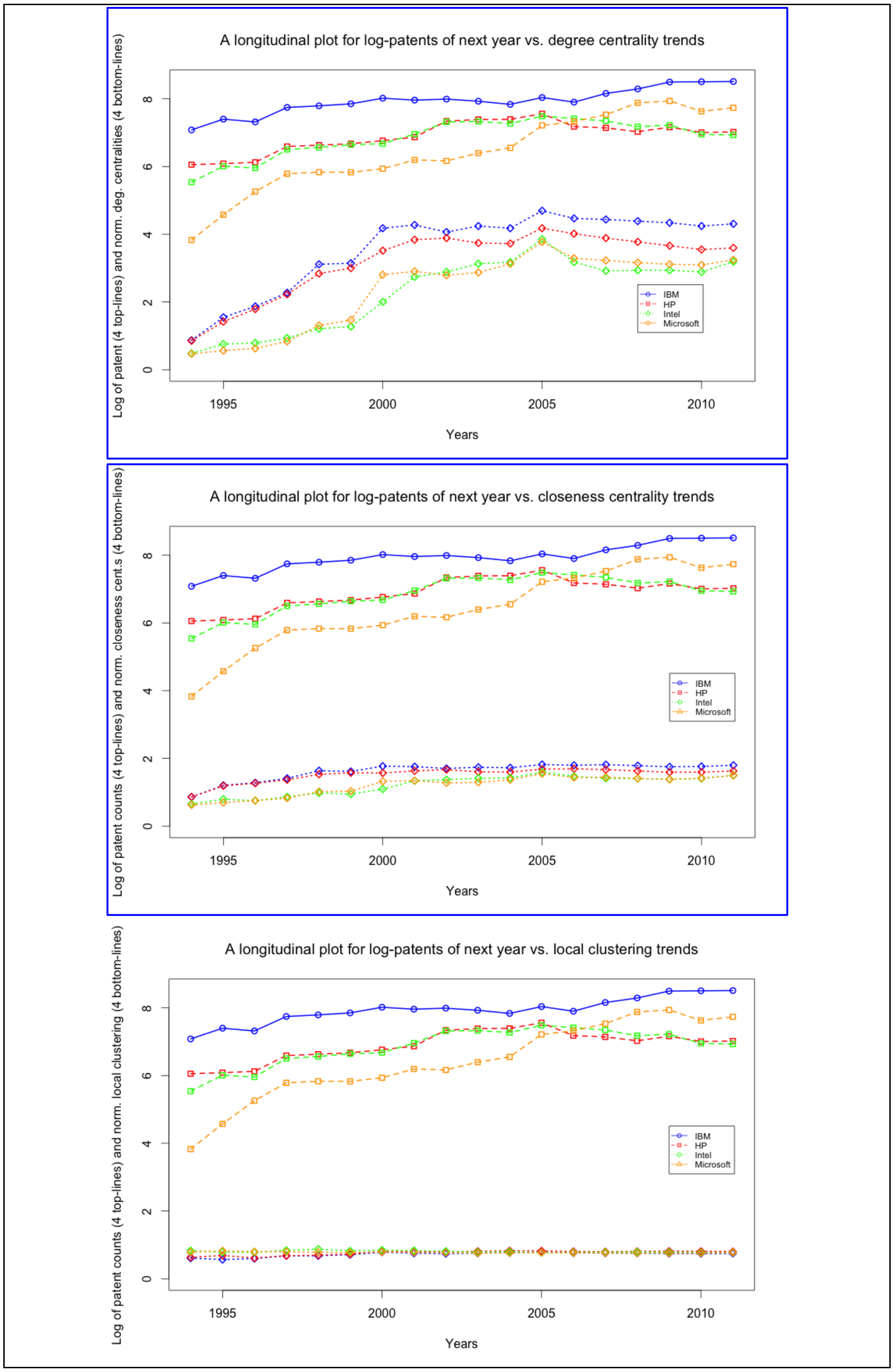


Figure 4.5. Longitudinal plots for *DC*, *CC* and *LCC* versus *Patents* for selected sample

Each year's correlation values for six pairs (*DC vs. Patents, CC vs. Patents, LCC vs. Patents, DC vs. CC, DC vs. LCC, CC vs. LCC*) are obtained and only minimum and maximum values are presented here. Table 4.2 presents the maximum (at the top) and the minimum (at the bottom) of all years' correlation values between the variables. Besides the correlation values, the year of maximum and minimum value is also presented nearby.

Table 4.2. Maximum and minimum correlation values of the variables

	$Patent_{it+1}$	DC_{it}	CC_{it}	LCC_{it}
$Patent_{it+1}$	1			
DC_{it}	0.667 $DC,_{2005}vsP,_{2006}$ -0.029 $DC,_{2011}vsP,_{2012}$	1		
CC_{it}	0.595 $CC,_{2005}vsP,_{2006}$ -0.022 $DC,_{2011}vsP,_{2012}$	0.963 $CC,_{2003}vsDC,_{2003}$ 0.899 $CC,_{2004}vsDC,_{2004}$	1	
LCC_{it}	0.340 $LCC,_{2005}vsP,_{2006}$ -0.0645 $LCC,_{2006}vsP,_{2007}$	0.481 $LCC,_{2003}vsDC,_{2003}$ 0.110 $CC,_{2010}vsDC,_{2010}$	0.750 $LCC,_{1994}vsDC,_{1995}$ 0.544 $LCC,_{2009}vsDC,_{2010}$	1

Big variances are noticeable between the minimum correlation values over the years. The only exception is the *DC-CC* correlation, where their correlation is found to be quite high, even at the minimum year. This, actually brings up another issue towards the model construction. Originally, *DC, CC and LCC* were to be used in the same regression model, however, *DC and CC* cannot be used together. Instead, two separate models are constructed. One model uses *DC, LCC and DCxLCC* interaction effect as predictors, the other one uses *CC, LCC and CCxLCC* interaction effect as predictors. Also, *KSE* is found to be so highly correlated with patents (up to r=0.98) that the random effect model could have fail to find the predictors' impact. I re-estimated the models using *PP* instead of both *PP* and *KSE* and constructed the *Fixed Effect* and *Random Effect Models.*

The following two plots show the heterogeneity across the organizations and across the years, which lays the basis for the assumptions to use the *FE* and *RE* models. Figure 4.6 shows

both plots where at the top, the boxplots for 63 organizations' response variable (*Patents*$_{it}$) are presented and at the bottom, the boxplots for 18 years' response variables are presented. Heterogeneity across organizations are controlled and across time-factor will be checked by another model.

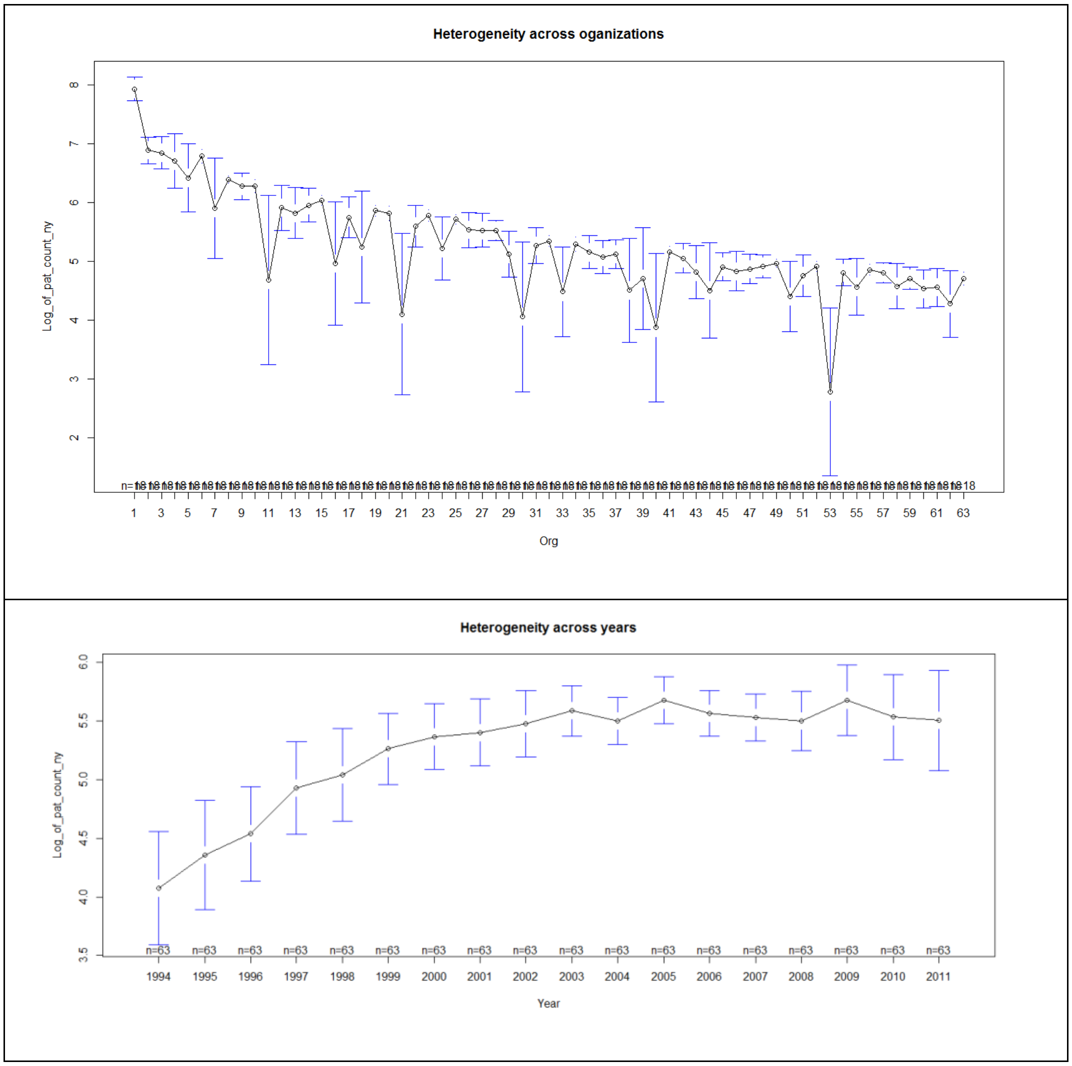


Figure 4.6. Heterogeneity in response variable across organizations and years

4.4.3. Regression Analyses

The two regression model results of the model *DC*, *LCC and DCxLCC* using *FE* and *RE* models are provided in Table 4.3. In the fixed effect model, both *DC* and *LCC* are found to have a significant positive impact whereas in the random effect model, the impact of *DC* is found to be not significant. The F-statistic ensures that the coefficients in the model are different than zero.

Table 4.3. Fixed effect and random effect regression summary for *DC*, *LCC and DCxLCC*

```
> Fixed_DC_LCC_int <- plm(Log_of_pat_count_ny~DC+LCC+DCxLCC+PP, data=PANEL10, index=c("Org", "Year"), model="within")
> summary(Fixed_DC_LCC_int)
Oneway (individual) effect Within Model

Call:
plm(formula = Log_of_pat_count_ny ~ DC + LCC + DCxLCC + PP, data = PANEL10,
    model = "within", index = c("Org", "Year"))

Balanced Panel: n=63, T=18, N=1134

Residuals :
   Min. 1st Qu.  Median 3rd Qu.    Max.
-6.1600 -0.2940  0.0235  0.4300  3.1700

Coefficients :
       Estimate Std. Error t-value  Pr(>|t|)
DC      1.52236    0.54969  2.7695  0.005712 **
LCC     1.23617    0.17791  6.9481 6.427e-12 ***
DCxLCC -1.09510    0.66268 -1.6525  0.098721 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1323.4
Residual Sum of Squares: 1125.7
R-Squared      :  0.14935
      Adj. R-Squared :  0.14066
F-statistic: 62.5039 on 3 and 1068 DF, p-value: < 2.22e-16
```

```
> Random_DC_LCC_int <- plm(Log_of_pat_count_ny~DC+LCC+DCxLCC+PP, data=PANEL10, index=c("Org", "Year"), model="random")
> summary(Random_DC_LCC_int)
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Log_of_pat_count_ny ~ DC + LCC + DCxLCC + PP, data = PANEL10,
    model = "random", index = c("Org", "Year"))

Balanced Panel: n=63, T=18, N=1134

Effects:
                 var std.dev share
idiosyncratic 1.0541  1.0267 0.784
individual    0.2908  0.5393 0.216
theta:  0.5906

Residuals :
   Min. 1st Qu.  Median 3rd Qu.    Max.
-6.5100 -0.2960  0.0795  0.4900  2.5700

Coefficients :
             Estimate Std. Error t-value  Pr(>|t|)
(Intercept)  3.557897   0.159553 22.2991 < 2.2e-16 ***
DC           1.129093   0.471097  2.3967    0.0167 *
LCC          0.959368   0.161995  5.9222 4.211e-09 ***
DCxLCC      -0.716964   0.581637 -1.2327    0.2180
PP           0.149235   0.031895  4.6789 3.232e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1458.9
Residual Sum of Squares: 1209.2
R-Squared      :  0.17116
      Adj. R-Squared :  0.1704
F-statistic: 58.2844 on 4 and 1129 DF, p-value: < 2.22e-16
```

The two regression model results of the model *CC, LCC and CCxLCC* using *FE* and *RE* models are provided in Table 4.4. In the fixed effect model, both *CC* and *LCC* are found to have a significant positive impact whereas in the random effect model, the impact of L*CC* is found to be in the borderlines of accepted significance. An interesting results, perhaps different thatn the previous two outputs is that the interaction effect (*CCxLCC*) is found to have a significant negative impact. Another interesting result is that the coefficient value of *CC* is quite high. For every unit of increase in the normalized CC value, the expected increase in the log of *Patents* is almost at the level of 4. The F-statistics in both models ensure that the coefficients in the model are different than zero.

Although both *FE* and *RE* models indicated similar outputs, the variations between the outputs are no negligible. The question is remaining for which one must be used to explain the impact level of the predictors. To decide between *FE* and *RE* models a "Hausman test" is applied for preferred model. It basically tests the fundamental difference between the two models: Are errors are correlated with the regressors? The null hypothesis is that the random effect is preferred, only a significant correlation result leads to the rejection of *RE* model. Table 4.5 shows the results of the "Hausman test" where fixed effect models is the preference in both *DC, LCC and DCxLCC* and *CC, LCC and CCxLCC* models.

As mentioned above, I can apply a further test to see if time fixed effects are needed. In a model where time factor is introduced, and the impact of predictors are controlled for. Each year's coefficient absorbs the heterogeneity do to time difference. Table 4.6 gives the results for time-fixed effects model. It is observable that the impact levels (the coefficient values and the significances) are intensified by controlling for years. The two tests, F test and Lagrange Multiplier tests that tells the necessity of the time-fixed effects (Table 4.7).

Table 4.4. Fixed effect and random effect regression summary for *CC*, *LCC and CCxLCC*

```
> Fixed_CC_LCC_int <- plm(Log_of_pat_count_ny~CC+LCC+CCxLCC+PP, data=PANEL10, index=c("Org", "Year"), model="within")
> summary(Fixed_CC_LCC_int)
Oneway (individual) effect Within Model

Call:
plm(formula = Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP, data = PANEL10,
    model = "within", index = c("Org", "Year"))

Balanced Panel: n=63, T=18, N=1134

Residuals :
   Min. 1st Qu.  Median 3rd Qu.    Max.
-6.1300 -0.2990  0.0189  0.4340  3.3600

Coefficients :
       Estimate Std. Error t-value  Pr(>|t|)
CC      3.94778    0.64225  6.1468 1.115e-09 ***
LCC     0.56487    0.22422  2.5193   0.01191 *
CCxLCC -1.68665    0.75882 -2.2227   0.02644 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1323.4
Residual Sum of Squares: 1056.9
R-Squared      :  0.20133
      Adj. R-Squared :  0.18961
F-statistic: 89.7394 on 3 and 1068 DF, p-value: < 2.22e-16
```

```
> Random_CC_LCC_int <- plm(Log_of_pat_count_ny~CC+LCC+CCxLCC+PP, data=PANEL10, index=c("Org", "Year"), model="random")
> summary(Random_CC_LCC_int)
Oneway (individual) effect Random Effect Model
   (Swamy-Arora's transformation)

Call:
plm(formula = Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP, data = PANEL10,
    model = "random", index = c("Org", "Year"))

Balanced Panel: n=63, T=18, N=1134

Effects:
                 var std.dev share
idiosyncratic 0.9896  0.9948 0.761
individual    0.3106  0.5573 0.239
theta:  0.6122

Residuals :
   Min. 1st Qu.  Median 3rd Qu.    Max.
-6.5200 -0.2990  0.0873  0.5020  2.7700

Coefficients :
             Estimate Std. Error t-value  Pr(>|t|)
(Intercept)  3.277219   0.172596 18.9878 < 2.2e-16 ***
CC           2.299520   0.567317  4.0533 5.395e-05 ***
LCC          0.381193   0.208508  1.8282   0.06778 .
CCxLCC      -0.550371   0.710784 -0.7743   0.43891
PP           0.129128   0.032912  3.9235 9.256e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1444.9
Residual Sum of Squares: 1169.3
R-Squared      :  0.19076
      Adj. R-Squared :  0.18992
F-statistic: 66.5347 on 4 and 1129 DF, p-value: < 2.22e-16
```

Table 4.5. Results of Hausman tests

```
> phtest(Fixed_DC_LCC_int, Random_DC_LCC_int)

 Hausman Test

data:  Log_of_pat_count_ny ~ DC + LCC + DCxLCC + PP
chisq = 22.602, df = 3, p-value = 4.888e-05
alternative hypothesis: one model is inconsistent

> phtest(Fixed_CC_LCC_int, Random_CC_LCC_int)

 Hausman Test

data:  Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP
chisq = 67.6254, df = 3, p-value = 1.376e-14
alternative hypothesis: one model is inconsistent
```

Table 4.6. Results for time-fixed effect model

```
> Fixed_CC_LCC_int.time = plm(Log_of_pat_count_ny~CC+LCC+CCxLCC+PP + factor(Year),  data=PANEL10, index=c("Org", "Year"), model="within")
> summary(Fixed_CC_LCC_int.time)
Oneway (individual) effect Within Model

Call:
plm(formula = Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP +
    factor(Year), data = PANEL10, model = "within", index = c("Org",
    "Year"))

Balanced Panel: n=63, T=18, N=1134

Residuals :
   Min. 1st Qu.  Median 3rd Qu.    Max.
 -6.040  -0.330  -0.010   0.438   2.990

Coefficients :
                  Estimate Std. Error t-value  Pr(>|t|)
CC                3.726289   0.643795  5.7880 9.395e-09 ***
LCC               0.699350   0.223515  3.1289 0.0018031 **
CCxLCC           -2.656558   0.772015 -3.4411 0.0006022 ***
factor(Year)1995  0.022047   0.177900  0.1239 0.9013968
factor(Year)1996  0.120442   0.180001  0.6691 0.5035658
factor(Year)1997  0.486323   0.181815  2.6748 0.0075931 **
factor(Year)1998  0.526199   0.186174  2.8264 0.0047969 **
factor(Year)1999  0.637723   0.190734  3.3435 0.0008564 ***
factor(Year)2000  0.699303   0.196522  3.5584 0.0003898 ***
factor(Year)2001  0.716831   0.197365  3.6320 0.0002948 ***
factor(Year)2002  0.758820   0.198043  3.8316 0.0001349 ***
factor(Year)2003  0.874568   0.198130  4.4141 1.119e-05 ***
factor(Year)2004  0.775348   0.199718  3.8822 0.0001100 ***
factor(Year)2005  0.826148   0.211470  3.9067 9.955e-05 ***
factor(Year)2006  0.752281   0.205861  3.6543 0.0002706 ***
factor(Year)2007  0.702966   0.206148  3.4100 0.0006743 ***
factor(Year)2008  0.698692   0.204318  3.4196 0.0006511 ***
factor(Year)2009  0.881800   0.203344  4.3365 1.587e-05 ***
factor(Year)2010  0.727800   0.204107  3.5658 0.0003791 ***
factor(Year)2011  0.642847   0.209254  3.0721 0.0021804 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares:    1323.4
Residual Sum of Squares: 1013.1
R-Squared      :  0.23445
      Adj. R-Squared :  0.21729
F-statistic: 16.093 on 20 and 1051 DF, p-value: < 2.22e-16
```

Table 4.7. Tests for time-fixed effect model

```
> pFtest(Fixed_CC_LCC_int.time, Fixed_CC_LCC_int)

 F test for individual effects

data:  Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP + factor(Year)
F = NaN, df1 = 0, df2 = 1051, p-value = NA
alternative hypothesis: significant effects

> plmtest(Fixed_CC_LCC_int, c("time"), type=("bp"))

 Lagrange Multiplier Test - time effects (Breusch-Pagan)

data:  Log_of_pat_count_ny ~ CC + LCC + CCxLCC + PP + factor(Year)
chisq = 9.1452, df = 1, p-value = 0.002494
alternative hypothesis: significant effects
```

Significant effects from the F test and Lagrange Multiplier tests tells that it is necessary to use the time-fixed effects.

4.4.4. Robustness for Tie Weight and $Patent_{it+2}$, $Patent_{it+3}$ Models

As the last piece of analyses, the results are re-estimated using different method of calculation for the predictor variables and different models.

Although it was hypothesized that the calculation methods to moderate the impact of the predictor variables, the correlation results of the Table 4.8 does not prove the claim. The correlation values in Table 4.8 are calculated by using un-weighted ties (binary method) whereas table whereas the correlation values are calculated by using weighted ties in Table 4.2.

Table 4.8. Maximum and minimum correlation values using non-weighted calculation method

	$Patent_{it+1}$	DC_{it}	CC_{it}	LCC_{it}
$Patent_{it+1}$	1			
DC_{it}	0.508 $DC,_{2005}vsP,_{2006}$ 0.012 $DC,_{2011}vsP,_{2012}$	1		
CC_{it}	0.504 $CC,_{2005}vsP,_{2006}$ 0.035 $DC,_{2011}vsP,_{2012}$	0.891 $CC,_{2002}vsDC,_{2002}$ 0.847 $CC,_{2004}vsDC,_{2004}$	1	
LCC_{it}	0.301 $LCC,_{1995}vsP,_{1996}$ -0.152 $LCC,_{2005}vsP,_{2006}$	0.361 $LCC,_{1994}vsDC,_{1995}$ 0.108 $CC,_{2009}vsDC,_{2009}$	0.661 $LCC,_{1994}vsDC,_{1995}$ 0.307 $LCC,_{2009}vsDC,_{2010}$	1

All the previous analysis (longitudinal plots, heterogeneity and regression analyses) are applied using the calculated predictors (*DC, CC and LCC*) that adopts the approach taking into account weights of connections. Interesting enough, predictor variable values from both calculation methods are so close that they did not affect the model estimation results significantly. In fact, the overall correlation coefficient between the values of Table 4.2 and Table 4.7 is calculated as r=0.952. Therefore, I did not present the other results with the non-weighted predictor values as they are very similar with the current calculations.

4.5. Discussion and Conclusions

This study analyzes the impact of the degree centrality, closeness centrality and local clustering coefficient of an alliance network of research and development joint ventures on the network members' invention performances. Research and Production Joint Venture (RJV) alliance data based on the notices of National Cooperative Research and Production Act of 1993

(NCRPA) which provides information regarding which organization has been a member of an RJV. Memberships in RJV are translated to the connectivity among the members of the RJV and subsequently, relationship maps are obtained.

It has been argued that high quantities of opportunities for technological knowledge inflow availability (*DC*) should facilitate greater invention counts. Similarly, the availability of diverse information in a close range (*CC*) must also facilitate greater invention counts. There are two competing arguments regarding locally dense pockets (*LCC*), which is sometimes seen as redundant hindering the novel knowledge creation but some studies suggested its ability for transmission capacity increase. Its facilitation of technological knowledge inflow and invention counts is also tested. Both the correlational analyses and regression model results indicated strong support for the impact of initial two arguments. The impact of clustering has also been visible but only after controlling for time variable, which is found to be the most viable model among all. These results are consistent with much of the theory developed in recent literature (1, 33). They are consistent with Schilling and Phelp's argument that combination of clustering and reach (another measure for the availability of diverse information in a close range) is associated with significantly higher invention performance. The results also have similar indications with Ahuja's findings regarding the negative interaction between the direct and indirect ties.

Another question of interest was the calculation method of these network positional metrics. One of the two main methods of calculation for all three metrics is using a binary approach and count only one tie even though there are more than bilateral connections. The other method is taking as many bilateral connections as there exist. Although my argument was that calculation of the two different methods arises significant discrepancies, an extremely high correlation is observed in both calculation methods, which did not affect the results significantly.

This result, however, indicates a contradictory picture with Opsahl's (27) examples of significant changes due to different calculation methods. The underlying conditions that lead to very similar outcomes calls for a more detailed analysis.

Along with the methodological contributions, this study also investigated important information like the number of alliances an organization is involved and duration of membership in alliances. The results add to this literature by suggesting that membership to successful research and development joint ventures last a little more than 10 years on the average based on the information from NCRPA database. This research also made visible the historical trends regarding alliance membership from 1994 to 2012.

This research also have several limitations. The findings may be influenced by the assumption that RJVs announced at Federal Register are representative of all existing RJVs. The results might have been biased if alliances in practice are more sparse, lasing shorter times. A strong assumption that follows is the equal treatment of all alliances. Different types of RJVs and different types of bilateral communications relationships may lead to significantly variant levels of facilitation for technological knowledge sharing. Although the strength is defined by the number of common memberships, I did not address the governance structure and the scope of the relationships. Although the sample that I worked exhibits high connectivity and represent the core of high invention performance organizations, it still constitutes a very small fraction of the existing organizations. Each of these limitations can be considered exciting areas for future research endeavor.

4.6. References

1. Ahuja G. Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. Administrative Science Quarterly 45(3): 425-455; 2000.

2. Audretsch D.B., Feldman M.P. Innovative Clusters and the Industry Life Cycle. Review of Industrial Organization 11: 253-273; 1996.

3. Basberg B.L. Patents and the measurement of technological change: A survey of the literature. Research Policy 16; 131-141; 1987.

4. Beaudry C., Schiffauerova A. Impacts of collaboration and network indicators on patent quality: The case of Canadian nanotechnology innovation. European Management Journal 29, 362– 376; 2011.

5. Bleeke, J.; Ernst, D. Is Your Strategic Alliance Really a Sale? Harvard Business Review 73(1): 97–105; 1995.

6. Blundell, R. R., R. Griffith, J. Van Reenen. Dynamic count data models of technological innovation. Econom. J. 105 333–344; 1995.

7. Borgatti, S. P.; Everett M. G.; Freeman L. C. UCINET for Windows: Software for social network analysis. Harvard, Analytic Technologies; 2002.

8. Burt R. S. Structural Holes: The Social Structure of Competition; Cambridge, MA: Harvard University Press; 1992.

9. Caliskan F., Reeves K.A., Ozcan O., Zeller D. The impact of collaboration network Position on innovative performance. Technology and Innovation, Vol. 14, pp. 351–363, 2012.

10. Chiu Y.T.H. How network competence and network location influence innovation performance. Journal of Business & Industrial Marketing. 24/1 46–55; 2009.

11. Coleman, J. S. Social capital in the creation of human capital. Amer. J. Sociol. 94 S95–S120; 1988.

12. Cowan, R.; Jonard, N.; Ozman, M. Knowledge dynamics in a network industry. Technological Forecasting & Social Change 71:469-484; 2004.

13. Cowan, R.; Jonard, N.; Zimmermann, J. Bilateral collaboraton and the emergence of innovation networks. Management Science 53(7):1051-1067; 2007.

14. Cowan R., Jonard N. Knowledge portfolios and the Organization of innovation networks Academy of Management Review. Vol. 34, No. 2, 320–342; 2009.

15. Debackere K., Clarysse B., Rappa M. Dismantling the Ivory Tower: The Influence of Networks on Innovative Output in Emerging Technologies. Technological Forecasting and Social Change 53, 139-154; 1996.

16. Federal Register, Articles Search Site. Retrieved in October 2013 from https://www.federalregister.gov/articles/search.; 2013.

17. Fritsch M., Kauffeld-Monz M., The impact of network structure on knowledge transfer: an application of social network analysis in the context of regional innovation networks. Ann Reg Sci. 44:21–38; 2010.

18. Granovetter, M.S. The strength of weak ties. American Journal of Sociology 78, 1360-1380; 1973.

19. Griliches Z., Patent Statistics as Economic Indicators: A Survey. (p. 287 - 343) In: R&D and Productivity: The Econometric Evidence. University of Chicago Press; January 1998.

20. Gulati R. Social Structure and Alliance Formation Patterns: A Longitudinal Analysis Administrative Science Quarterly, 40: 619-652; 1995.

21. Hausman, J., B. Hall, Z. Griliches. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52 909–938; 1984.

22. Iandoli, L.; Marchione. E.; Ponsiglione, C.; Zollo, G. Knowledge exchange processes in industrial districts and the emergence of networks. Cejor 20:231-250; 2012.

23. Kogut, B. The Stability of Joint Ventures: Reciprocity and Competitive Rivalry. Journal of Industrial Economics 38(2): 183–198; 1989.

24. Lovejoy, W. S.; Sinha, A. Efficient structures for innovative social networks. Management Science 1-9; 2010.

25. Monthe J.D.L., Link. A., Networks, Alliances and Partnerships in the Innovation Process. Norwell, MA: Kluwer Academic Publishers Group; 2002.

26. Noteboom, B., Haverbeke W.V., Duysters G., Gilsing V., Oord A. Optimal cognitive distance and absorptive capacity. Research Policy 36 1016–1034; 2007.

27. Opsahl, T., Agneessens, F., Skvoretz, J. Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32 (3), 245-251; 2010.

28. Owen-Smith J., Powell W.W., Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. Organization Science: Vol. 15, No. 1, pp. 5–21; January–February 2004.

29. Padula G. Enhancing the Innovation Performance of Firms by Balancing Cohesiveness and Bridging Ties. Long Range Planning 41. 395e419; 2008.

30. Powell W.W. Kenneth W. Koput and Laurel Smith-Doerr Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. Administrative Science Quarterly, Vol. 41, No. 1; pp. 116-145; March 1996.

31. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

32. Sabidussi, G. The centrality index of a graph. Psychometrica 31: 581-603; 1966.

33. Schilling, M. A. Phelps C. C., Interfirm collaboration networks: The Impact of large-scale network structure on firm innovation. Management Science 53(7):1113-1126; 2007.

34. Stuart T.E. Network Positions and Propensities to Collaborate: An Investigation of Strategic Alliance Formation in a High-Technology Industry. Administrative Science Quarterly, Vol. 43, No. 3; pp. 668-698 (Sep., 1998).

35. The US Department of Justice: Filing a Notification Under the NCRPA. Retrieved from December 2012 to October 2013: http://www.justice.gov/atr/public/guidelines/ncrpa.html.; 2013.

36. Torres-Reyna, O. Panel Data Analysis: Fixed and Random Effects. Princeton University Data and Statistical services. Retrieved in October 2013 from http://www.princeton.edu/~otorres/Panel101.pdf; 2013.

37. Tsai W. Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance. The Academy of Management Journal 44(5): 996-1004; 2001.

38. USPTO Patent Full Text and Image Database. Explanation of Data. Retrieved in October 2013 from http://www.uspto.gov/web/offices/ac/ido/oeip/taf/stcasga/explan_stcorg.htm; 2013.

39. USPTO Patent Full Text and Image Database. Patent Statistics by Organization. Retrieved October 2013 from http://www.uspto.gov/web/offices/ac/ido/oeip/taf/stcasga/usa_stcorg.htm; 2013.

40. Wasserman, S.; Faust K., Social Network Analysis: Methods and Applications. New York, NY: Cambridge University Press; 1994.

CHAPTER 5: CONCLUSION AND CONTRIBUTION

The two fundamental questions that inspired this dissertation research were: 1) What kind of strategies allow organizations to attain the required effective invention results in a relationship network? 2) What is the most effective structural positioning in such a network of relationships? Three independent studies were employed to answer these questions using three different methodologies.

The simulation study sought to explore the effectiveness of three different collaboration strategies. This research not only indicated the most effective and consistent strategies, but also indicated where they are most useful both in terms of changing environments of technological dynamism and in terms of invention type developed in the context of collaboration. These exploratory results bring unique contributions to both companies and governments motivated to collaborate for the purpose of solving industrial problems and bring solutions to society. Besides its broader impact, it also makes methodological contributions. The simulation is robust in terms of allowing multilateral collaborations between the collaborating agents. It also allows for various numbers of combinations in the strategies of the members. Perhaps more importantly, it is open for further development in order to mimic the practical behavior of collaborating organizations.

The survey study was administered within a regional economy across all industries. It provided cross sectional data from the most inventive organizations in the State of Florida. As the second research question states, the impact of positional metrics on invention performance

was investigated. A methodological contribution of this study is the introduction of five different relationship types in the construction of the network map. Although very sensitive to the network data, the results give insight regarding what type of positioning is associated with higher effectiveness in invention performance. The two prominent metrics, degree and closeness centralities, are found to have the highest impact are being close in terms of network distances and being in collaboration with highest performers.

The Federal Register database that announces notices of National Cooperative Research and Production Act provides a unique opportunity to construct the alliance relationship network of organizations since 1994. The longitudinal data provided information for an array of analyses from the investigation of research joint ventures to the network positional impacts on inventive performance. Similar to the previous study, closeness centrality is found to be highly and significantly associated with the invention performance of organizations. In future research, it may be interesting to investigate the ways of reducing distances in between the members of the network to access technological knowledge. It is also important that local clustering is found to impact the invention performance, which corroborates the supporters of the benefit of local clustering rather than those who emphasize its negative impact due to its redundancy. The interaction of closeness and clustering, however, is found to have a significant negative impact, which calls for further investigation.

It is true that sometimes inventions become the subject of competition. However, solutions to our health problems, communication or transportation needs etc. depend on how well we recombine the existing and enhancing technological knowledge. Inventions are not usually the product of fortunate events. They are the product of an effective process for the recombination of technological knowledge. Also, effective invention performance is not only

important for organizations individually, it is crucial for governments that are concerned with the problems of society. In a broader sense, solutions to many problems can be as easy as the utilization of technological knowledge in a network of relationships or utilization of untapped collaboration opportunities. In this research, effective ways to facilitate recombination of technological knowledge is addressed and presented to inform both companies and policy makers. Better understanding on the dynamics of the invention process may bring better solutions to our existing problems.

APPENDICES

Appendix A: R Code for the Simulation of Collaborative Invention Networks

```
    ### COLLABORATIVE INNOVATION NETWORKS - SIMULATION MODEL ###

## Start with:

install.packages(ggplot2); require(ggplot2)
install.packages(ggplot) ; require(ggplot)
install.packages(igraph) ; require(igraph)
install.packages(tnet)   ; require(tnet)
install.packages(MASS)   ; require(MASS)
install.packages(lmtest) ; require(lmtest)
install.packages(gplots) ; require(gplots)

MAIN.TABLE = data.frame(matrix(0, nrow=18, ncol=12))
for (i in seq(1, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "CC"
for (i in seq(2, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "RC"
for (i in seq(3, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "SC"

# THE MAIN TABLE IS BEING FILLED NOW
for (q in c(0, 0.2, 0.4, 0.6, 0.8, 1))
 for (r in c(0, 0.2, 0.4, 0.6, 0.8, 1)) {

# Replications of experiments
Performance.table = NULL
for (n in 1:10) {

#Global parameters
N=24  # Number of organizations (!! must be multiple of 3, for now !!)
k=5   # Technological knowledge elements (the last two elements of the CCs are initialized separately, in case of a parameter change)
K=200 # Initial (normal random) knowledge endowment average. Std. dev= K/5.
alpha = q # Proposal acceptance rate. It is a binomial probability (number of accepted connections are binomially distributed)
lambda=0.5 # (1-lambda) is termination rate. Number of continuing connections are binomially distributed
ph=2  # The number of projects that an organization proposes when the latter is highly desirable to work with
pl=1  # The number of projects that an organization proposes when the latter is moderately desirable to work with
cyc = 20 # Number of cycles (terms) each replication will run
beta = r  # The difference level at which the knowledge complementariness is awarded (pays off for invention)

#This simulation consists of 6 stages:
# 1) Initialization of agents' states and connection matrix
# 2) Partnership proposal (preparation for joint venturing)
# 3) Response to proposal (formation of network connections)
# 4) Invention count increase
# 5) Technological Knowledge increase
# 6) Presentation (only when necessary)

# Functions
is.wholenumber <- function(x, tol = .Machine$double.eps^0.5)  abs(x - round(x)) < tol

# ==================================== 1) Initialization Stage ====================================

#List of agents in network simulation: CC: Cognitive Cooperator, RC: Relational Cooperator, SC: Structural Cooperator.
agent.list = NULL
for (i in (N/N):(N/3)) {  agent.list = c(agent.list, i) ; agent.list[i] = sub("^", "CC", i) }
```

```
for (i in ((N/3)+1):(N/(3/2))) { agent.list = c(agent.list, i) ; agent.list[i] = sub("^", "RC", (i-N/3)) }
for (i in ((N/(3/2))+1):N) { agent.list = c(agent.list, i) ; agent.list[i] = sub("^", "SC", (i-(N/(3/2)))) }

#Assign each agent's ID, strategy and initialize the values of knowledge elements and invention counts
Agents.state = NULL
for (i in (N/N):(N/3)) {
  Agents.state <- cbind(Agents.state, c(ID=i, Str.type=1, Knowledge=round(runif(k, 0, K)), InventionCount=sample(1000000,1)))
}
for (i in ((N/3)+1):(N/(3/2))) {
  Agents.state <- cbind(Agents.state, c(ID=i, Str.type=2, Knowledge=round(runif(k, 0, K)), InventionCount=sample(1000000,1)))
}
for (i in ((N/(3/2))+1):N) {
  Agents.state <- cbind(Agents.state, c(ID=i, Str.type=3, Knowledge=round(runif(k, 0, K)), InventionCount=sample(1000000,1)))
}

colnames(Agents.state) <- agent.list
for (i in 1:k+3)
 for (j in 1:N)
if (Agents.state[i,j] < 0) Agents.state[i,j] = 0

#A random initial Network Matrix and a zero Proposal Matrix are created

Initial.N.Matrix <- matrix(0, nrow = length(agent.list), ncol = length(agent.list))
for (i in 1:N) {
 Initial.N.Matrix[i,] <- sample(c(2, 1 , sample(0, N-2, replace=TRUE)))
 while (Initial.N.Matrix[i,i] != 0) Initial.N.Matrix[i,] <- sample(c(2, 1 , sample(0, N-2, replace=TRUE)))
}
colnames(Initial.N.Matrix) <- agent.list
Current.N.Matrix <- Initial.N.Matrix
Proposal.Matrix <- matrix(0, nrow = length(agent.list), ncol = length(agent.list))
colnames(Proposal.Matrix) <- agent.list

# ================================= End of Initialization Stage =================================

# Some performance variables are initialized here

Mean.perf.CC = rep(0, cyc) ; Mean.perf.RC = rep(0, cyc) ; Mean.perf.SC = rep(0, cyc)


#Let's run for cyc number of cycles
for (m in 1:cyc) {

# There are 3 stages in this cycle: 2)Partnership proposal, 3) Response to Proposal, 4) Inventions

# ============================== 2) Partnership Proposals Stage ================================

# Step-1: CCs identify their worst knowledge element, i.e., smallest in value
Worst.elements = NULL
for (i in (N/N):(N/3)) {
 Worst.elements <- cbind(Worst.elements, c(order(Agents.state[3:(k+2),i])[1:2]))
}
colnames(Worst.elements) <- agent.list[(N/N):(N/3)]

# Step-2: CCs want to know who has the best for their worst knowledge element, i.e., highest in value
Best.partners = NULL
for (i in (N/N):(N/3)) {
```

```r
 Best.partners <- cbind(Best.partners,
 c(if (order(Agents.state[Worst.elements[1,i]+2,-i])[N-1] >= i) order(Agents.state[Worst.elements[1,i]+2,-i])[N-1]+1
   else order(Agents.state[Worst.elements[1,i]+2,-i])[N-1],

   if (order(Agents.state[Worst.elements[2,i]+2,-i])[N-1] >= i) order(Agents.state[Worst.elements[2,i]+2,-i])[N-1]+1
   else order(Agents.state[Worst.elements[2,i]+2,-i])[N-1]))
}

# Step-3: CCs' proposals are temporarily recorded. But first we need to set all entries to zero for the second cycle and on.
Proposal.Matrix[,] = 0
for (i in (N/N):(N/3)) {
 Proposal.Matrix[i,Best.partners[1,i]] <- ph
 Proposal.Matrix[i,Best.partners[2,i]] <- pl
}

# Step-4: RCs send proposals only to their current connections
Trusted.partners = NULL
for (i in ((N/3)+1):(N/(3/2))) {
                    Trusted.indices = NULL
                    Trusted.indices = which(Current.N.Matrix[i,]==max(Current.N.Matrix[i,]))
 if  (length(Trusted.indices) > 2)
     Trusted.partners <- rbind(Trusted.partners, c(sample(order(Current.N.Matrix[i,])[N:(1+N-length(Trusted.indices))])[1:2]))

 if  (length(Trusted.indices) == 2)    # the two members have the same number of ties - not to confuse with different number of ties
     Trusted.partners <- rbind(Trusted.partners, c(sample(order(Current.N.Matrix[i,])[N:(N-1)])))

 if  ((length(Trusted.indices) == 1) & (length(which(Current.N.Matrix[i,][-Trusted.indices] == max(Current.N.Matrix[i,][-Trusted.indices]))) == (N-
1)) )
     Trusted.partners <- rbind(Trusted.partners, c(order(Current.N.Matrix[i,])[N], order(Current.N.Matrix[i,])[N] ))

 if  ((length(Trusted.indices) == 1) & (length(which(Current.N.Matrix[i,][-Trusted.indices] == max(Current.N.Matrix[i,][-Trusted.indices]))) != (N-
1)) )
    Trusted.partners <- rbind(Trusted.partners, c(order(Current.N.Matrix[i,])[N], order(Current.N.Matrix[i,])[N-1]))

 if (Trusted.partners[i+1-((N/3)+1),1] == Trusted.partners[i+1-((N/3)+1),2])
    Proposal.Matrix[i,Trusted.partners[i+1-((N/3)+1),1]] <- ph+pl
 else {
   Proposal.Matrix[i,Trusted.partners[i+1-((N/3)+1),1]] <- ph
   Proposal.Matrix[i,Trusted.partners[i+1-((N/3)+1),2]] <- pl   }
}

# Step-5: SCs want to know who has currently the highest number of connections: order(colSums(Initial.N.Matrix))[N]
#       OR
#       *** SCs also want to know who has the highest number of invention count: order(Agents.state[k+3,])[N]
#       SCs' proposals are temporarily recorded

#for (i in ((N/(3/2))+1):N) {
#  Proposal.Matrix[i,order(colSums(rbind(colSums(Current.N.Matrix), rowSums(Current.N.Matrix))))[N]] <- 3
#  if (Proposal.Matrix[i,i] != 0) { Proposal.Matrix[i,i] = 0
#                              Proposal.Matrix[i,order(colSums(rbind(colSums(Current.N.Matrix), rowSums(Current.N.Matrix))))[N-1]] <- 3
#                             }
#   else Proposal.Matrix[i,order(colSums(rbind(colSums(Current.N.Matrix), rowSums(Current.N.Matrix))))[N-1]] <- 1
#      if (Proposal.Matrix[i,i] != 0) { Proposal.Matrix[i,i] = 0
#                                  Proposal.Matrix[i,order(colSums(rbind(colSums(Current.N.Matrix), rowSums(Current.N.Matrix))))[N-2]] <- 1
#                                 }
#}
```

```
for ( i in ((N/(3/2))+1):N) {
 Proposal.Matrix[i, order(Agents.state[k+3,])[N]] <- ph
 if (Proposal.Matrix[i,i] != 0) { Proposal.Matrix[i,i] = 0 ; Proposal.Matrix[i, order(Agents.state[k+3,])[N-1]] <- ph }
 else Proposal.Matrix[i, order(Agents.state[k+3,])[N-1]] <- pl
  if (Proposal.Matrix[i,i] != 0) { Proposal.Matrix[i,i] = 0 ; Proposal.Matrix[i, order(Agents.state[k+3,])[N-2]] <- pl }
}

# ============================== End of Partnership Proposals Stage ==============================

# =============================== 3) Responsing to Proposals Stage ===============================

# Previously Proposal Matrix was finalized. Now 1) Each  proposal is accepted or rejected (or partially accepted).
#                                   2) Current Network Matrix is partially reduced (some ventures are terminated)
#                                   3) Responses are superimposed (added) to Current Network Matrix.

# Step-1:
Response.Matrix = Proposal.Matrix
for (i in 1:N) {
 for (j in 1:N) {
  t = Proposal.Matrix[i,j]
  Response.Matrix[i,j] = rbinom(1, t, alpha)
 }
}

# Step-2:
for (i in 1:N) {
 for (j in 1:N) {
  t = Current.N.Matrix[i,j]
  Current.N.Matrix[i,j] = rbinom(1, t, lambda)
 }
}

# Step-3:
Current.N.Matrix <- Current.N.Matrix + Response.Matrix

# ============================ End of Responsing to Proposals Stage ===============================


# Temporarily generated invention counts are set to 0 to avoid the impact of initial endowments (done only in the 1st cycle)
if (m==1) Agents.state[k+3,] = 0


# ============================= 4) Invention Counts Increase Stage ================================

# Now, practically, 1) More resources means more chances of innovation
#                   2) Better knowledge capacity means more chances of innovation
#                   3) Knowledge complementariness is hypothesized to provide advantage

# Step-1: Resources are translated into connections so (Degree centrality impact is taken into account)
Agents.state[k+3,] <- Agents.state[k+3,] + round(colSums(rbind(colSums(Current.N.Matrix), rowSums(Current.N.Matrix)))/2)

# Step-2: Organizations with knowledge elements at certain percentiles make inventions increasingly. Knowledge stock is multiplied by 1-alpha
parameter
for (i in 3:(k+2)) {
 Quantile5 <- quantile(Agents.state[i,], probs = seq(0, 1, 0.25), na.rm = FALSE, type = 3)
```

```r
 for (j in 1:N) {
        KS.increase=seq(0, 0, length.out = N)
        for (l in 1:5)
         if (Agents.state[i,j] >= Quantile5[l]) KS.increase[j] <- KS.increase[j] + 1
           Agents.state[k+3,j] <- Agents.state[k+3,j] + round((KS.increase[j]))
       }
}

# Step-3: Invention increase due to knowledge complementariness (complementariness impact is to be investigated)

for (i in 3:(k+2)) {
 Tech.range <- max(Agents.state[i,]) - min(Agents.state[i,])
 for (j in 1:N)
  for (l in 1:N) {
   k.check = NULL ;
   if (Current.N.Matrix[j,l] != 0) {
    for (p in 3:(k+2)) {
     k.check = c(k.check, Agents.state[p,j] > Agents.state[p,l])
            } ;
     if (length(unique(k.check)) > 1) {
       Ind.percentage <- abs(Agents.state[i,j] - Agents.state[i,l])/Tech.range
       Agents.state[k+3,l] <- Agents.state[k+3,l] + round((1 - abs(Ind.percentage-beta))*Current.N.Matrix[j,l])
       Agents.state[k+3,j] <- Agents.state[k+3,j] + round((1 - abs(Ind.percentage-beta))*Current.N.Matrix[j,l])
                }
}
}
}

# =========================== End of Invention Invention Counts Increase Stage ========================

# ============================ 5) Technological Knowledge Increase Stage ==========================

# Step-1: Increase due to partnerships: The organization who has a lower knowledge element benefits
Max.coll = max(Current.N.Matrix[,])
for (i in 3:(k+2)) {
 for (j in 1:N)
  for (l in 1:N)
   if ((Current.N.Matrix[j,l] != 0) & (Agents.state[i,l] - Agents.state[i,j] > 0))
     Agents.state[i,j] <- Agents.state[i,j] + round((((Current.N.Matrix[j,l]/Max.coll)*(Agents.state[i,l] - Agents.state[i,j]))/2)
   if ((Current.N.Matrix[j,l] != 0) & (Agents.state[i,l] - Agents.state[i,j] < 0))
     Agents.state[i,l] <- Agents.state[i,l] + round((((Current.N.Matrix[j,l]/Max.coll)*(Agents.state[i,j] - Agents.state[i,l]))/2)
}

# Step-2: Inventions cause knowledge increase too (NOT ACCOMPLISHED YET: LIMITATION)

# Step-3: Increase/decrease due to other factors. Assumed uniformly distributed with mean 0 and std.dev=K/..

for (i in 3:(k+2))
 for (j in 1:N) {
  Agents.state[i,j] <- Agents.state[i,j] + round(runif(1, -K/2, K/2))
  if (Agents.state[i,j] < 0) Agents.state[i,j] = 0
        }

# ======================== End of Technological Knowledge Increase Stage ===========================

# =============================== 6) Presentation Stage ====================================
```

```
# Longitudinal plot variables created Mean/SD

Mean.perf.CC[m] = round(mean(Agents.state[k+3,(N/N):(N/3)]))
Mean.perf.RC[m] = round(mean(Agents.state[k+3,((N/3)+1):(N/(3/2))]))
Mean.perf.SC[m] = round(mean(Agents.state[k+3,((N/(3/2))+1):N]))

#SD.perf.CC[m] = round(sd(Agents.state[k+3,(N/N):(N/3)]))
#SD.perf.RC[m] = round(sd(Agents.state[k+3,((N/3)+1):(N/(3/2))]))
#SD.perf.SC[m] = round(sd(Agents.state[k+3,((N/(3/2))+1):N]))

# Presantation at the end of some cycles


if (is.wholenumber(m/571)) {

# Presentation of the network

 g1 <- graph.adjacency(Current.N.Matrix)
 par(mfrow=c(1,2))
 plot(g1, layout=layout.auto, vertex.size=4,
     vertex.label.dist=0, vertex.label.degree=pi, vertex.color="red", edge.arrow.size=0.01)

# Presentation of the invention counts table

 CCinventions = NULL; RCinventions = NULL; SCinventions = NULL
 for (i in (N/N):(N/3)) CCinventions <- rbind(CCinventions, Agents.state[k+3,i])
 for (i in ((N/3)+1):(N/(3/2))) RCinventions <- rbind(RCinventions, Agents.state[k+3,i])
 for (i in ((N/(3/2))+1): N) SCinventions <- rbind(SCinventions, Agents.state[k+3,i])
 Matrix.represantation = NULL
 Matrix.represantation <- cbind(CCinventions, RCinventions, SCinventions)
 colnames(Matrix.represantation) <- c("CC", "RC", "SC")
 Rows.names = NULL; for (i in 1:(N/3)) Rows.names <- cbind(Rows.names, i)
 rownames(Matrix.represantation) <- Rows.names
 textplot(Matrix.represantation, cex=1.0, valign="top"); title("Invention Counts")
}

# Cycles Mean/SD Plot

if (is.wholenumber(m/571)) {

par(mfrow=c(1,2))
g_range <- range(0, Mean.perf.CC, Mean.perf.RC, Mean.perf.SC)

plot(Mean.perf.CC, type="o", col="blue", ylim=g_range,
   axes=TRUE, ann=FALSE)
lines(Mean.perf.RC, type="o", pch=22, lty=2, col="red")
lines(Mean.perf.SC, type="o", pch=23, lty=3, col="green")

title(xlab="Cycles")
title(ylab="Mean invention count")

legend(1, g_range[2], c("CCinventions","RCinventions","SCinventions"), cex=0.8,
    col=c("blue","red","green"), pch=21:23, lty=1:3);

title(main="Average Performance of Strategy Groups", col.main="black", font.main=2)
```

```
}

} # end of cyc count

#graph.density(g1)
#centralization.degree(g1)

# Presentation of winners table in each replication Mean/SD

par(mfrow=c(1,1))

Winner=NULL
if (max(Mean.perf.CC[m], Mean.perf.RC[m], Mean.perf.SC[m]) == Mean.perf.CC[m]) Winner=c("CC")
if (max(Mean.perf.CC[m], Mean.perf.RC[m], Mean.perf.SC[m]) == Mean.perf.RC[m]) Winner=c("RC")
if (max(Mean.perf.CC[m], Mean.perf.RC[m], Mean.perf.SC[m]) == Mean.perf.SC[m]) Winner=c("SC")
Performance.table = rbind(Performance.table, cbind(Mean.perf.CC[m], Mean.perf.RC[m], Mean.perf.SC[m], Winner))
colnames(Performance.table) = c("CC", "RC", "SC", "Winner")
textplot(Performance.table, cex=1.0, valign="top") ; title("Leaders at Mean performances over 10 replications")

#if (min(SD.perf.CC[m], SD.perf.RC[m], SD.perf.SC[m]) == SD.perf.CC[m]) Winner=c("CC")
#if (min(SD.perf.CC[m], SD.perf.RC[m], SD.perf.SC[m]) == SD.perf.RC[m]) Winner=c("RC")
#if (min(SD.perf.CC[m], SD.perf.RC[m], SD.perf.SC[m]) == SD.perf.SC[m]) Winner=c("SC")
#Performance.table = rbind(Performance.table, cbind(Mean.perf.CC[m], Mean.perf.RC[m], Mean.perf.SC[m], Winner))
#colnames(Performance.table) = c("CC", "RC", "SC", "The Smallest SD")
#textplot(Performance.table, cex=1.0, valign="top") ; title("Standsrd deviation performances over 10 replications")

} # end of 10 replications set

if (r==0) MTrow=16; if (r==0.2) MTrow=13; if (r==0.4) MTrow=10; if (r==0.6) MTrow=7; if (r==0.8) MTrow=4; if (r==1) MTrow=1;
MTcol = (q+.1)*10

MAIN.TABLE[MTrow:((MTrow-1)+nrow(as.data.frame(table(factor(Performance.table[,4], lev=c("CC", "RC", "SC")))))),MTcol:(MTcol+1)] =
as.data.frame(table(factor(Performance.table[,4], lev=c("CC", "RC", "SC"))))

for (i in seq(1, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "CC"
for (i in seq(2, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "RC"
for (i in seq(3, 18, 3)) for (j in seq(1, 12, 2)) MAIN.TABLE[i,j] = "SC"

textplot(MAIN.TABLE, cex=1.0, valign="top") ; title("Mean invention performance (replication winners) at various alpha vs. beta")

#textplot(MAIN.TABLE, cex=1.0, valign="top") ; title("The SD performances at various alpha vs. beta")

write.csv(MAIN.TABLE, "MAIN.TABLE.csv")
} # end of Main Table construction

#require(xlsx)
write.csv(MAIN.TABLE, "MAIN.TABLE.csv")

# ================================== End of Presentations Stage ==============================

# ================================== END OF SIMULATION ===============================
```

Appendix B: Image of the Written Permission for Published Portions of Chapter 3

lmiranda@cognizantcommunication.com Oct 14

to me, Christine

Dear Fethullah:

We can grant permission for one-time use of the requested information. Please note that proper reference to the original source of publication is required. This should include listing of the DOI and publisher reference in addition to volume number and page location. **The article(s) and/or figure(s) cannot be used for resale.**

Thank you for contacting us. Please let me know if I can further assist in any way.

Best regards,

-----Original Message-----
From: "Christine Stagg" <cstagg@twcny.rr.com>
Sent: Monday, October 14, 2013 7:39am
To: "Bob Miranda" <Cogcomm@aol.com>, "Lori Miranda" <lmiranda@cognizantcommunication.com>
Subject: Fwd: Permission

Permission.eml
31K View Download

Fethullah Caliskan <fcaliska@mail.usf.edu> Oct 15

to lmiranda

Sure, I will include the proper referencing. Thank you!

Appendix C: The Complete List of Organizations Used in the Analyses of Chapter 2

Table A1. The complete list of organizations used in the analyses of Chapter 2

1	IBM Int. Business Machines
2	Hewlett-Packard & Hp Dev.
3	Intel Corporation
4	Micron Technology, Inc.
5	Microsoft Corporation
6	General Electric Company
7	Motorola, Inc.
8	Texas Instruments, Inc.
9	Eastman Kodak Company
10	Xerox Corporation
11	Lucent Inc. (Alcatel After 2006)
12	AT&T Corp.
13	Advanced Micro Devices, Inc.
14	General Motors
15	3m Minnesota Mining Manuf.
16	Cisco Technology, Inc.
17	Honeywell International Inc.
18	Sun Micr. Inc. (Oracle After 2009)
19	Ford Motors
20	University Of California
21	Broadcom Corporation
22	Boeing Company
23	E. I. Du Pont De Nemours And Co.
24	Qualcomm, Inc.
25	United States Of America, Navy
26	Applied Materials, Inc.
27	Siemens
28	Procter & Gamble Company
29	Apple, Inc.
30	Delphi Technologies, Inc.
31	Seagate Technology, Llc
32	Exxon (Mobil, Chemical, All)
33	Oracle International Corp.
34	Caterpillar Inc.
35	Lockheed Martin Corporation
36	Philips Electronics
37	Medtronic Inc.
38	LSI Logic Corporation
39	Kimberly-Clark Worldwide, Inc.
40	Agilent Technologies, Inc.

Table A1. (Continued)

41	National Semiconductor Corp.
42	Sony Corporation
43	Raytheon Company
44	Halliburton Energy Services, Inc.
45	United Technologies Corporation
46	Xilinx, Inc.
47	Baker Hughes Incorporated
48	Illinois Tool Works Inc.
49	United States Of America, Army
50	Hitachi
51	Schlumberger Technology Corp.
52	Eaton Corporation
53	Freescale Semiconductor
54	Corning Incorporated
55	Emc Corporation
56	Massachusetts Institute Of Tech.
57	Dell Products, L.P.
58	Pioneer Hi-Bred International, Inc.
59	Northrop Grumman Corporation
60	Monsanto Technology, Llc
61	Chrysler Motors Corporation
62	Cardiac Pacemakers, Inc.
63	Usa, Dep. Of Health & Human Serv.

ABOUT THE AUTHOR

Fethullah Caliskan earned his BS in Industrial Engineering in 2001 and his MA in Human Resources Management from Marmara University in 2006. He worked as a management consultant at The Scientific and Technological Research Council of Turkey (TUBITAK) from 2001 to 2007. He earned his Masters in Industrial Engineering from University of South Florida (USF) and currently pursuing his Ph.D. His research interests include analysis of collaborative innovation models and social network analysis. He led a research team that received a $5,000 grant from USF Challenge Grant Program. He has two recognition awards from USF Provost Office for teaching in College of Engineering including Probability and Statistics for Engineers and Engineering Economics.