2007

The role of networks in labor markets

Nongnuch Soonthornchawakan
Louisiana State University and Agricultural and Mechanical College

THE ROLE OF NETWORKS IN LABOR MARKETS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by
Nongnuch Soonthornchawakan
B.A., Thammasat University, Thailand, 1987
M.A., Thammasat University, Thailand, 1990
M.E., North Carolina State University, USA, 2005
M.S., Louisiana State University, USA, 2006
August, 2007

To my dad, my aunts, Krongthip, and Chor

ACKNOWLEDGMENTS

There are absolutely no words in any language that can adequately express my deepest gratitude and appreciation that I have for my advisor, Dr. Sudipta Sarangi, and co-advisor, Dr. Tibor Besedes. I am most indebted to them for providing assistance, guidance, knowledge, and all supports. This dissertation would not have been possible without them. I will never forget everything that they did for me. As a tribute to them, I will give my future students the same help and all supports that were afforded to me when given the opportunities. Thank you for everything that you two have done for me! I am grateful as well to distinguished members of my dissertation committee, Dr. Robert J. Newman, Dr. David M. Brasington, and Dr. Belinda C. Davis, for providing suggestion, and insightful comments that help me enhancing the quality of my work.

I would like to thank Thammasat University and Faculty of Economics at Thammasat University for financial support. I am indebted to Dr. Sukrita and Dr. Savaraj Sachchamarga, Dr. Sukanya Nitungkorn, and Dr. Saipin Chintrakulchai, for stimulating, and encouraging me to pursue the doctoral program. I also offer heartfelt thanks to them for on-going support throughout the studying period. I would particularly like to thank Dr. Bhanupongse Nidhiprabha, and Dr. Arayah Preechametta, for help and encouragement.

My special thanks go to Pisut Kulthanavit, and Supanit Tangsangasaksri, for always cheering me up. Their good will kept me going during these past years. Many thanks to my best friend, Fan Chen, for the valuable and on-going help he has given to me. Fan, thanks for being around in every situation! I wish to thank Suwandee Chaiwarut, Mitsunori Yokoyama, Dr. Tung-Hsiao Yang, Taehee Han, and several other friends and colleagues whose names I cannot continue listing for their friendships and support. Their friendships meant a great deal to me.

Above all, I owe a huge debt of gratitude to my aunts, who have devoted themselves to raise me after my parents passed away, and my sisters, for their love and all support throughout my entire life. I could not be here without them. My special thanks go to my sister, Krongthip, for helping my aunts take good care of me since I was young. No matter what I do, she always supports and understands me. Thanks also to Chor who always be here for me. Thanks for your love and your care. Especially, thank you for the belief and confidence that you have had in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Networks of relationships play an important role in the social and economic operation of the labor market. Social connections have been shown to be crucial in influencing the transition and efficiency in the labor market because they can quickly spread information over large segments of society. In particular in "*small world*" networks everyone can connect to others through very few intermediaries and information can spread far and fast over such a small-world network. The first chapter of this dissertation starts with the formal elements of social network analysis and graph theory. It then provides an overview of the emerging literature on models of small worlds. Networks characterized by very small *characteristic path lengths*, yet high *clustering coefficients*, are said to exhibit the small-world phenomenon.

Since interactions or links in the academic labor market are observed easier than other labor markets, the second chapter investigates the labor market for academic economists from a social network perspective. The sample includes the top two hundred economics departments in the world and provides a separate analysis of the subset pertaining to North America. The data indicates the stronger links between higher ranked universities than between the universities in the higher and lower ranked universities. The obvious pattern of interaction in the network is that the top-ranked grantors place their Ph.D. economists mostly in group ranked below them.

The small-world properties of this network are examined in the third chapter. The data confirms the small-world phenomenon in the economics academic network. Any two ranked universities can be connected through approximately three links only. Although it is shown that there is inequality in the placement of Ph.D. students, there are many centers of connections in the network. However, most of the influential universities in terms of centrality of the network are not the ones influential in granting doctoral degrees.

CHAPTER 1: SURVEY OF MODELS OF THE SMALL-WORLD PHENOMENON

1.1 Introduction

Networks of relationships play an important role in the response to job opportunities because they can quickly spread information over large segments of society. Hence, the structure of a social network is a key factor in determining labor market interaction. This paper will study the structure and properties of such a network, as well as explore its role in the labor market.

The infectious diseases, news and rumors can widely spread by interpersonal contact, from coast to coast, country to country, continent to continent over a social network, in which the average number of separation is very few or short links, faster than over a network, which has a hundred degrees of separation or too many links. In this paper, particular attention will be given to the network model for analyzing the small-world phenomenon proposed by Watts and Strogatz (1998) and Watts (1999a and 1999b). Using graph theory, the authors demonstrate that everyone in the world can be reached through a short chain of social acquaintances, a fact which results in subjective perceptions of a small world. Watts (1999b) defines a "small" network as one in which "almost every member of the network is somehow 'close' to almost every other member, even those that are perceived as likely to be far away." (p. 495) In other words, even if one travels to a completely strange place and encounters a local person with whom they have absolutely no friends in common, there still exist a small number of "friends of friends" that bridge the social gap between them. The two essential, and related, properties of this "'small world" network are:

1. The average distance between each two members is much shorter than in a network generated by a random process.

2. The number of connections between members is much greater in comparison to a random network.

This chapter is organized as follows: Section 2 will summarize briefly the existing literature dealing with the small-world network. Section 3 will introduce some basic notations in graph theory and their definitions. Their uses in explaining the small-world network phenomenon proposed by Watts and Strogatz (1998) and Watts (1999a and 1999b) will be reviewed. Sections 4 and 5 will illustrate the characteristics and applicability of small-world models that occur at a possible topology. A summary of the paper is presented in the last section.

1.2 Literature Review

The small-world phenomenon was first examined by Stanley Milgram (1967) through a curious social experiment. He addressed a number of letters to stockbrokers in Boston, Massachusetts and randomly distributed them to people in Nebraska, referred to as the starting participants. In an attempt to generate an acquaintance chain from each starter to the target, he instructed individuals to pass the letter down only to a first-name acquaintance of theirs. If the participant did not know the final addressee in person, he was not allowed to send the letter directly to them. Instead, he could send the letter to an acquaintance judged to be more likely to know the target. To prevent people from sending the document to someone who had already participated, the list of prior addressees was attached. All participants were volunteers. They were not paid or rewarded as an incentive for the completion of a chain. In order to keep track of a letter's progress and collect information at every level, Milgram asked each participant to fill out a "tracer" and mail it back to him. He found that the average "chain length," i. e., the number of intermediaries required to link the starter and the target, is about five. This resulted in his conclusion that Americans have only "six degrees of separation" (as cited in Watts 1999b, p.

493) from anybody else. In other words, any two people in the United States of America can be connected through only six links.

Although Milgram's findings were undermined because the experiment contained many possible sources of error, the small-world phenomenon remains a viable concept. It has been applied to social networks other than those based on kinship and friendship. After Brett C. Tjaden published a computer game, “The Six Degrees of Kevin Bacon” on a University of Virginia's Web site based on the small-world problem, “six degrees of separation” became a popular notion in the entertainment industry (Newman 2000).

Tjaden used the Internet Movie Database (IMDB) to document connections between different actors. He postulated, facetiously, that the actor Kevin Bacon was somehow at the center of the movie actor network (as cited in Watts 1999a, p. 3). The Bacon-number represented the fewest number of steps, through roles in films, by which an individual actor or actress is separated from Kevin Bacon. As illustrated by Martin (2005), Kevin Costner had been in film with Kevin Bacon in JFK. Costner’s Bacon-number is, then, one. Henry Winkler, on the other hand, had never been in a film with Kevin Bacon, but he appeared with Michael Keaton in Night Shift (1982) who had been in film with Kim Basinger in Batman (1989) who had been in film with Mickey Rourke in 9 ½ weeks (1986) who had been in film with Kevin Bacon in Diner (1982). This would yield a Bacon-number of four for Henry Winkler. Except, Henry Winkler had been in film with Clint Howard in Little Nickey (2000) and he appeared with Kevin Bacon in My Dog Skip (2000). So, Henry Winkler’s Bacon-number is only two (see Figure 1.1). Thus, the Bacon-number measures the shortest path between two agents in the Hollywood network.

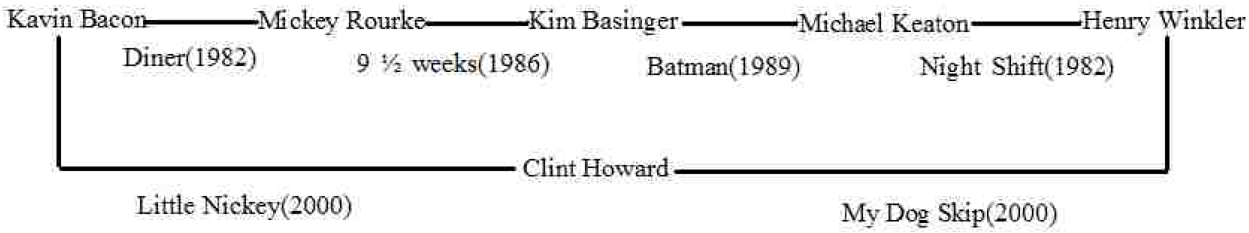


Figure 1.1: Sample of Bacon-number in Film Network

Initially, Tjaden hypothesized that no-one who has ever been in movie in an American film has a Bacon-number greater than four but later found that the highest known Bacon-number is, in fact, eight (as cited in Watts, 1999a, p. 3).

Another famous primary measurement of the small-world phenomenon is the Erdos-number project. While the Bacon-number measures the distance between an individual and Kevin Bacon in Hollywood, Erdos-number, which was most likely first defined by Casper Goffman (1969), measures the collaborative distance between any mathematician and Paul Erdos, the great mathematician of the 20th Century. Two mathematicians are considered to be connected if they have at least one joint paper together. Erdos published about 1401 papers in Mathematical Reviews and had 502 direct collaborators (Martin, 2005). These are the people with Erdos-number one. People who have collaborated with them, but not directly with Erdos, have Erdos-number two, and so forth. For example, Albert Einstein has Erdos-number two, since he did not collaborate with Paul Erdos, but he did publish joint research with Ernst Straus, who was one of Paul's major collaborators. (Grossman, 1996) Both the Kevin Bacon Game and the Erdos-number Project verified that only a short chain of intermediate acquaintances connect any two people in the network, presenting particularly clear perspectives on the small-world phenomenon.

The small-world phenomenon was first examined mathematically by Manfred Kochen and Ithiel de Sola Pool (1978). Under the assumptions of independence of connections and no definite structure of the network, they calculated expected values of the probability (p_i) that two randomly selected elements of a network would be connected via a shortest path consisting of *i* intermediaries. They estimated that any people is likely to be connected to any other with number of 2 intermediaries (three lengths of chain). Kochen (1989) improved on Kochen and Pool's work and confirmed their conclusion that the world was probably as small as Milgram had shown.[1]

Rapoport and Solomonoff (1951) was the one of the pioneer theoretical investigations of the small-world phenomenon. The study looked at distance in a social network. Under the assumption that every element had the same number of connections, Rapoport and Solomonoff developed the idea of a disease spreading in a randomly connected network. Starting from a first small infected set, they calculated the expected fraction of population to be affected eventually. The total infected fraction was then determined by exponentially projecting the spread of the disease ahead in time. The additional importance of this work lies in that the induction took into account some of the structural specifications of a social network. These include the tendency of people in the network to connect with others who bear similarities to them, the fact that connections are bilateral (people know each other), and that one's acquaintances tend to know each other directly, forming a dense cluster of 'inter-acquainted' individuals.[2]

Building on this study, more improved calculations were constructed by Ferraro and Sunshine (1964) and Skvoretz (1985).[3] In their work, they accounted for the variable number of connections each person possesses in a social network. Common sense dictates that different

[1] Both Kochen and Pool (1978) and Kochen (1989) was discussed by Watts (1999b, p. 497).
[2] Rapoport and Solomonoff (1951) was discussed by Watts (1999a, pp. 12-13).
[3] Ferraro and Sunshine (1964) and Skvoretz (1985) were discussed by Watts, 1999a, pp. 13-14.

individuals' number of acquaintances may range widely. A person who becomes infected with the disease may have a number of connections equal, or less than, the person who infected him or her. This introduces fluctuations in the speed at which the disease spreads. Moreover, the spreading of the disease also depends on the strength of a person's social ties, as introduced by Granovetter (1973).

Both Granovetter's strength of ties and Barnes' (1969) "density" (as cited in Watts, 1999a, p. 15), which measures how intense the connections in the network are, come very close to the concept of "clustering," a phenomenon defined by Duncan Watts and Steven Strogatz (1998) and Watts (1999a and 1999b). They were the first to propose an alternative model for the small-world phenomenon by using graph theory. They began by building an initial one-dimension graph, which contains a specific number of elements *n*. Each element has *k* links to other elements. Then, they randomly rewired each link with probability *p.* They measured the fraction of connections in the graph to the possible number of connections, called "the clustering coefficient." They then measured the average shortest path by which each element can reach any other. This was referred to as "the characteristic path length," and was used to determine whether or not the graph describes a small-world network. The idea is that random rewiring introduces increasing amounts of "disorder connections," i. e. shortcuts between elements that super cede the sequential connections in the initial constructed graph. The chains between elements in the network get progressively shorter. Not only does the average path between elements in the network shrink dramatically after the introduction of shortcuts, it drops to a value which is almost as small as the probability that any member can meet anyone else by chance.

Even though any element might be able to reach any other, even those in a remote cluster, through much shorter chains, the connections within each subset of elements remain unchanged.

This means that the introduction of shortcuts has a very small impact on the average density of connections in the entire graph. Thus, the clustering coefficient of a randomly rewired network can remain almost as high as the clustering coefficient in the initial graph. It is, in any case, far greater than the clustering coefficient in a network where no-one knows anyone else and can only encounter others by accident.

It is this kind of network, with very small path length and high clustering that Watts and Strogatz called a “small-world” network. They proceeded to investigate scientifically three practical cases: the network of actors in feature films, the electric power grid of the Western United States and the neural network of the nematode worm *Caenorhabditis elegans*. They found that all three networks exhibited a much higher degree of clustering and slightly higher characteristic path lengths relative to a network of strangers that meet by chance. Thus, they concluded that all three networks exhibit features of the small-world phenomenon. Following the theoretical groundwork established by Watts and Strogatz (1998) and Watt (1999a and 1999b), Zlatic et al. (2006) examined the Wikipedia network of articles and concluded that it also exhibits signature traits of the small-world phenomenon.

1.3 Graph Theory and Network Definitions

Real-life examples of small-world networks are currently studied as occurring at a possible topology somewhere between two extreme models. The two theoretical extremes are the abovementioned network of complete strangers who meet accidentally and the network where everyone knows everyone else. This section provides background information on graph theory and introduces some standard technical terms relevant to this kind of network analysis.

A graph G contains a set of elements in a network, shown as points in the graph and called *vertices* or *nodes*. The number of vertices in graph G is n. The total set of vertices in graph

G is $V(G) = \{1, 2, ..., n\}$. Vertices, representing the network's discrete elements, may stand for people, animals, computer terminals, organizations, institutes, etc. Any two vertices may be linked directly to each other by a line called *edge* or *link*, visible in the graph as the line segment connecting two vertices. $E(G)$ is the total set of edges in graph *G.* An edge represents some sort of relationship between the connected elements, be it friendship, alliance, or the peculiar interplay between predator and prey. Figure 1.2 provides a simple example of a graph G containing a set of vertices $V(G)=\{A,B,C,D,E\}$ and a set of edges $E(G)=\{e_{A,B}, e_{B,C}, e_{B,D}, e_{B,E}, e_{C,D}, e_{D,E}\}$.

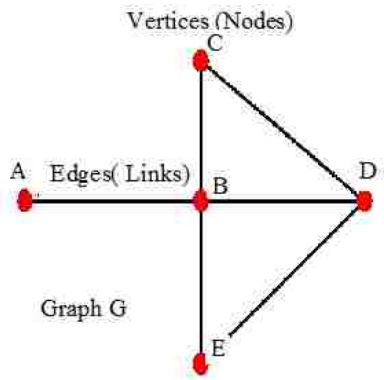
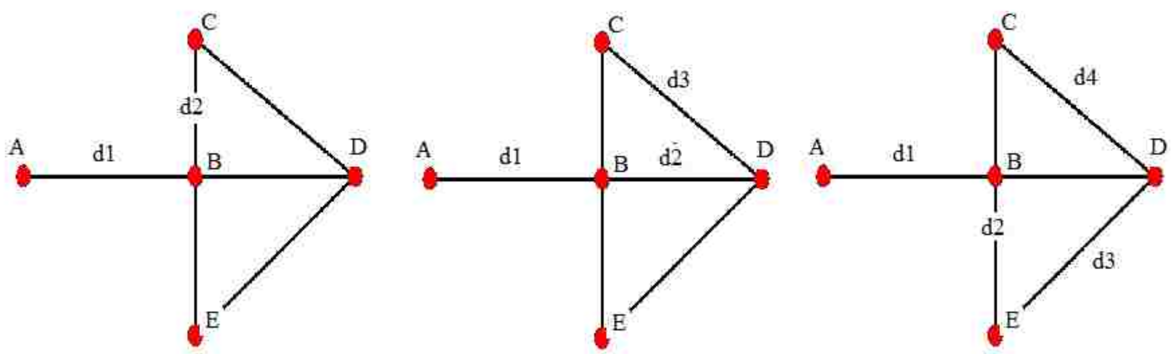


Figure 1.2: Graph *G* containing vertices and edges

The vertices in graph *G* may be indirectly connected through a sequence of vertices and edges called a *walk*. A walk in which each point and each line are used only once is called a *path*. A *closed path* is a walk which begins and ends at the same vertex. A closed path with at least three distinct vertices is called a *cycle*. The *path length* is measured by the number of edges connecting the vertices to each other. Let $d(e)$ be the sequential number of edges passing through vertices to connect one vertex to any other. In Figure 1.2, vertices A and C are not directly connected by an *edge*. Still, there exist an infinite number of *walks* from A to C. They are, however, connected through only three *paths*, illustrated in Figure 1.3. One possible path through which vertex A can reach vertex C is by using the two edges $e_{A,B}$ and $e_{B,C}$. We can see in Figure

1.3a that if the sequential paths *d1* and *d2* are used, then the path length is 2. The second path (Figure 1.3b) shows that vertex A can reach vertex *C* by using two edges $e_{A,B}$, $e_{B,D}$ and $e_{D,C,}$ through the sequential paths *d1*, *d2* and *d3*, which would result in a path length of 3. Figure 1.3c illustrates a path from A to C with a length of 4. Therefore, the path length between vertices A and C can be 2, 3, or 4.



a. b. c.

Figure 1.3: A to C Paths and Path Length

Distance (or *geodesic distance*) is the minimum number of edges traversed between vertex *i* to vertex *j,* denoted by $d\ (i, j), i, j \in V(G)$. In other words, *distance* is the shortest path length between one vertex and another. If vertex *i* cannot reach vertex *j* in the graph (or no path between them), then $d(i, j) = \infty$. Figure 1.4 displays the geodesic distance of all vertices in graph *G*. For instance, the distance between vertices A and C is 2. That is $d(A,C) = 2$.

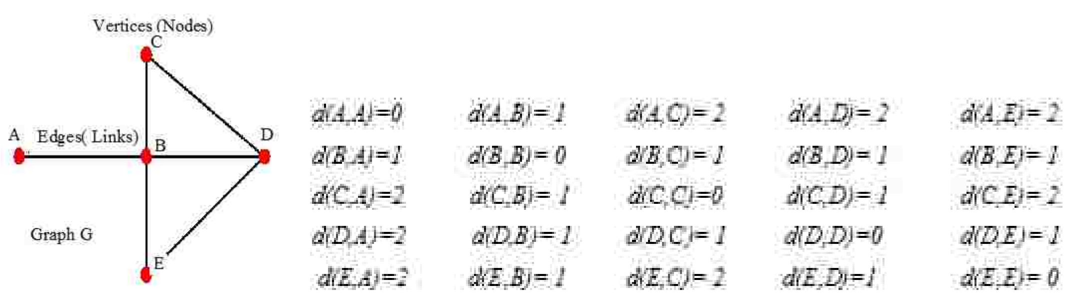


Figure 1.4: Geodesic distance

In graph G, the longest or maximum *distance* between any pair of vertices is called a *diameter*. It is denoted by $D(G)$ where $D(G) = max_{i,j \in V(G)} d(i, j)$. For the graph in Figure 1.4, the diameter equals 2 (D(G)=2).

A graph is said to be *connected* if any vertex can be reached by any other vertex in the graph. That is, any vertex can be reached by a finite number of connecting edges, forming a path in the graph. Vertices i and j may be connected either directly, or through a path formed by the set of intermediaries j_1, j_2, $j_3, \ldots, j_n$. A *complete graph* is one in which each vertex is directly connected to every other vertex. The overall number of connections among vertices in the entire graph is a measure of the network *density*. The more vertices are connected to one another directly, the denser the graph will be. Density is more specifically defined as the proportion of actual edges to the possible number of edges in a completed graph. If the total number of vertices in the graph (its *size*) is n, then the total number of possible edges is n^2-n. The density value is high if there are many vertices directly connected to one another, and approaches 1 if nearly all vertices are directly connected to one another. A completed graph has density equal to 1.

Even if a graph is not connected, parts of it may be. A graph may be partitioned into various groups, called *components***.** Within a component, all vertices can reach each other directly or indirectly, by paths. There is no path between all pairs of vertices in different components. Some isolated vertices and/or components will have no connection with the rest. Only the largest component is considered for social network analysis because the distance between two vertices inside the largest component can be measured in a finite number.

Let the network partition of a given graph be denoted by $CM(G)$. The graph can be divided into z components, $CM(G)=\{CM_1,CM_2,\ldots,CM_z\}$, where z$\geq 1$. The graph is *disconnected* if more than one component exists within it. A connected graph is composed of a single component.

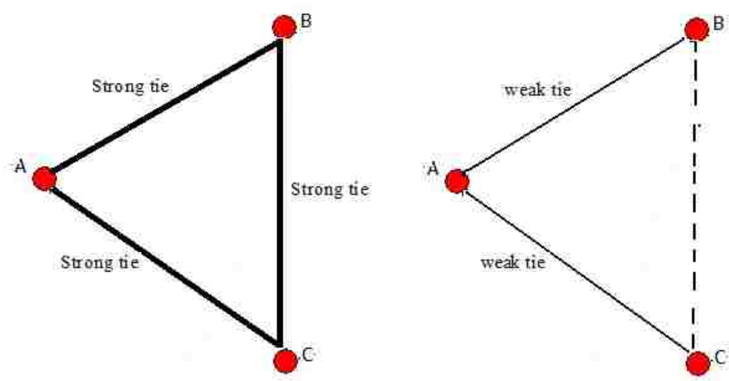
A component can be thought of as a *connected subgraph*, i.e., a selected subset of interconnecting vertices. A *complete subgraph* is known as a *clique*. A *clique,* then, is a subset of vertices in which every member is directly connected to every other. Within a single component, there may exist any number of *cliques* connected to each other. When a *clique* is under consideration as a section within its host component, it is referred to as *local*.

The above framework attempts to capture the fact that, in social networks, members of the same clique are often acquainted with all other members. A one-to-one connection between two members of every social network is considered to be a relationship such as that between friends, neighbors, relatives, or coworkers.

Montgomery (1991) argued that the pattern of social ties between individuals may be an essential factor in explaining labor market dynamics. After all, employee referrals are often crucial sources of employment information. They affect not only the response to job opportunities but also the screening of job applicants. Granovetter (1973) ventured that the *strength* of a *tie* between two people is a function of the amount of time spent together, emotional intensity, degree of intimacy, and reciprocal services. *Strong ties* are said to exist when two people knew each other well over long periods of time. They indicate an intimate relationship.

Strong ties are considered more likely to possess the transitivity property because they suggest an affinity between two people of similar character, which often 'overflows' to include a third person. A social network exhibits a number of *triads* in which each person knows the other two people. In any part of the graph where two edges connect three vertices, there is potential for a triad to form. If strong ties connect A with B and with C, then B and C are more likely to connect to each other and complete the triad. Vertices connected by strong ties are more likely to be similar to each other, which means both vertices B and C are similar to A and therefore likely to build a direct relationship. Within the same community or clique, one-to-one ties connecting each pair of members tend to be strong and clustered.

Weak ties are said to connect people who contact one another less frequently and whose relationships are acquainted or casual. Weak ties are less likely to possess the transitivity property. The vertex B, if connected to A by a weaker tie, is less likely to connect to vertex C which is connected to A by a similar weak tie. The chance of *triad* completion should be less in the case of A-B and A-C weak ties than strong ties (see Figure 1.5).



a. Strong ties and triad completion b. Triad is less likely to occur with weak ties

Figure 1.5: Triad

While edges connecting vertices within the same clique are viewed to be strong ties, edges spanning two different cliques, known as *bridges***,** are considered to be weak. A *bridge* is defined as a link in a network that provides the only path between two vertices. If the bridge is removed, a new component (or subgraph) will be formed. This is because the bridge is the only route through which information can flow between two cliques in the network. However, in a large network such as a human society, there may be multiple routes connecting two vertices in different cliques. With the increase of distance following the removal of a direct link, the communication between these two vertices may become costly and subject to distortion along the way.

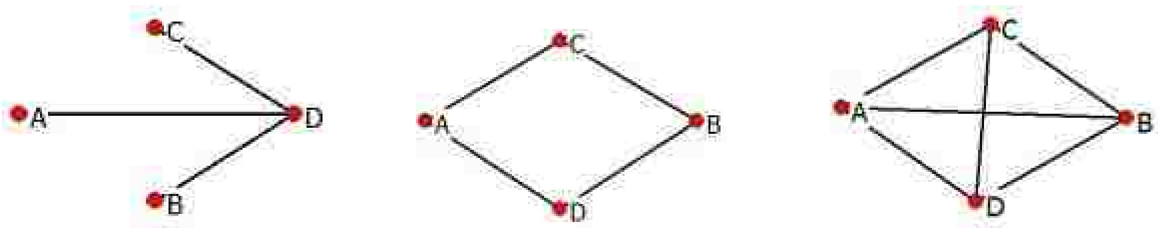
Figure 1.5a illustrates that vertices A and B are connected with a strong tie. A is simultaneously connected to C with a strong tie, creating the possibility of a tie between B and C. If a triad is formed, A can now choose to reach C by traversing through B, who will then connect to C. The path between A and B, thus is not a *bridge*. However, in Figure 1.5b, the ties between vertices A and B, and between A and C, are weak. A direct connection between B and C is unlikely. In this case, the path AB may perform the function of a *bridge*, as a sole connection between B and C, and their respective cliques. We can say then, that strong ties are less likely and weak ties are more likely to be bridges.

How does this model inform our understanding of labor market dynamics within a social network? Here is an example illustrated by Bonacich (2006). A person's close friends are in his own social circle or subgroup, all connected by strong ties and likely to know one another. Since people connected by strong ties are more likely to share sources of information, the information they can provide in turn is quite homogenous. The tip a job seeker receives from friends or relatives is likely to be more than necessary. On the other hand, one's acquaintances are likely to move in different social circles or subgroups, only accessible through weak ties. Yet, it is through such ties that information between subgroups flows. Although a few weak ties may result in very little communication between two cliques, they can be highly effective in bridging social distance. This is so because a person's casual or remote acquaintances are much more likely to be exposed to different sources. The job seeker, then, can receive new information from them that is unavailable from his friends, relatives, or close acquaintances.

Granovetter (1973) made a case for "the strength of weak ties." He selected a random sample of professional, technical, and managerial job changers living in a Boston suburb and asked them how they arrived at the job information. He found that most employees had found their present jobs through an acquaintance that they seldom contact. He concluded that weak ties are an important resource in increasing a person's job mobility and exposure to new job opportunities.

Quantitative measure of the strength of ties is based primarily on frequency, which is embedded in the definition of strength. In the case of collaborations between pairs of scientists, studied by Newman (2004), the frequency of co-authorship between two scientists, estimated as the number of co-authored papers over a given time period, can be used as a proxy for the strength of their tie. To refine this estimation, one needs to assign those papers different weight proportional to the time the two authors spent working together. If a paper has *n* co-authors then, while working on this paper, author *i* divides his time between his *n*-1 co-authors. The strength of tie between two authors working on a paper with *n* co-authors should be discounted by $\frac{1}{n-1}$, for instance. If two authors wrote two papers together, but on one of them had 8 more co-authors, then the strength of their tie is given by $1+1(\frac{1}{9})=1.11$.

Network graphs can also be presented by way of an *adjacency matrix* or *adjacency list*. Vertex i is adjacent to j if vertex i is directly connected to vertex j and not just connected to vertex *j* through other vertices. The adjacency matrix *A(G)* is an $n \times n$ matrix, where $A_{i,j}$ is the number of edges that vertex i directly connects to vertex j. The adjacency list, then, is the list of vertices directly linked to each vertex i. Examples of graphs and their adjacency lists are presented in Figure 1.6. From any graph, we can derive an adjacency list and vice versa.



$$A_{4*4} = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}, \quad A_{4*4} = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \end{array}, \quad A_{4*4} = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Figure 1.6: Graphs and their adjacency matrices

The vertices adjacent to a particular vertex *i* are said to be its *neighborhood***.** The neighborhood ($\Gamma(i)$) of a vertex *i* is the subgraph (S). It consists of the vertices adjacent to *i*, excluding *i* itself. In other words, the neighborhood of *i* is formed by all vertices with direct connections with vertex *i*.

Let S be a connected subgraph of graph G. The neighborhood ($\Gamma(S)$) of the connected subgraph (S) is all vertices directly connect to any vertices in S. These all vertices are not included in S. If $\Gamma(i)$=S, then $\Gamma(\Gamma(i)) = \Gamma(S)$. The 1st neighborhood of vertex *i* is the set of vertices directly connected to vertex *i*. The 2nd neighborhood of vertex *i* is the set of vertices which connect to vertex *i* through the 1st neighborhood of vertex *i*. The j^{th} neighborhood of *i* is given by $\Gamma^{j}(i)$. The 0th neighborhood of vertex *i is* the vertex *i* itself. In the connected graph, each vertex *i* can have at most j_{max} neighborhoods and can sequentially disperse information throughout the graph through its friends from the 1st neighborhood up to its $j_{\max}^{th}$ neighborhood.

The number of the size of any vertex's neighborhood is called a *degree*. The number of vertices that vertex *i* directly connects to is the degree of vertex *i*, denoted k_i.

Let the set of vertices connected to *i* in graph G be $V_i(G) = \{ j \in V(G) : e_{i,j} = 1\}$. Then, the degree of vertex *i* is $(k_i) = |V_i(G)|$. The average degree of all vertices in graph G, denoted by k, is defined as $k = \sum_{i \in V(G)} \frac{|V_i(G)|}{n}$.

If it has a high degree as a result of being "well-connected," a vertex might be regarded as possessing centrality within a graph. However, the centrality of a given vertex is determined not only by the size of its immediate neighborhood but also by that of its distant ones. If *i* performs a crucial role as information channel between a large number of vertices in the graph, it can be seen as lying "between" them. This is known as *betweenness centrality*. The betweenness centrality of vertex *i*, also known as *node betweenness*, is measured as the frequency of vertex *i* that falls in the shortest paths between other pairs of vertices. The greater this number is, the higher the node betweenness score of vertex *i* will be.

In the collaboration network analyzed by Newman (2004), a few scientists have much higher node betweenness score than the majority. The distribution of node betweenness scores in the scientist collaboration network approximately follows a power law. Let *q(x)* be the probability distribution of the betweenness score. A power law relationship can be represented as $q(x) = Cx^{-\alpha}$, where C and α are constants. The equation can be rewritten by taking logs of both side, $\log(q(x)) = \log C - \alpha \log(x)$, and can be represented as a linear line on a log-log graph. The cumulative distribution *Prob*($X \geq x$)[4] follows the power law as given by: $Prob(X \geq x) = \int_x^{\infty} Cx^{-\alpha} dx = \frac{C}{\alpha - 1} x^{-(\alpha-1)}$. The scale of cumulative distribution on a log-log graph,

[4] Newman (2005) demonstrated that the cumulative distribution under power law sometimes follows the Pareto distribution.

given by the power law, is also a straight line. The power law indicates that the probability of finding vertices with a large number of degrees is significant.

Since the node betweenness score distribution of the scientist collaboration network is said to approximately follow a power law, this translates into a few influential individuals and a multitude of peripheral actors. Removal of any of these influential individuals from the network will cause it to be disconnected. Moreover, Goh et al. (2003) found that those few influential scientists tend to collaborate primarily with others of the same caliber.[5]

1.3.1 Network Definitions

A number of different types of network graphs will be presented by using the standard definition as follows:

In an **undirected graph** connections between vertices have no direction (no arrow on edges, as in Figures 1.1-1.6 and Figure 1.7a). This implies that relationships between two vertices are symmetric, or reciprocal. You are my friend and thus I am your friend. For instance, in Figure 1.7a, an arrowless line between points A and B represents that vertex A relates to vertex B the same way that vertex B relates to vertex A. Both act simultaneously as receivers and senders of information or goods.

In a **directed graph** original, or source actors, reach target actors as represented by arrows. The two directions are counted as being distinct edges. The directed graph indicates the directional ties between two vertices. In Figure 1.7b, vertex A relates to vertex B, but vertex B does not relate to vertex A. Or, it can be interpreted that vertex A sends information or goods to vertex B without receiving anything in return. Vertex A is a sender (or source of information) while vertex B is a receiver (or target of information).

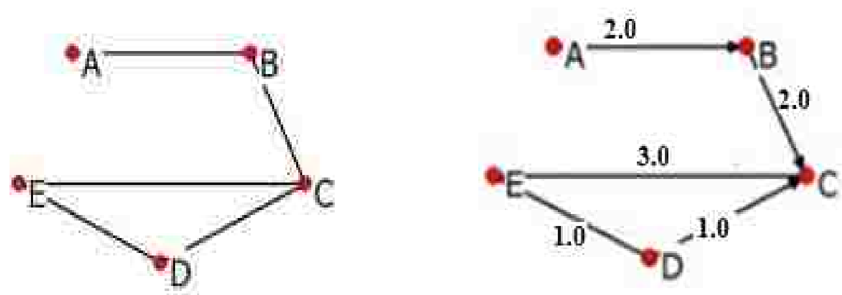
[5] Newman (2004, p.12)

For the directed graph, while the size of any vertex's receivers or targets of information can be measured and called "out-degree", the size of senders or sources of information of any vertex can also be measured and called "in-degree." Let the set of all other vertices connected by vertex i be $V_i(G)=\{j \in V(G):\ e_{i,j=1}\}$, the out-degree of any vertex i, denoted as k_i^{Out}, is $k_i^{Out} = |V_i(G)|$. In other words, the out-degree, which is the number of connections from a vertex to all other vertices in the neighborhood, can be computed as the sum of ones within vertex *i*'s row in the adjacency matrix.

Let the set of all other vertices connecting to vertex i be $V_i(G)=\{j \in V(G):\ e_{j,i=1}\}$, the in-degree of any vertex *i*, denoted as k_i^{In}, is $k_i^{In} = |V_i(G)|$. In other words, the in-degree of any vertex *i*, denoted as is the number of connections received by vertex *i* from all other vertices and can be computed as the sum of ones within vertex *i*'s column in the adjacency matrix.

In an **unweighted graph** edges are not assigned any strength, or value, of ties. An adjacency matrix derived from an unweighted graph is also unweighted—it informs us only whether there exist edges between vertices or not. The connection between any pair of vertices $i, j \in$ V(G), would be $e_{i,j} = \{0,1\}$, where $e_{i,j} = 1$ means that there is a connection between them and $e_{i,j} = 0$ means there is no connection between them. Examples of unweighted graphs are found in Figures 1.6 and 1.7a. In the latter, edges are unweighted, only indicating that vertices A and B are connected to each other. The amount of information shared between them is therefore implicitly considered to be equal. In other words, vertex A knows everything vertex B knows and vice versa.

In a **weighted graph (or valued graph)** numerical values, such as the value of a relationship or frequency of interaction, the magnitude of goods or the information being exchanged, are assigned to each connection. Entries in a derivative adjacency matrix are not simply zero or one, but can be any integer number. The weighted graph can be either directed or undirected. Figure 1.7b is an example of a weighted directed graph (or digraph), in which there is a set of integer values attached to each arrow. The weighted directed graph in Figure 1.7b represents both a directional and valued relationship, such as vertex A sending 2 units of goods to vertex B. B acts as a receiver from A but also serves in the capacity of a sender, relaying the 2 units to vertex C.



a. Unweighted undirected graph b. weighted directed graph

Figure 1.7: Examples of directed, undirected, weighted and unweighted graphs

An unweighted and undirected graph containing no loops (i. e. no edge starts and ends at the same vertex) and one in which no multiple edges exist between any two vertices is a **simple graph**. In other words, vertices i and j connect to each other with only one edge and vertex i has no edges to itself. Figure 1.7a is a simple graph. The adjacent matrix of a simple graph is a matrix with rows and columns labeled by graph vertices, in which all entries would be either 1 or 0. If vertex i is directly connected to vertex j, the entry is 1. If there is no direct connection between vertices i and j, the entry value is 0. Since a simple graph is an undirected and an un-looped graph, the adjacency matrix must have zeros on its main diagonal and must also be symmetric.

A **sparse graph** is a graph in which the number of edges connected to other vertices is much less than the number of possible edges in a completed graph of the same size. The maximum number of possible edges in a connected graph is denoted by *M*. The number of edges in the fully connected (completed) graph is $\binom{n}{2} = \frac{n(n-1)}{2}$. The sparseness condition is met when M $<< \frac{n(n-1)}{2}$.

A graph in which all the degrees of all vertices are equal (i. e., same number of neighbors as in: $k_1 = k_2 = k_3 = k_n = k,$) is called a *k*-**regular graph.** A 0-regular graph consists of all isolated vertices. In a 1-regular graph, each vertex must connect to only one another vertex. A 2-regular graph is a cycle or ring. A 3-regular graph and a 4-regular graph are called as a *cubic*, and a *quartic* graph, respectively. The relationship between the average degree (k), maximum number of edges (M), and the number of vertices (n) in a network is given by: M = $\left(\frac{n \cdot k}{2}\right)$ or $k = \frac{M}{n} \cdot 2$ in any graph of size *n* where each vertex, or actor, has the same degree *k*. That this relationship is valid can be clearly seen in a 2-regular graph where $n = 6, k = 2$ and M=$\frac{6 \cdot 2}{2} = 3$. A 4-regular graph with $n = 6, k = 4$ has M =12 (see Figures 1.8a and 1.8b).

a. 2-regular graph with $n = 6, k = 2$, M=3 b. 4-regular graph with $n = 6, k = 4$, M =12

Figure 1.8: k -regular graphs

An **empty graph** is one of size *n* containing *n* isolated vertices. There is no edge connecting any of the *n* vertices in the graph. An empty graph is known as a 0-regular graph.

A **cycle graph** is one containing at least one path that returns to its starting point. In general, the cycle graph contains no self-loop. Examples of a cycle graph include the triangle and square graphs.

A **tree graph** is a simple, undirected, connected graph in which any two vertices are linked by only one path and no cycles. The points of connections at the end of two or more segments in a tree graph are called *forks* and the segments are called *branches*. Final vertices at the end of tree graph are called *tree leaves*. A tree with one vertex at its center is called a *central tree*. This type of graph is relevant to the representation of certain kinds of social networks, such as the organization of institutional and social hierarchies.

A **Cayley tree** is a tree graph in which each non-leaf vertex contains a fixed number of branches *y*, forming an *y*-Cayley tree.

A **path graph** is a tree graph with *n* vertices where two vertices have a degree of 1 and the rest n-2 vertices have degrees of 2. The two vertices with degree 1 appear as the starting or ending points of the graph. A path graph with 3 vertices is an example of a 2-Cayley tree.

An ***n*-star graph** is a tree graph with *n* vertices in which one vertex, with a degree of *n*-1, plays the role of a center, connecting to all other vertices. The other *n*-1 vertices, having degrees of 1, are only connected to the center vertex. The star graph is considered as a *y*-Cayley tree with *y*+1 vertices in the graph.

A **Moore graph** is a regular graph of vertex degree $k > 2$ in which the upper boundary of the number of reachable vertices in the graph is given by $k\sum_{d=0}^{D-1}(k-1)^{d-1}$ where *d* is the distance of any vertex to any other reachable vertices. The Moore Graph is constructed as the tree graph which consists of branches and leaves. There is some number of vertices in the graph of size *n*, which are not met with any other vertices in the tree graph. Therefore, the Moore graph is considered as perfectly expanding in the sense that at the end of the tree graph none of its vertices are adjacent with each other.

Watts and Strogatz (1998) and Watts (1999a, 1999b) analyzed the small-world phenomenon by utilizing graph theory. They limited their use to undirected, unweighted, sparse, and connected graphs. Under these constraints, the first property that a small-world network needs to have is that, with a fixed population size n, it must satisfy the condition $n >> k >> 1$. They measured the average distance, or characteristic path length (*L*), and defined the clustered characteristic of a graph, known as the clustering coefficient (*C*), in order to explain the properties of small-world networks. What follows are definitions and explanations of their analysis, accompanied by other related statistics.

The *characteristic path length*, or the average distance in the entire graph, denoted by *L(G)* is the average of the average path lengths between any two vertices. Since any vertex *i* in the connected graph G of size *n* can reach any other *n*-1 vertex, an average distance between *i* and any other vertex *j*, denoted by *L(i),* is given by $L(i) = \frac{\sum_{j \in V(G)} d(i,j)}{n-1}$. The characteristic path length *L(G)* in the connected graph is computed from the average *L(i)* for every vertex, so that

$$L(G) = \frac{\sum_{i \in V(G)} L(i)}{n} = \frac{\sum_{i \in V(G)} \sum_{j \in V(G)} d(i,j)}{n(n-1)}$$

The characteristic path length reflects the global structure of a network graph. Since the characteristic path length is the average number of edges that crosses over the path between any pair of vertices in graph G, it can be interpreted as an indicator of the average number of chains that connect any two people. In an unconnected graph, the characteristic path length is infinite. In the real world, a global social network may not be fully connected. Another way of saying it is that there may be more than one component in it. In that case, in order to study the small-world phenomenon in a social network, the characteristic path length of its largest component, what Goyal et al (2004) called a "giant component," must be measured, relative to that of all other small components.

In the case of a network with very large *n*, it is difficult to measure the average characteristic path length. Huber (1996) suggested that it is more efficient to measure the *median path length***,** denoted by L_{median}, by using the random sampling technique (Watts 1999a, p29). In other words, estimating the median of path lengths obtained from randomly selected vertex pairs is more practical than estimating the average path length of a network with a great number of members.

To obtain L_{median}, first we must randomly select *S* samples containing n_S vertices. We then estimate the shortest path length between each vertex (*v*) in that sample (*S*) to all other vertices in *S*, $v \in V(S)$. Thus, all distances within the random sample $d(v,j)$ $\forall$ $V(S)$ are calculated. Next, we calculate the average shortest path length connecting any vertex *v* in sample size n_S with the remaining n_S -1 vertices. Finally, the characteristic path length *(L)* can be assumed to approximate the median average distance L_{median} of all vertices $v \in V(S)$:

$$L \approx L_{median} = \underset{v,j \in V(S)}{median} \frac{d(v,j)}{n_S - 1}$$

The *clustering coefficient* (*C*), capturing the peculiarity that an individual's friends often know one another, is quantified by using the concept of neighborhood. The clustering coefficient indicates what proportion of the neighborhood of vertex *i*, ($\Gamma(i)$) is adjacent to each other. Let the number of edges in the neighborhood of vertex *i* be $|E(\Gamma(i))|$. Each vertex *i* has a number of edges, or degree . The total number of possible edges in $\Gamma(i)$ is $\binom{k_i}{2}$. Since we wish to calculate the clustering coefficient, we are only interested in 'friends' of vertex i that know each other. We therefore include only those vertices in the neighborhood that have at least two edges: $i \in V'(G) \equiv \{i \in V(G), k_i \geq 2\}$. The clustering coefficient of vertex i, formulated by Watts and Strogatz (1998) and Watts (1999a, 1999b) is:

$$C(i) = \frac{|E(\Gamma(i))|}{\binom{k_i}{2}} \quad \forall i \in V'(G)$$

Using the undirected graph condition, $|E(\Gamma(i))| \ = \dfrac{\sum_{l \in V_i(G)} \sum_{h \in V_i(G)} e_{l,h}}{2}$. Then,

$$C(i) = \frac{\sum_{i \in V_i(G)} \sum_{h \in V_i(G)} e_{i,h}}{k_i(k_i - 1)} \quad \forall i \in V'(G).$$

$C(i)$ is the proportion of total actual connections between a pair of any vertex's neighbors relative to all the possible connections within vertex i's neighborhood. Analogously, the clustering coefficient *(C)* of the whole graph, or network, is the proportion of total actual connections within all vertices' neighborhoods relative to all possible connections between vertices in the graph. The value of *(C)* should be equal to the averaged *C(i)* for each vertex in the graph. The clustering coefficient (*C)*, can then be thought of as the probability that a pair of vertices in the graph will be connected directly, given that they share a mutual friend:

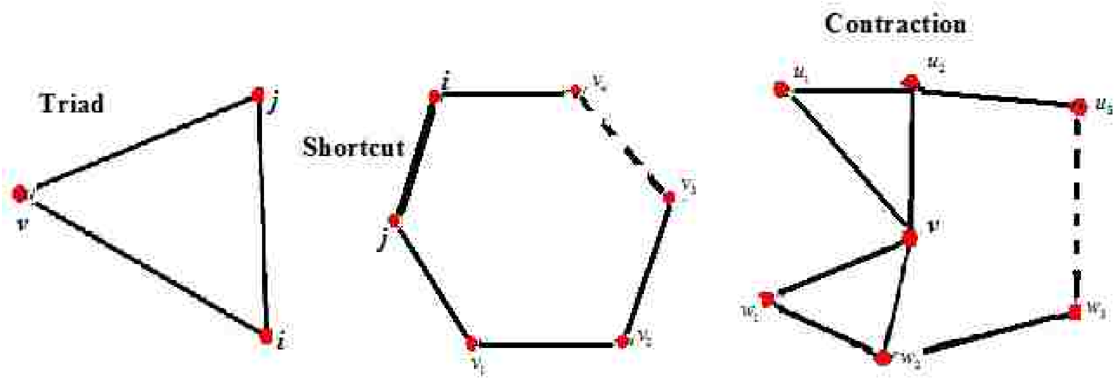
$$C = \frac{\sum_{i \in V'(G)} C(i)}{\sum_{i \in V'(G)} k_i(k_i - 1)} \quad \forall i \in V'(G)$$

C=0 would imply an empty graph and mean that no neighbor of any vertex i is directly connected to any other neighbor of i. In other words, no friend of mine knows any other friend of mine. On the other hand, C=1 would imply a complete $\frac{n}{k-1}$ subgraph, or *clique*, in which every member knows every other member directly.

The *range* of an edge $(R_{i,j})$ is the shortest path length between i and j in the absence of the edge (i, j). In other words, if there exists a direct link between any two vertices, $R_{i,j}$ is the second shortest path between them. Watts (1999a) called an edge *(i, j)* as *r-edge* if its range equals *r*. In the case where *R(i , j)* = 1 exists and *R(i , j)* = 2 is formed, the new path *(i, j)* tends to complete a triad in a highly clustered graph with strong ties. *R-edges* with *r>2* do not necessarily complete triads. However, since they indirectly connect more widely separated vertices, a direct connection between the same pair would be called a *shortcut*. In Figure 1.9b, vertices *i* and *j* do not connect to any mutual vertex and $r \geq 7$. The 1-edge *(i, j)* would then be a shortcut. It allows vertices *i* and *j* to meet directly instead of going through the sequential edges $e_{j,v_1}, e_{v_1,v_2}, ..., e_{v_n,i}$.

If two vertices *u* and *w* share vertex *v* as mutual friend, and the second shortest path length between them (other than the 2-edge through *v*) meets the condition $d_v(u,w) > 2$, then Watts (1999a) said that *u* and *w* are *contracted* by *v*, and the pair *(u,w)* is a *contraction*. In Figure 1.9c, both vertices u_1 and w_1 connect directly to vertex *v*. Although they are not directly connected to each other, they can connect indirectly through a number of vertices (such as u_2, w_2, and u_3, w_3). But those edges pass through at least two other vertices and have $r \geq 3$. Because the 2-edge through *v*, their only mutual friend, is the shortest path between u_1 and w_1, *v* is considered as their *contractor*. There is no shortcut, in Figure 1.9 c, connecting the pair of vertices (u_1, w_1), although they share a common friend *v.* A triad does not occur between these three vertices, which makes the *contractor* possible.

On the other hand, multiple triads are formed between vertices u_1, u_2 and v and between vertices w_1, w_2 and v. The pair (u_1,w_1) is not directly connected to each other but they are indirectly connected through other vertices in multiple triads via their only one mutual friend and contractor v. The set of vertices u and w is called a *contraction*. The number of vertex pairs that are not directly connected to each other but can connect through a common friend can be measured in order to investigate the role of contraction in a graph.



a. b. c.

Figure 1.9: Triad, shortcut and contraction compared

The fraction of all vertex pairs in the network that are not connected directly and have one mutual neighbor (contractor) relative to their neighborhood's possible edges, given by $\frac{k(k-1)}{2}$, is denoted by Ψ. A high value of Ψ means more contraction in a graph. For a graph G, with maximum possible number of edges $(M = \frac{k \cdot n}{2})$, the fraction of the edges which perform as shortcuts relative to the total edges M is denoted by ϕ. Its value is indicative of the role of shortcuts in the graph. The higher the value of ϕ, the more edges perform as shortcuts.

Ψ and ϕ will both be useful in graphical representations of a small-world network. These two variables make the study of small-world network properties easier to understand because they behave independently of the parameters used in model construction. At the same time, they are useful indicators of change within the network model. When one more shortcut occurs in the graph, reflected in an increase of ϕ, the clustering coefficient is decreased because one triad is deleted. The clustering coefficient, then decreases linearly with the increase in ϕ.

If a shortcut occurs between two very remote vertices, it can reduce dramatically the characteristic path length L in the entire graph. Thus, a small increase in ϕ may lead to a nonlinear reduction in L. The intensity of the small-world phenomenon, prevalent networks with high clustering coefficients and low characteristic path lengths, is then dependent on a high shortcut index ϕ.

However, increase in the number of shortcuts is not the only change that can bring about reduction of the characteristic path length. If a decrease is observed, it may also be due to any two vertices being connected via a new mutual friend, or contractor. A sparse graph can become denser through contraction without a significant effect on the clustering coefficient. The term Ψ is just as essential in analyzing the phenomenon of small-world networks.

Let us observe how contraction without shortcuts can take place within the structure of a social network. Each cluster in a social network consists of highly-interconnected, dense group of friends. But relationships between clusters may be rare. The distance between different clusters may be spanned by a single shortcut to another cluster. Alternatively, a bundle of edges from many individual members of one cluster may connect to a single vertex in another cluster as their common friend. Since many triads are formed, this connection is not a shortcut, but a contraction leading to an increase in Ψ. For example, when two strangers meet, a shortcut between their families is formed. But, if these two people get married and the bride meets everyone in the groom's family, without their families having a chance to directly meet each other, then she becomes a contractor between members of the groom's family and her own (Figure 1.10).

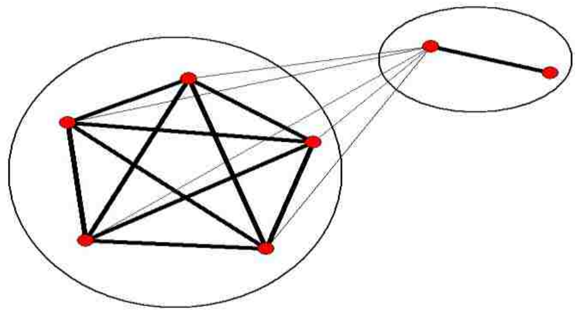


Figure 1.10: Edge connecting one single vertex in one clique to all vertices in another clique.

Contraction can be considered a function of shortcut formation if shortcuts are created between two clusters already connected by contraction. Using the same example of marriage between two families, we can picture that the family of the bride meets the groom, and any member of the bride's family may directly meet a member of the groom's family at the wedding. The bride's brother (*v*), who is too shy to meet the groom's family directly, has *k-b* edges connecting him with other members of his family. When they meet the groom's people directly, a number of shortcuts (*b*) are formed to the groom's family. The connections of the bride's shy brother to the groom's family are expressed as contractions *b(k-b)*.

If the bride's family members (n_b people) create shortcuts to the groom's family, the contraction will spread widely to $n_b b(k - n_b b)$ contractions between the two families, and $\frac{n_b b(n_b - 1)b}{2}$ contractions within their own cluster. The sum total of contraction in this case is given by $n_b b(k - n_b b) + \frac{n_b b(n_b - 1)b}{2}$. The bundle of shortcuts is comprised of many individual shortcuts ϕ reflected in each person's degree. Therefore, $n_b b = \phi k$. The total contraction can be computed in term of shortcut as $\frac{k(2k-b)\phi-(k\phi)^2}{2}$. Therefore, the fraction of contraction over the neighborhood's possible edges (Ψ) occurring when shortcut bundle is formed, is calculated over total contraction over the possible number of edges in the neighborhood $(\frac{k(k-1)}{2})$ as $\Psi(\phi)=\frac{(2k-b)\phi-k\phi^2}{k-1}$. The contraction Ψ, thus, can be calculated in terms of the shortcut variable ϕ. When shortcuts are created within a network, vertices can be directly connected, and the number of indirect contractions decreases.

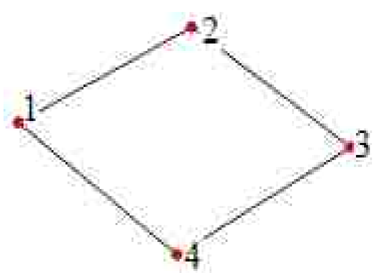
The initial topology of a graph, in which vertices and links between them are defined, is called a *substrate*. An initial graph can be a *tree substrate*, a *star substrate*, etc. The example in Figure 1.11a is a *one-lattice* graph of a *k*=2 substrate creating a *ring substrate*.

A *lattice* graph is a highly *regular*, *unweighted*, *undirected* and *simple* graph in which every vertex *i* joins with only a few neighbors. A *d-lattice* graph refers to a d-dimensional Euclidian lattice graph. For example, a one-dimensional lattice graph, a two-dimensional lattice graph and a three-dimensional lattice graph are set of vertices arranged in a straight line, a square, and a cubic, respectively. A 1-lattice with *k=2* is a one-dimensional structure in which each vertex connects to two other vertices in a row. The edges in this particular lattice graph form a *ring*. A 2-lattice with k=4 is a two-dimensional square grid (Figure 1.11b). This kind of lattice graph is sometimes called a *grid* graph.

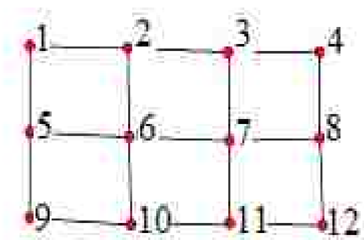
The characteristic path length and clustering coefficient of a 1-lattice graph with k≥2 can be calculated using the following equations:

$$L = \frac{n(n+k-2)}{2k(n-1)}$$
$$C = \frac{3}{4}\frac{(k-2)}{(k-1)}$$

The characteristic path length and clustering properties of any 1-lattice graph are thus characterized by the following relationships: for sparse graph the length scales linearly increases in the increase in graph size and decreases with respect to degree, but the clustering coefficient is independent of size of graph and is independent of degree when degree is large ($C \to \frac{3}{4}$).



a. 1-lattice with *k*=2



b. 2-lattice with *k*=4

Figure 1.11: d-lattice graphs

The 1-lattice is a good example of a completely *ordered* graph — one in which there is ordering of connections between the actors in the graph. For instance, if three individual actors n_i, n_j, n_k, where $i<j<k$, are ordered so that n_i is adjacent to n_j, and n_j is adjacent to n_k, but n_i is not adjacent to n_k then ordering is said to be incomplete. In Figure 1.11b, the 2-lattice with $k=4$ is one in which vertex 6 directly connects to the vertex 7 and the vertex 7 directly connects to the vertex 11, but vertex 6 does not directly connect to vertex 11. However, in the one-dimensional lattice network of Figure 1.11a, vertex 1 is adjacent to vertex 2, vertex 2 is adjacent to vertex 3, vertex 3 is adjacent to vertex 4, and the ordering is completed by vertex 4 adjoining to vertex 1. This is by far the simplest example of a completely ordered network. The opposite of the completely ordered graph is called a *random* graph, in which connections might be formed

between any vertex and any other vertices. A graph in which connections are incompletely ordered is also considered random, to a degree.

1.4 Models of a Small World

A network can be modeled as either completely or incompletely ordered. Real-world cases, such as technological networks, economic networks, social networks, or natural sciences networks, appear to fall somewhere between these two categories. Representations of small-world networks in the real world can occur at a possible topology between a completely ordered and a random graph.

Many social network theories use the idea of "social space." The adjacency metric was constructed by measuring the distances of existing between people in the network. This concept is extremely problematic due to difficulties in definition of the space and the metric. Previous methods of characterizing human values and relationship are not consistent, and can become rather complicated. To avoid such problems, Watts and Strogatz (1998) and Watts (1999a) developed a simple, comprehensible, and sustainable social network model. It is based on the following assumptions:

1. All networks can be represented only in terms of the connections between their elements.
2. All elements are identical.
3. All edges are equal and symmetry (i. e., lending themselves to representation in an unweighted and undirected graph such as the ones above).
4. A new edge can be formed by the influence of the already existing pattern of edges. In other words, people gain new acquaintances through the introduction of a mutual acquaintance.

These assumptions yield the simplicity of the model because it is not necessary to deal with the slippery question of space and metrics. We need to concern ourselves only with relations between actors, or vertices in the graph. The third assumption greatly simplifies the model in comparison to a real-world network. A social network falls somewhere between the completely connected network where everybody knows everybody else (sometimes called a *caveman world*), and the other extreme (known as a *Solaria* world), where everyone is a stranger and new friendships occur by chance. In the real-world, a social network obviously integrates many different circles of friends. Within the circle, most people are interconnected, but interactions across circles can be relatively rare. Most people in the real world have a chance to meet people from other groups through the connections of a friend (contractor).

They explored the model of a small-world network, in which everyone can possibly meet or know everyone else directly or indirectly via a short distance. In their model clustering is high and disorder is introduced by randomly rewiring a completely ordered network. Their fourth assumption relieves them from having to burden the small-world model with explanations of how people in the real world can make a new friend or create a new connection. They found that rewired networks are highly clustered, like the ordered networks from which they had been derived, but at the same time have a small characteristic path length like random networks. In order to be able to follow their reasoning, we need to know more details about the completely ordered, random, and small-world models, presented side by side.

1.4.1 Completely Ordered Models

Watts (1999a and 1999b) proposed a regular network in which each vertex has the same number of edges (same degree) and adjoins a small number of neighbors in a highly clustered pattern. In a small, caveman community, my two friends cannot help but know each other well. Throughout the cave we cohabit, everyone knows everyone else — it is a completely connected cluster. The regular network proposed by Watts is a caveman world. In it people become acquainted only through the introduction of one or more mutual friends. When they share no friends, the propensity to become connected directly is very small. Once they have only one mutual friend, the probability of becoming connected becomes suddenly high and plateaus there regardless of how many additional mutual friends they may acquire. Figure 1.12 illustrates this feature. The two axes represent the ordered extreme and the random extreme network model. The scale gives mutual friends as a fraction of total friends. For either extreme, the propensity to become friends starts from near zero, rises suddenly, and settles at one.

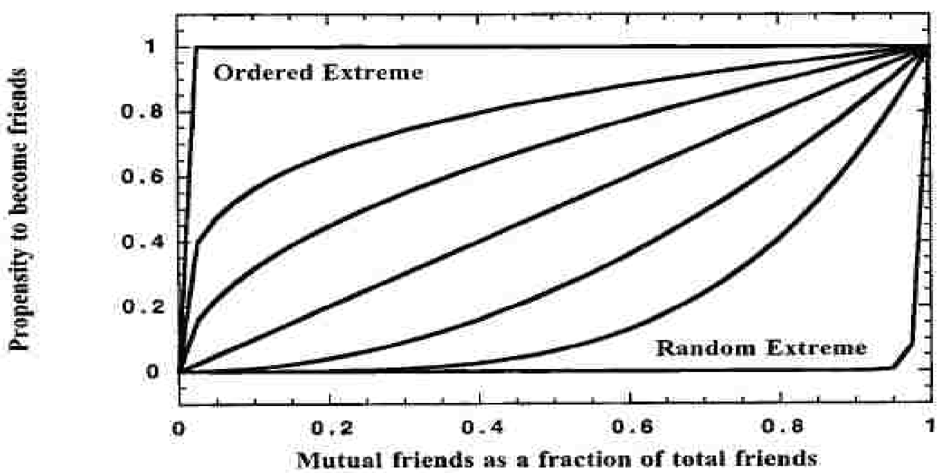
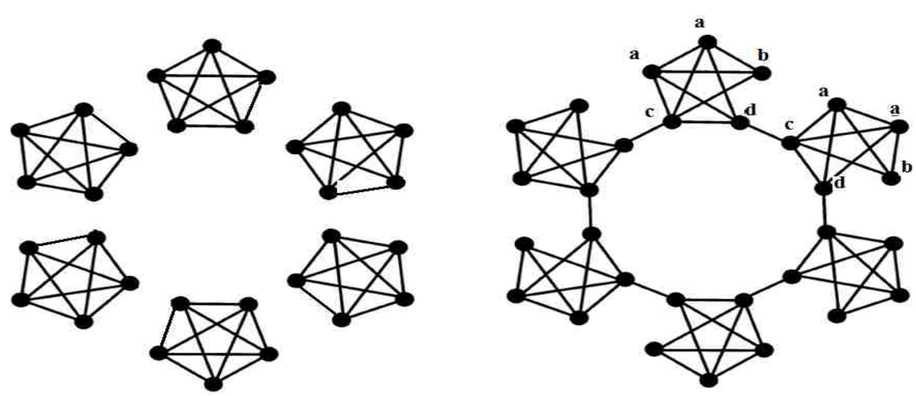


Figure 1.12: Real-world networks lie between two extremes
Source: Watts (1999b, p.504)

In Watts' caveman model, where the total number of people is n, each person has k degrees and each cave has $k+1$ people, $n_{local} = k + 1$ vertices. The network consists of many isolated $\frac{n}{k+1}$ caves, or *cliques*, or $n_{global} = \frac{n}{k+1}$ clusters. The clustering coefficient of each cave in the caveman model is one $(C_C = 1)$ because everyone in the same cave knows each other. Hence all edges in a cave form the multiple triads, and there are no shortcuts in a cave or edges between the caves (Figure 1.13a). The network is locally dense and globally sparse because, in the entire graph, the number of people is much greater than the average degree, $n >> k$. The caveman model does not satisfy the completely connected condition because it has more than one component. Watts turned this graph into a *connected-caveman* model by extracting one edge from each cave and using it to connect to a neighboring cave, such that all disconnected caves eventually form a single, unbroken loop. The connected-caveman model, a less extreme example, is used as a benchmark for ordered, highly clustered networks. Each cluster consists of n_{local}, which is the number of vertices in same cluster, and n_{global}, which is the number of clusters – their values are still the same as in the caveman world. However, in the connected-caveman world, not every vertex has the same degree k because certain vertices in each cave have developed connections external to their locale:



a. Caveman world   b. Connected- Caveman world

Figure 1.13: Caveman and connected-caveman worlds
Source: Watts (1999a, pp. 103-104)

In Figure 1.13b, each cluster of the connected-caveman world has two vertices (type *a*) of k degree connecting only to vertices within the cluster, one vertex (type *b*) of k-1 degree connecting only to vertices within the cluster, one vertex (type *c*) of k+1 degree connecting to vertices both inside and outside the cluster, and one vertex (type *d*) of k degree connecting to both vertices inside and outside. The clustering coefficient measures the ratio of directly connected members of one's neighborhood to the possible connections between them. Since members of one cluster now have different degrees, each vertex will yield a different clustering coefficient. Let C_a, C_b, C_c, and C_d denote the clustering coefficients of people type *a*, *b*, *c*, and *d*, respectively. The clustering coefficient in the connected-caveman world, denoted as C_{CC}, is a weighted average of the coefficients of *k*+1 vertices, given by:

$$C_{CC} = \frac{1}{k+1}((k-2)C_a + C_b + C_c + C_d)$$

Since one edge is missing from vertex *a*'s neighborhood:

$$C_a = (\binom{k}{2} - 1) / \binom{k}{2} = \frac{2}{k(k-1)}(\frac{k(k-1)}{2} - 1),$$

All k neighbors of vertex b connect to each other: $C_b = \binom{k}{2} / \binom{k}{2} = 1,$ Edges *k* are missing from vertex c's neighborhood, c has k+1 neighbors:

$$C_c = (\binom{k+1}{2} - (k+1)) / \binom{k+1}{2} = \frac{2}{(k+1)k}(\frac{(k+1)k}{2} - (k+1)),$$

$$\text{and} \quad C_d = (\binom{k}{2} - (k-1)) / \binom{k}{2} = \frac{2}{k(k-1)}(\frac{k(k-1)}{2} - (k-1))$$

because edges k-1 are missing from the neighborhood of vertex d. The final result is:

$$C_{CC} \approx 1 - \frac{6}{(k^2-1)}, \text{and} \quad C_{CC} \to 1 \text{ as } k \to \text{large}$$

It can be verified easily that, if each vertex has sufficiently high degree, the connected-caveman world will have a very high clustering coefficient (C_{CC} approaching one).

When the caveman graph was modified into a connected caveman graph, shortcuts were introduced. The effective local degree, k_{local}, is the average number of connected edges which are part of triads within each cluster and not shortcuts. The effective clustering degree, $k_{cluster}$, quantifies how many of vertex *i*'s neighbors are connected to another neighbor of *i*:

$$k_{local_{CC}} = k - \frac{2}{k+1}$$

$$k_{cluster_{CC}} \approx k - \frac{4}{k}$$

The average degree of each vertex within each cluster in the connected-caveman model $(k_{local_{CC}})$ is less than the average degree of each vertex in the caveman world (k). Similarly, the average effective clustering degree in the connected-caveman model $k_{cluster_{CC}}$ is less than in the caveman world (k).

The clustering coefficient, which is the proportion of actual direct connections in the neighborhood of each vertex to all possible connections, can be calculated in terms of k_{local} and $k_{cluster}$ as follows:

$$C = \frac{2}{k(k-1)} \frac{k_{local}(k_{cluster}-1)}{2}$$

$$C = \frac{k_{local}(k_{cluster}-1)}{k(k-1)}$$

High clustering means a great number of non-shortcut connections between vertices. As we have already seen, the clustering coefficient in the caveman world of Figure 1.13a, in which everyone has degree k, is equal to one ($C_C = 1$). By comparison, the clustering coefficient of the connected-caveman model (C_{CC}) is less than and very close to one ($C_{CC} < 1, C_{CC} \to 1$) because its k_{local} and $k_{cluster}$ are less than *k*. In other words, if in each cluster all vertices have mutual friends, then there is a large number of dense edges that cannot be shortcuts. In addition, the equations above suggest that the clustering coefficient is based on localized clusters in a network. This makes it very different from the characteristic path length, which tells us about the "global" structure because it measures the ability of actors to reach any other actor in the network.

Let us examine closely the characteristic path length in a connected-caveman model. The network consists of many connected clusters. Let L_{local} stand for characteristic path length within a single cluster, and L_{global} for the characteristic path length between clusters in the entire graph. Let d_{local} represent the average distance between pairs of vertices in the same cluster, and d_{global} the average distance between two vertices who are members of different clusters. Since most of the vertices are directly connected to each other, L_{local} is equivalent to d_{local}. For the degree is far greater than one $(k >> 1)$, $d_{local} \approx 1$. d_{global} is influenced by L_{local} and L_{global} because these quantify the path of a vertex traversing from one cluster to connect to a vertex in another cluster. When $n >> k >> 1$, d_{global} roughly approximates to $\frac{n}{2(k+1)}$. In the connected-caveman model, the number of vertex pairs connected in the same cluster is denoted by N_{local}, the number of pairs connected across clusters is denoted by N_{global}, and the total number of pairs of vertices connected in the network is *N*. N_{local} is calculated as the number of pairs connected in one cluster multiplied by the number of clusters in the network. N_{global} quantifies the number of edges which can traverse from one cluster to another cluster. As such, it represents the number of clusters that connect to each other, as well as the number of vertices with connections outside their own cluster. The total number of edges in the network (*N)* is measured as the sum of the number of edges within each cluster, N_{local}, and the number of edges connecting clusters, N_{global}.

$$N_{local} = \frac{(k+1)k}{2} \cdot \frac{n}{k+1} = \frac{n \cdot k}{2}$$

$$N_{global} = \frac{\left(\frac{n}{k+1}\right) \cdot \left(\frac{n}{k+1} - 1\right)}{2} \cdot (k+1) \cdot (k+1) = \frac{n(n-k-1)}{2}$$

$$N = N_{local} + N_{global} = \frac{n(n-1)}{2}$$

It can be seen that the number of edges in a cluster, N_{local}, is equal to the possible number of edges and the total number of edges in the connected-caveman graph is equal the maximum number of edges in the fully connected graph or complete graph.

The average distance between all pairs in a connected caveman graph (L_{CC}) is given by:

$$L_{CC} = \frac{1}{N}\left[N_{local} \cdot d_{local} + N_{global} \cdot d_{global}\right]$$

$$L_{CC} = \frac{2}{n(n-1)}\left[\left(\frac{n \cdot k}{2}\right) \cdot 1 + \left(\frac{n(n-k-1)}{2}\right)\left(\frac{n}{2(k+1)}\right)\right] \approx \frac{n}{2(k+1)} \approx d_{global}$$

It has been observed that characteristic path length (L_{CC}) of the connected-caveman model approximates the distance between one vertex in one cluster and any other vertex in a different cluster (d_{global}). In other words, when clusters are traversed by a single edge as in Figure 1.13b, L_{CC} is dominated by d_{global}. Hence, the characteristic path length is said to be reflexive of the global structure of the graph. For $n >> k >> 1$, L_{CC} is large, and L_{CC} increases linearly with *n*.

1.4.2 Random Models

Almost all of random graph theory analyses employ two models G(n, K) and G(n, p) along with the techniques proposed by Paul Erdos and Alfred Renyi in the 1950's and 1960's (as cited in Watts, 1999a and Newman, 2002). G(n, K) signifies a graph with n vertices and a number of randomly chosen edges K. G(n, p) signifies a random graph with n vertices in which the probability of random edges is expressed by p. According to Newman (2002), the most commonly used random graph is G(n, p) which consists of n vertices joined by links placed uniformly at random. Every one of the possible edges in the random model $\left(\frac{n(n-1)}{2}\right)$ may occur with independent probability $0<p<1$. G(n, p) is, then, a labeled graph of n vertices in which each edge appears with probability independent of any other edge. Since the average number of edges in the graph is $\frac{n(n-1)p}{2}$, and since each edge connects two vertices, the average degree of a vertex is given by $k = \frac{2n(n-1)p}{2n} \approx np$, for large n.

Let us compare this model to the random graph denoted as G(n, K), which is a labeled graph of n vertices having K randomly chosen edges. Edges are independent of any other edges in the G(n, p) model whereas the number of randomly chosen edges in a G(n, K) model is fixed. Although it is more convenient to use G(n, p) in random graph analyses, in practice, G(n, K) and G(n, p) are proposed to be interchangeable (Watts 1999a, p.40). Because $K \cong np \approx k$, once the size of graph (n) is known, any random graph built by specifying p, can be also specified by k.

There is evidence that the degree distribution among vertices in a random network is quite different from the degree distribution in real-world networks (Barrat and Weigt 2000, Newman 2002, 2004, and 2005). As noted by Newman (2002, p. 4), in an Erdos and Ranyi random graph with the exact degree *z*, the binomial distribution gives the probability of a vertex as follow:

$$p_z = \binom{n-1}{z} p^z (1-p)^{n-1-z}, \qquad \text{where } n >> k >> 1$$

However, in graphs with very large *n*, the probability distribution will become a Poisson distribution, $p_z = \frac{k^z e^{-k}}{z!}$, where $n >> kz$. Both distributions are strongly peaked about the mean *k*, and fall off rapidly as they move away from the peak value. The probability of a vertex having a large *k* is negligibly small.

Barabasi and Bonabeau (2003) observe that the connectivity in the World Wide Web network is dominated by a few highly connected web pages. They argue that the network's degree distribution follows the power law. Newman (2001, 2002, 2004, 2005) and Zlatic et al. (2006) also found that many networks in the real world, such as the network of scientific collaboration, web hits, copies of books sold in America, Wikipedia, etc., appear to follow a power law distribution (with some deviations). This degree distribution is reflected in a linear tail appearing in log-log scales of cumulated distribution. The power law distribution gradually decays due to a small number of non-negligible vertices having very large degrees.

In a random network, each vertex is connected to vertices existing anywhere in the network by chance. Insofar as no member has any propensity to connect to anyone in particular, the random network can be thought of as a "Solaria."[6] In this extreme case of a network, people can form new friendships with a stranger without regard for any mutual friends they may or may not already share. Figure 1.12 indicates that, within this kind of network, the propensity to become friends relative to the number of existing mutual friends curves near zero and stays near

[6] Watts (1999a, p. 44)

zero up until all of a pair's friends are mutual. At that point, the curve suddenly increases to one. This is so because within this kind of network everyone is a stranger to one another and there is no increased chance to become friends unless everyone's friends have mutual friends.

In Erdos and Renyi random graph, the probability of a pair of vertices becoming connected by an edge is higher if they have a mutual friend than if they do not have any mutual friends. Since the clustering coefficient (C), as defined by Watts and Strogatz (1998), can be thought of as the average probability that two neighbors of a given vertex are also direct neighbors, the clustering coefficient in the random models (C_{random}) can be viewed as[7]:

$$C_{random} = p \approx \frac{k}{n}$$

Obviously, under the sparseness condition $(n>>k)$, C_{random} is very small. Random models do a poor job of capturing the clustering property of networks. Since the clustering coefficient of a connected caveman network (C_{CC}) is very large and approaches the unity $(C_{CC} \to 1)$, C_{CC} is certainly greater than C_{random}.

$$C_{CC} >> C_{random} \approx \frac{k}{n}$$

For any random graph that has no single largest component and none of whose vertices are adjacent to each other, distance and characteristic path length are meaningless because there is no path between any pair of vertices. However, for a random graph that has a largest component, the distance and the characteristic length can be measured. Erdos and Renyi have expressed the characteristic length of a random graph whose *n* vertices are detached from any vertex within distance with an average degree *k*. In the limits of large *n* and *k*, the asymptotic

[7] As discussed and simplified by Newman (2002, p.2)

approximation of the characteristic path length is $L_{random} \sim \frac{\ln(n)}{\ln(k)}$.[8] For any $n >> k >> 1$, the characteristic path length is small.

Comparing random networks to connected caveman networks, we note that the characteristic path length in the former increases logarithmically with the graph size *n*. In the latter, it increases linearly. Also, for a connected-caveman model with large *n*, the more *n* is greater than k, the larger the characteristic path length is: $\left(L_{CC} \approx \frac{n}{2(k+1)} \right)$. If we were to compare the characteristic path length of these two extreme models, we would have:

$$L_{random} \sim \frac{\ln(n)}{\ln(k)} << L_{CC} \approx \frac{n}{2(k+1)} \qquad \text{for any } n >> k >> 1$$

This comparison reveals that the characteristic path length in random networks increases rather slowly. As *n* becomes larger, they lag further behind connected-caveman networks of the same size, whose characteristic path length grows much faster.

Random models, which are locally disordered graphs, are characterized by a low clustering coefficient and, especially with large *n*, short characteristic path length. The characteristic path length in the random models is invariably much smaller than in same-sized connected-caveman models, which are highly ordered graphs with high clustering coefficients and large characteristic path lengths.

In order to approximate the characteristic length in any random model, Watts (1999b) employed the Moore graph instead. Let S be the number of vertices included within a distance D (diameter). The shortest path between any vertex and any other within S must be shorter than D and is given by $\sum_{d=1}^{D-1} d \cdot k(k-1)^{d-1}$. The number of vertices that can reach others in D steps by

[8] Newman (2002, p.9)

traversing through vertices belonging to S is: n-S-1. Therefore, the summation of all distances from any vertex to any other vertices in the Moore graph is: $\sum_{d=1}^{D-1} d \cdot k(k-1)^{d-1} + (n-S-1)D.$ The average of the distances between all vertices, called characteristic path length, must be averaged over n-1 vertices. Hence, the computation of the characteristic path length in the Moore graph (L_{MG}) is:

$$L_{MG} = \frac{1}{n-1}\left[\sum_{d=1}^{D-1} d \cdot k(k-1)^{d-1} + (n-S-1)D\right]$$

$$L_{MG} = D - \frac{k(k-1)^{D}}{(n-1)(k-2)^{2}} + \frac{k(D(k-2)+1)}{(n-1)(k-2)^{2}}, \text{ when } k>2.$$

Watts (1999a) claimed that even though the degree of vertex in the Moore graph is not exactly same but increases in every steps, the Moore graph can be a good approximation of a pseudo-random graph, i.e., a random graph created by constructing a ring of n vertices and randomly adding edges until there is a total of K edges. One important difference between a pseudo-random and a random graph is that, at $k \approx \ln(n)$, the characteristic length of the former is distinguishable from the latter. Therefore, it seems reasonable to use a Moore graph construction to approximate the characteristic length of a random graph where $k \geq \ln(n) >> 1$. When the average degree is greater or equal to $\ln(n)$, but far greater than one, the characteristic path length of the random graph (L_{random}) can indeed approximate the characteristic path length of Moore graph (L_{MG}).

$$L_{random} \approx L_{MG}, \quad \text{where } k \geq \ln(n) >> 1$$

1.4.3 Application of Small-World Models

In the real world, people become acquainted not only through introduction by a mutual friend, but also by bumping into complete strangers. A relationship can be formed both by recommendation and accident. Small-world networks, in which people can be connected by a brief chain of acquaintances, like the one suggested by Milgram (1967), or the network of movie actors, are commonplace. As Figure 1.12, real-world networks lie somewhere between the completely ordered and random network models. How does one go about modeling a real-life small-world network more precisely?

Watts and Strogatz (1998) examined a possible small-world network by using what they called a "random rewiring procedure." They started rewiring a ring lattice (one-dimensional lattice) with n vertices and k degrees, by replacing the near-neighbor edges with edges to randomly selected vertices throughout the network, chosen uniformly with probability p. They measured alterations in both characteristic path length and clustering coefficient as the rewiring probability increased. They claimed that the small-world phenomenon occurs when the probability of random rewiring increases from $p=0$ to a value between $0<p<1$. At $p=0$, there is no random rewiring and the lattice is unchanged. With $p=1$, the lattice is transformed into a completely random model, much like the Erdos-Renyi one (see Figure 1.14). With a probability of random rewiring near zero ($p\rightarrow 0$), the network is highly clustered and characteristic path length grows linearly with n because $L \sim \frac{n}{2k} >> 1$. The clustering coefficient would then be $C \sim \frac{3}{4}$. At probability of random rewiring set near unity ($p\rightarrow 1$), the network is poorly dense and L increases logarithmically with n because $L \approx L_{random} \sim \frac{\ln(n)}{\ln(k)}$. The clustering coefficient then becomes $C \approx C_{random} \sim \frac{k}{n} << 1$. The small-world phenomenon, thus, when cast into graph theory, occurs in a sparsely connected graph that exhibits a characteristic path length close to that of an asymptotic approximation of the random graph $(L \approx L_{random})$ with a much higher clustering coefficient $(C >> C_{random})$.

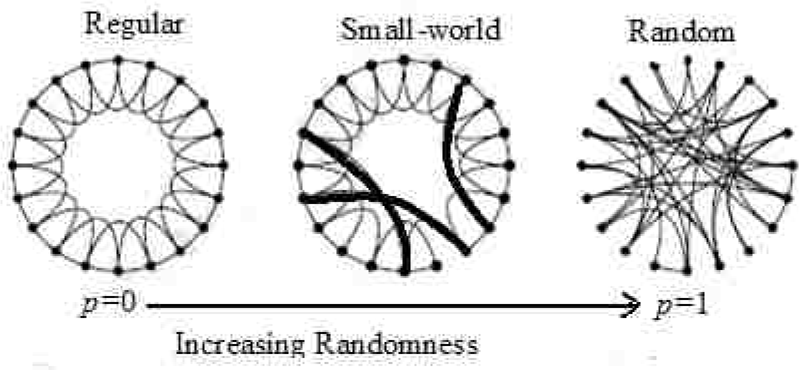


Figure 1.14: Random rewiring across the spectrum of probability
Source: Watts and Strogatz (1998, p. 441)

To conform with the small-world phenomenon, a network must be wired in such a way that any two vertices can be connected to each other by just a few links. For this property to occur, the following conditions (modified from Goyal et al., 2004) must be satisfied:

1. $n >> k >> \ln(n) >> 1$, would meet the condition of sparseness; $k>>1$ guarantees that the network is connected.
2. The network must be connected or must have a giant component for the characteristic path length to be measured.
3. $L \approx L_{random} \sim \frac{\ln(n)}{\ln(k)}$. That is, L must be almost as small as L_{random}.
4. *L* must be on the order of ln(*n*) and should increase logarithmically with *n,* as in a random network, since there is some degree of randomness in a small-world network.

Even though the characteristic path lengths of some small-world networks are greater than those in their random counterparts, their value must be comparatively low.

5. $C >> C_{random} \sim \frac{k}{n}$, satisfies the condition that the clustering coefficient of a small-world network is much greater than that of a random one.

Watts and Strogatz (1998) tested this small-world model using the collaboration graph of film actors in Hollywood (n=226,000 and k= 61), the power grid of the western United States (n= 4,941 and k= 2.94) and the neural network of worm C.elegans (n= 282, k=14). The characteristic path lengths of these three networks are 3.65, 18.7, and 2.65, respectively. The clustering coefficients are 0.79, 0.80, and 0.28, respectively. The researchers compared the characteristic path length (L_{actual}) and clustering coefficient (C_{actual}) for these three networks to a Erdos-Renyi random model built theoretically with the same number of vertices (n) and the same average number of edges per vertex (k). L_{random} of the Erdos-Renyi networks were reasonably close, while the clustering coefficients were much lower than in the real-life networks (Table 1.1). Watts and Strogatz concluded that because all three networks have $L_{actual} \geq L_{random}$ and $C_{actual} >> C_{random}$, they satisfy the conditions of the small-world property. Similarly, many other real-world networks analyzed by Newman (2002) can be shown to have actual clustering coefficients much higher than the magnitudes calculated from mirroring random models.

Table 1.1: Small-world vs. random networks

	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005
C. elegans	2.65	2.25	0.28	0.05

Source: Watts and Strogatz (1998, p. 441)

Amaral et al. (2000) tested the statistical properties of a variety of diverse real-world networks for occurrences of the small world phenomenon. They proposed three classes of small-world networks: *scale-free, broad-scale*, and *single-scale*.

Scale-free networks are those whose degree distribution follows the power law no matter what their scale is. In other words, the shape of the degree distribution curve will follow the power law form (decaying slowly from a straight line on a log-log plot) regardless of the scale of measurement.

Broad-scale networks are characterized by a connectivity distribution that obeys a power law until a sharp cutoff occurs. This type of small-world networks exhibits typical power law distribution on a small to medium scale. However, when large scales are reached, degree distribution is truncated. Figure 1.15 illustrates the log-log plot of a cumulative distribution in a broad-scale network. Initially, it manifests Power law decay but is truncated to Gaussian decay of tail when the number of connections becomes large. The movie actor network is considered a broad-scale network because it behaves this way when the number of collaborations exceeds a certain limit.

Finally, single-scale networks are characterized by a connectivity distribution with a rapidly decaying tail. Degree distributions in this class of small-world networks do not follow a power law; instead, they decay much faster. In Figure 1.15, the tail of connectivity distribution in log-log plots of the electric power grid of southern California, the network of world airports, and the neuronal network of the worm C. elegans, are shown to decay exponentially, faster than if governed by a power law. The log-log plot of degree distribution in the friendship network of high-school students (Figure 1.15) seems to conform to a Gaussian distribution. Thus, all of the above examples are classified as single-scale networks.

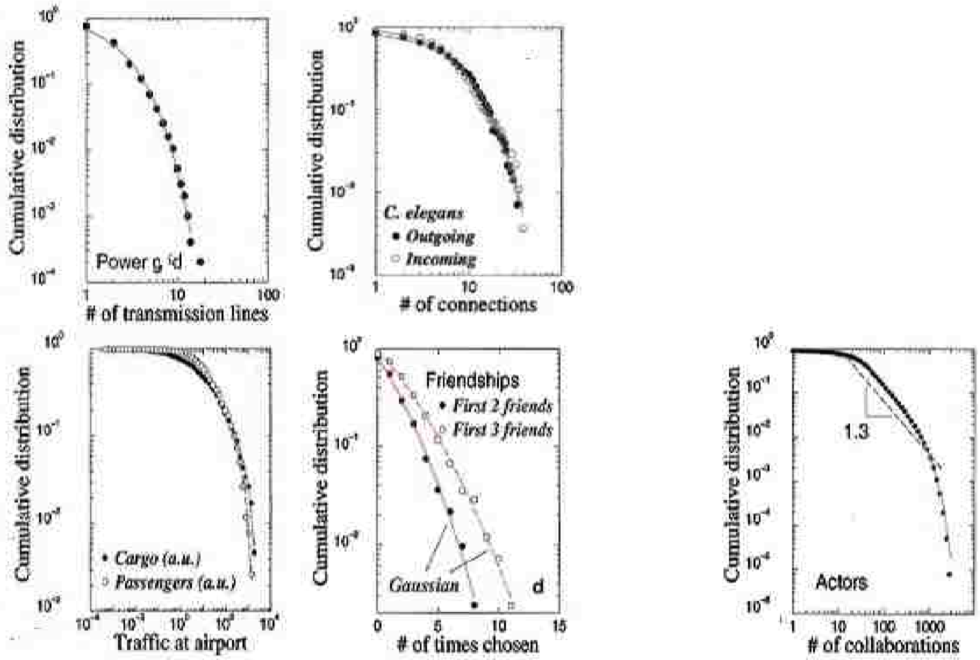


Figure 1.15: Classes of small-world networks with degree distributions on a log-log plot. (Electric power grid, C.elegans neural network, Airport traffic, and School Friendship are all single-scale networks, while the movie actor network is broad-scale.)

Source: Amaral et al. (2000, pp. 11150-51)

Amaral et al. (2000) found two effects hindering these small-world networks from achieving a power law distribution of connectivity. One is the effect of *aging* vertices. A very highly connected vertex will eventually stop creating a new connection. An example from the Hollywood network would be a retired actor. Another effect is the cost of adding new links to a vertex, especially the physical cost or the limited capacity of a vertex to host new links. An example from the network of world airports would be space and time constraints that prevent certain airports from becoming a hub to more airlines.

1.5 The Relational Model and the Small World Phenomenon

A graph constructed by interpolation between a completely connected and an incompletely connected, or random, graph, can be built using two different construction algorithms. Based on which algorithm was used as discussed by Watts (1999a and 1999b), the graph can belong to one of two categories: relational and spatial. The characteristic path length and clustering coefficients of both categories are determined only on n and k. However, the construction rule for a relational graph depends on distances between vertices that are computed only in terms of the network structure. New edge construction in this category is the function of preexisting edges.

In a spatial graph, on the other hand, new edge construction is a function of physical distance in the graph. Thus, in general, it can be said that the small world phenomenon can occur only in a relational graph. A spatial graph, in which vertices are constrained by physical distance to access only nearby neighbors, cannot reach remote members easily. The following section will examine the relevance of the relational graph to representations of the small-world phenomenon.

1.5.1 The Relational Graph

To analyze the small world phenomenon, three following types of relational graph proposed by Watts (1999a and 1999b) can be constructed. First is the α-model, which is formed as an imitation of a real-world social network. Second is the β-model, which generalizes the α-graph by stripping inessential properties of social networks. Third is ϕ-model, which combines the observed properties of the α- and β-models. Let us see how the features of a small-world network—high clustering and low characteristic path length— are manifested in those three models.

1.5.1.1 The α-Model

An α-model is constructed in order to include the character of connections in a social network. They have some level of order, but are neither completely ordered, as in the caveman world, nor incompletely ordered (or unordered) as in a random network. Because the level of 'orderedness' in social networks falls between these extremes, they both need to be investigated. In a relative network construction algorithm, α is used as a parameter controlling for the probability of connection (0≤*p*≤1) between any vertex and any other in the network, (0≤*α*≤∞). At α=0, all edges are formed in completely ordered. Increasing α, the propensity of connecting between any two vertices increases. This is why the result is called an α-model. To construct an α-model, the two extreme models are interpolated by a random rewiring procedure in the form of a graph with specified *n*, *k*, and *α,* and set p to a very small value.

Rewiring will take place after the probability that any vertex connects to any others $(P_{i,j})$ is computed. Interpolation between the completely connected model and the random model is done by first fixing a vertex *i* and calculating $R_{i,j}$ which is a measure of vertex *i*'s particular propensity to connect to vertex *j*. The $R_{i,j}$ can be computed as follows:

$$R_{i,j} = \begin{cases} 1, m_{i,j} \geq k \\ \left[\dfrac{m_{i,j}}{k}\right]^{\alpha}(1-p)+p, k > m_{i,j} > 0\ p \\ p, m_{i,j} = 0 \end{cases}$$

where $m_{i,j}$ is the number of mutual friend of vertices *i* and *j*, k is average degree of the graph, p is a baseline probability for edge (*i, j*) to exist ($p << \binom{n}{2}^{-1}$), (meaning that p is set when *n* is specified), and *α* is an adjustable parameter, 0≤*α* ≤∞. $R_{i,j} = 0$ if vertex *i* and vertex *j* are already connected. Next, the probability of connection between *i* and *j* is obtained by: $P_{i,j} = R_{i,j} / \sum_{l \neq i} R_{i,l}$. $P_{i,j}$ ranges over the interval (0, 1). Then, they use $P_{i,j}$ in a uniform random process to connect vertices until the number of edges $M = ((k \cdot n)/2)$ has been constructed. If the value of α is changed, the value of $P_{i,j}$ is changed as well. The α-graph is constructed by increasing the value of α from 0 to 20.

The α-model was analyzed through a process of computer-based numerical experiments. For small α, the propensity of connection $R_{i,j}$ is dominated by p because the edge between vertices i and j could be formed only by the baseline random opportunity (p is set to be very small but non zero). When two consecutive edges are formed, two vertices now have a mutual friend and friendships in the network can start to expand. Potential triad formation now has a role in connecting network elements.

As α progresses towards infinity (α→∞), all probabilities of connections, $P_{i,j}$, will go to the uniform random limit value p showing that a large-α-graph would look like a random graph. In this α-model, the existing average edge degree k determines new link formation and clustering during the rewiring process. The number of vertices connected to the same component, thus, grows linearly only with respect to the average degree k. This might violate the sparseness condition. A violation will not occur if the graph is disconnected. However, characteristic path length cannot be measured in a disconnected graph so that point is moot. In order to avoid the sparseness violation problem and ensure that the connectivity condition is met, a connected substrate must be built before processing the rewiring construction algorithm. The ring substrate is preferred as the basic topology of this model. Unlike star, tree, and path substrates, which have centers, roots, and end points, respectively, the ring substrate has no 'special' vertices and no more than the necessary edges to ensure a connected graph. The ring substrate can control for the sparseness violation problem even in the range of a very large α.

To identify the moment when the small-world phenomenon appears, Watts and Strogatz (1998) and Watts (1999a and 1999b) built a low-lattice with a ring substrate and constructed α-models with n=1000 and k=10, $p = 10^{-10}$, and 0≤α≤20. Then, they measured L with respect to α ($L(\alpha)$) and the clustering coefficient with respect to α ($C(\alpha)$). The result was that, at a large value of α, networks had short characteristic path length but still high clustering coefficient. The transition of the clustering coefficient from low to high marked the emergence of the small-world phenomenon. Figure 1.16 indicates that, as α increases, L(α) increases to a maximum and then drops rapidly to its asymptotic random model limits. The length would approach to the random limit when α≥11. There exists a class of graphs in the intermediate region of α, where the clustering coefficient is high but the characteristic path length is small and equivalent to random graphs. (high C and $L \sim L_{random}$). This is the class of small-world networks. The breakdown of length scaling and clustering in an α-model is as follows:

1. For α=0, both L(α) and C(α) are large. For fixed n and k, an increase in α causes both L(α) and C(α) to increase until they reach their maximum values at the small value of α.. When α is small, the properties of L(α) and C(α) are best explained with a 1-lattice graph.
2. For intermediate α, once both L(α) and C(α) reach their maximum value, or *cliff*, they drop rapidly to small values, then lean towards the asymptotic limit. However, the clustering cliff clearly occurs after the length cliff. In this region of α, there exists the high clustered graph with low characteristic path length. This disparity makes the small-world phenomenon possible.
3. For large α, L(α) decreases logarithmically with n and approaches the random model limits of $L_{random} \sim \dfrac{\ln(n)}{\ln(k)}$ and $\lim_{\alpha \to \infty} C(\alpha) \sim \dfrac{k}{n}$. In other words, a random graph can match a large-α graph in both characteristic path length and clustering properties.
4. In addition, for small α, $L(n,k,\alpha)$ increases linearly with increasing in n. In the large-α region, $L(n,k,\alpha)$ unsurprisingly scales with n again. $L(n,k,\alpha)$ decreases with increasing in k at any level of α (both small or large). Moreover, as n becomes very large (in the order of

millions and billions of elements), there is distinctive difference in the characteristic path length of small-α and large-α ranges.

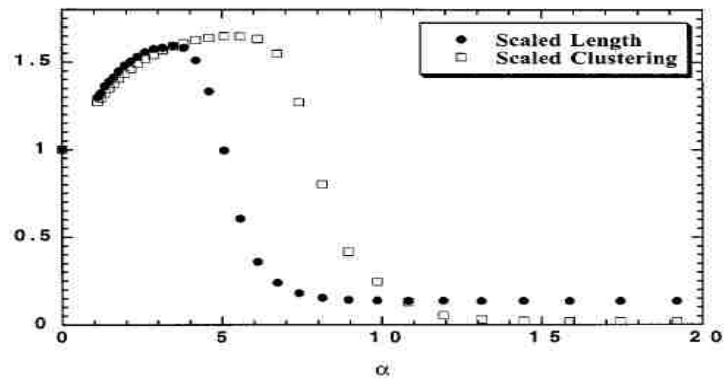


Figure 1.16: Length and clustering in α-models
Source: Watts (1999a, p.58)

For robustness check, Watts (1999a) built a number of other α-models, using as substrate a 2-dimensional lattice, a 2-Cayley tree substrate, a random substrate (in which edges are randomly linked until the whole graph is connected), and no-substrate at all (under the construction algorithm, this means considering only the value of α in order to obtain a connected graph). At low α, there are some discrepancies between these other kinds of substrates and a ring substrate. But when α is high enough, the length of an α-model built on a ring substrate and of any other kind substrates do not show any difference. That is, for sufficiently high α, all substrates yield similar length values and clustering coefficients. Although the length of no-substrate graph is most similar to the length of ring substrate graph, it is similar only when the graph is connected and the value of α is low. In addition, a ring substrate yields the most variation in lengths (compared to others, a fixed *n* and *k* ring substrate exhibits the maximum length). It is therefore easier to observe and distinguish the behavior of characteristic path length in ring substrates. As a result, the ring substrate is the preferred topology in α-models capturing the small-world phenomenon. However, due to some discrepancies between ring substrate and other kinds of substrate in the α-model (when the value of α is relatively low), there is a need to construct a more generalized model that is less complicated in terms of dependency on social motivation than the α-model. The β-model also uses a ring substrate as initial topology of the graph but differs in other important respects and affords more opportunities to theorize the small-world phenomenon.

1.5.1.2 The β-Model

Like the α-model, the β-model is an interpolation between the completely ordered and the completely random graph. It begins with a simple ring substrate stripped of social motivation such as mutual friends, acquaintance or other social circles. The building algorithm of a β-model starts with a perfect one-dimension lattice with *k* degree ([illegible] on either side) in which any chosen vertex *i* is connected to its nearest neighbor in a clockwise direction (*i*, *i*+1). In order to not duplicate connection in the graph in the rewiring process, a uniform distribution probability value *r* of one vertex *i* connecting to any other is initially generated. This probability is compared it to the probability of random rewiring β. If $\beta \leq r$, the edge (*i*, *i*+1) is unchanged, but if $\beta > r$, the edge (*i*, *i*+1) is deleted and uniformly randomly rewired such that vertex *i* connects to another vertex *j* with probability β over the entire graph. The value of β needs to be measured against the value of *r* because the rewiring edges cannot be self-connecting and cannot be the same as the initial edges. After a vertex *i* is considered, the procedure is redone for the nearest neighbor to consider vertex *i* and vertex *i*+2, vertex *i* and vertex *i+3*, and so on, until all edges in the graph are rewired.

The parameter β is the basic probability value used in the rewiring algorithm. As such, the parameter β can range from zero to one. The graph was observed at different values of β. When β=0, the 1-lattice graph does not change because all edges remain unchanged. When β=1, all edges are randomly connected. When $0<\beta<1$, the graph was changed into a random graph. The result, visible in Figure 1.17, indicates the occurrence of the small-world phenomenon at low lengths and high clustering coefficients. The β-graph captures the rapid transition of the characteristic length with respect to the probability of random rewiring *L(β)* and the clustering coefficient with respect to the probability of random rewiring *C(β).* As in the α-graph, all the 'action' occurs very close to β=0. The length cliff and the cluster cliff in the small-β region occur as in the α-model. However, in small β region, the length scale changes logarithmically with respect to *n*. In the α-model, for small α region, the length increases linearly with *n*. When the probability of random rewiring is high (β approaches unity, $\beta \rightarrow 1$), the characteristic length and clustering coefficient are transformed into the asymptotically random graph's values. The transition in characteristic length and clustering coefficient occurs at a value of β that is different from the value of α. Clustering, *C(β)*, still stays at high value after the length *L(β)* becomes close to L_{random}. In the large-*n* range, the existence of small-world graphs (sparsely connected, decentralized, clustered), seems clearer in the β-model than in the α- model.

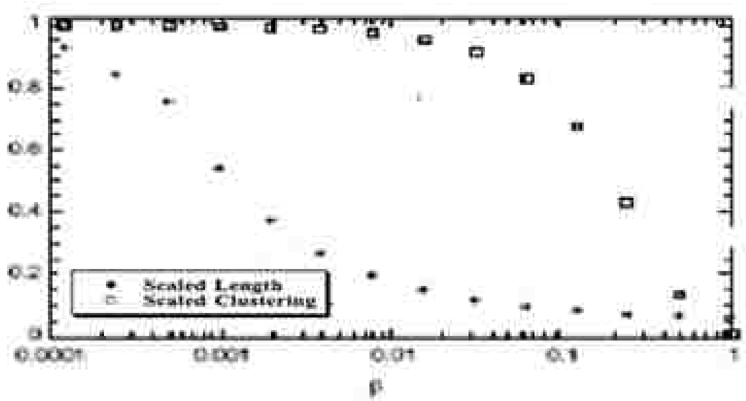


Figure 1.17: Characteristic Length and Clustering Coefficient as functions of β
Source: Watts (1999a, p70)

Although the β-model and α-model are built using different construction mechanisms, both exhibit similar data structure when approaching limiting values and when rapid transitions occur. Since there is no social motivation in β-model, it can be interpreted much clearer than the α-model. As such, the β-model is more adept at illustrating the small world phenomenon than the α-model. Regardless of the differences in these two models, there are parameters independent from network construction that makes the job of comparing them much easier. The remainder of this subsection compares the α and β models in term of shortcut, ϕ, and contraction, Ψ.

In a relational graph, the rewiring algorithm introduces shortcuts and contractions. Both are means of bridging the distance between remote vertices. The small-world phenomenon can be observed more clearly in terms of shortcuts and contractions than in terms of α or β values. Both the characteristic path length and the clustering coefficient, represented in terms of shortcut, ϕ, and contraction, Ψ, can well explain the small-world phenomenon. Of the two, shortcuts provide the simpler means of accounting for it. In other words, characteristic path length and clustering coefficient, when computerized, produce less complicated results in terms of shortcuts than in terms of contractions. Even though it is more precise to capture the small-world phenomenon by contraction, shortcut is more practical.

In rewiring, the first few shortcuts randomly introduced in a large graph connect two widely separated vertices, thereby making the distance between these two vertices much shorter. These first few shortcuts are likely to have as high impact as in a rapid drop of the graph’s characteristic length. However, when the average distance in a graph has become smaller, new shortcuts have much less impact on the characteristic length because at this time it has already decreased to its asymptotic value. The introduced shortcut reduces triads in the graph. The clustering coefficient, defined as our friends knowing each other, is also reduced as a result of decreased triad formation. In a highly clustered graph, the deletion of one edge from a triad has less impact on the clustering coefficient. Since the network rewiring algorithm starts from a completely connected network with a very high clustering coefficient, the clustering coefficient, then, remains high even when more shortcuts (the sufficiently small fraction of shortcuts) are created. Hence, the co-existence of low characteristic length and high clustering coefficient in a connected and sparse graph containing a sufficient small fraction (but not too small) of shortcuts is the environment where the small-world phenomenon occurs.

In the initial ring substrate of an α-model, when an additional edge was randomly added, it acted as a bridge connecting two distant parts of the 'rim.' This edge is classified as a shortcut because distant vertices on both sides of the substrate seem to be in communication through it even when they are not directly connected. In the low-β-model, the additional edge formed itself as a new real shortcut that connects formerly distant vertices. For very small ϕ, then, both characteristic path length in terms of shortcut ($L(\phi)$) and the clustering coefficient in terms of shortcut ($C(\phi)$) in the α-model perform in ways different from the β-model. Once again, the visibility of the small-world phenomenon is greater in the β-model. Watts (1999a) found that, with all algorithm constructions including a 1-dimension lattice ϕ-model, connected networks with a sufficiently but not too small ϕ can yield small length and high clustering coefficient as seen in Figure 1.18:

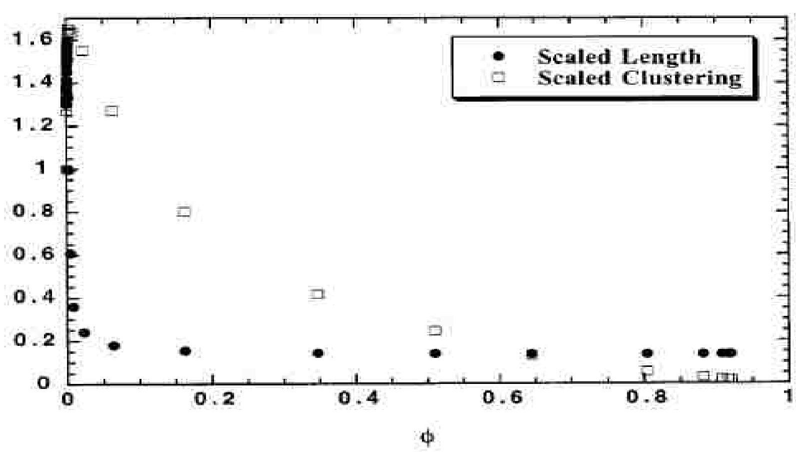


Figure 1.18: Characteristic length and clustering as functions of ϕ
Source: Watts (1999b, p. 510)

Like shortcuts, contractions are able to bring distant vertices closer together, by introducing a common neighbor, or mutual friend, bridging the gap. Instead of forming their own edge as a shortcut, two remote vertices may be contracted by a randomly added edge. This new edge can contract distance in the graph with very little impact on clustering. Hence, analysis in terms of the fraction of contractions, Ψ, gives a clearer picture of the small-world phenomenon than analysis in terms of shortcuts. Watts (1999) concluded that, in a connected and sparse graph, low characteristic path length can coexist with high clustering coefficient over a large region of Ψ. This domain of small-world graphs is outlined in Figure 1.19.

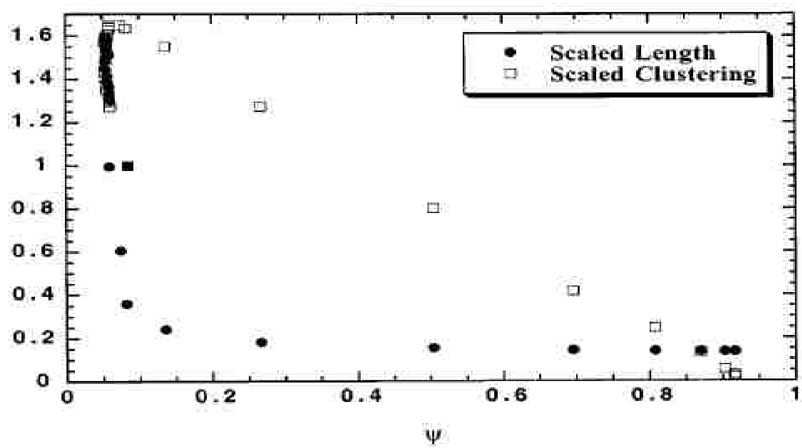


Figure 1.19: Characteristic length and clustering as functions of Ψ
Source: Watts (1999b), p. 513.

1.5.1.3 Transition in the Relational Graph

The transition in this relational graph occurs when edges in the highly ordered graph, such as the connected-caveman network, are removed from one cluster to randomly connect vertices in different clusters. The random rewiring process shifts the local edge in one cave to be global edge connecting to another cave, as shown in Figure 1.20. This process introduces shortcuts into the graph. The local length increases but the global length decreases.

It is obvious that the new edge in a randomly rewired graph does not traverse from cluster to cluster by the connection of type *c* and type *d* vertices in neighboring clusters as it does in the connected-caveman graph. Many edges traverse through clusters in the network that are far apart. Those edges transform into global links. When edges are shifted randomly from local to global scale, we cannot say that the characteristic length *L* is dominated by d_{global}. Rather, we should say that *L* is determined by both the length that a vertex traverses through its own cluster (L_{local}) and the length that it traverses globally to connect to a vertex in another cluster L_{global}.

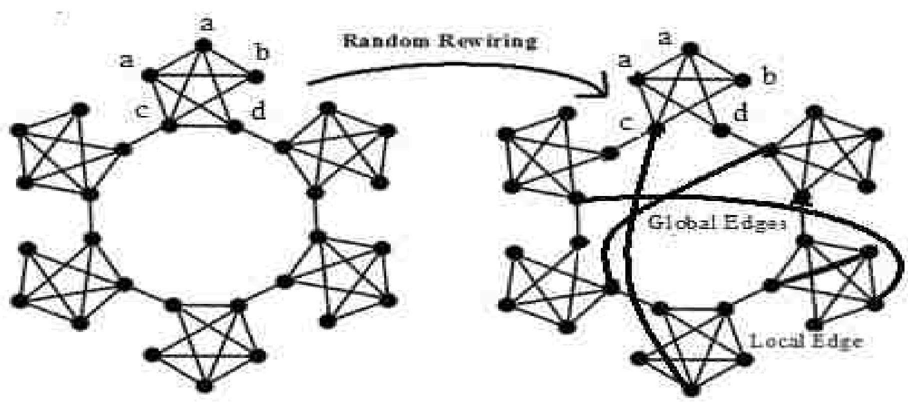


Figure 1.20: Randomly rewire the connected-caveman graph

After rewiring, due to the influence of shortcuts, the characteristic path length of the relational graph (L_r), with k>2, decreases nonlinearly and can be approximated by the length of the Moore graph (L_{MG}) where $L \propto \frac{\ln(n)}{\ln(k)}$. Hence, in a transitional graph, the characteristic path length could increase logarithmically with the size of (n) for ϕ>0. The increase in shortcuts leads to a linear decrease in both k_{local} and $k_{cluster}$. The clustering coefficient of a relational graph (C_r), measured in terms of shortcuts, can approximate $(1-2\phi)$. And, for any large k, and small ϕ, it is very high ([illegible]). This confirms the observation we made in section 5.1.3 that after randomly rewiring a completed graph, the first few new shortcuts have a highly nonlinear impact on the characteristic path length, $L(\phi)$, but linear impact on the clustering coefficient, $C(\phi)$. Thus, the small-world graph occurs at the introduction of these first few shortcuts.

1.5.2 The Spatial Graph

The construction rules for a relational graph presented above do not depend on any physical distance in the graph. This enables shortcuts and contractions introduced by the rewiring process to shrink distance throughout, without any constraint. The transitional process in a relational graph, thus, can exhibit the small-world phenomenon. However, if the graph construction algorithm was constrained by physical distances between vertices, a so-called spatial graph is produced. Spatial constraints in the construction algorithm prevent shortcuts or contractions from having significant influence in shrinking the characteristic length of the graph after rewiring. The small-world phenomenon is thus less likely to occur in a spatial graph. Still, an exhaustive study of small-world networks must review the spatial, as well as relational graph. Although global communications have affected the social networks of most societies, there still exist communities where physical distance between members is the crucial factor in forming relationships.

A spatial graph is one that considers physical distance as determined by a graph parameter, ξ, called the spatial distance parameter. In the process of random rewiring, new edges are created as a function of the probability distribution of the spatial distance parameter. In other words, in a construction of spatial graph, the probability of connection between two vertices is co-determined by ξ.

In the case of constructing a one-dimension graph using uniform distribution in rewiring edges, the physical distance between each vertex is ξ=w/2, where w is width of distribution. If normal distribution is chosen for the random rewiring algorithm, the spatial distance in the graph is $\xi=3\sigma$ (i.e., the connection will be made within $\pm 3\sigma$ of the initial vertex). Distribution of spatial graph presents a characteristic known as 'finite cutoff' because ξ is finite. Thus we ensure that new edges will occur only between vertices preexisting within the physical distance of ξ.

In a spatial graph, vertex *i* can never connect directly farther than the external length scale specified by ξ. However, if the external length scale is long enough, vertices in the entire graph may become connected. Even though shortcuts and contraction occur, edges can connect the vertices only within external length scale ξ. Before the length scale extends to be sufficient large, clustering in a spatial graph already decreases to its random limits. The characteristic path length as the function of external length displays in the same way as the clustering coefficient.

Spatial networks, which are constructed under the assumption of a uniform distribution, limit vertices within a ξ radius. Since all new edges only occur in local scale, $k_{local} = k$ and $k_{global} = 0$. As the spatial distance ξ increases, the number of vertices connected in local scale increase, and the number of vertices in the global scale decreases since $n_{local} = 2\xi + 1$ and $n_{global} = \frac{n}{2\xi + 1}$. Nevertheless, the increase in the size of graph (*n*) has a linear impact on n_{global}. Thus, L_{global} linearly rises when *n* increases. After random rewiring, the characteristic path length of a spatial graph is the product of local and global lengths, when the connection occurs outside the local area. However, the characteristic path length in a spatial graph, L_S, is most likely dominated by the global length. The characteristic path length in a spatial graph, L_S, thus increases linearly with the size (*n*), but decreases monotonically with increase of the physical distance (ξ).

When a spatial graph is randomly rewired, every edge remains local. Shortcuts and contraction have a small impact on the length of the entire graph because they connect only vertices within the fixed external length. They might contract a pair of widely separated vertices, but at a distance which is only a little further than the cutoff value (ξ). The clustering coefficient of a spatial graph, C_s, then, can be calculated in terms of ξ. As $k >> 1$ and $\xi >> 1$, the clustering coefficient of the spatial graph varies as $C_s \propto \frac{k}{\xi}$. The clustering coefficient of a spatial graph decreases monotonically with the physical distance ξ, as the characteristic path length does. In addition, for $\xi \approx n$, the clustering coefficient of spatial graph, C_s, approximates the random graph limit value ($C_S \approx \frac{k}{n}$). For ξ=(k/2), the clustering coefficient of a spatial graph is equal to the clustering coefficient of a one-dimension lattice graph ($C_S = C_{1-lattice}$), which is a high value.

Watts (1999a) compared the characteristic path length and clustering coefficient of a spatial graph with respect to the physical distance ($L_s(\xi)$ and $C_s(\xi)$). The result is that $L_s(\xi)$ and $C_s(\xi)$ are lined in the same pattern both in the numerical calculation and analytical approach. This means that, with the uniform distribution having a cutoff value, the small-world phenomenon of networks with low characteristic length and high clustering coefficients cannot occur in a spatial graph.

However, if the spatial graph is constructed using infinite variance, or Cauchy distribution, instead of uniform, or Gaussian distribution, it will generate a broad range of graphs with high C(ξ) and low L(ξ). We can say, then, that spatial graphs with infinite cutoff, can exhibit the small world phenomenon.

1.6 Summary

Many economic situations cannot be explained by market mechanisms, but can be well explained by the network of relationships. The structure of social networks is, then, an important key in explaining market interactions. The small-world network, first studied by Milgram (1967), describes the phenomenon that everyone in the world can be reached through a short chain of social acquaintances. Following his work there have been a number of discussions, both empirical and theoretical, examining the small-world phenomenon. This chapter studied the small-world model proposed by Watts and Strogatz (1998) and Watts (1999a and 1999b). Using techniques from graph theory, they interpolated between two extreme networks (completely connected and completely unconnected) by starting with a simple one-dimensional ring-lattice. Then, they randomly rewired some edges by a probability distribution and measured the so-called *characteristic path length* and *clustering coefficient*. They observed a sudden transition in the characteristic path length as the rewiring probability increased due to the role of the first few *shortcuts*. In contrast to the characteristic path length, the clustering coefficient remains high until a high rewiring probability is reached. The first few shortcuts have a significant impact in contracting the distance between two widely separated regions in the graph. Networks characterized by very small path lengths, yet high clustering coefficients, are said to exhibit the small-world phenomenon. Watts and Strogatz (1998) applied their model to real-world networks and identified the small-world phenomenon, as Milgram's social experiment did. The elements in the networks Watts and Strogatz studied empirically could communicate within even shorter chains than the "six degrees of separation" Milgram had sensationally proposed in 1967.

CHAPTER 2: DATA DESCRIPTION AND PRELIMINARY ANALYSIS

2.1 Introduction

Most studies of job search or job mobility in classical labor economics theory neglect the social interaction process through which workers acquire jobs and employers hire workers. Even though personal ability is justly considered one of the main determinants of job mobility, a number of researchers have stressed the key role of social relationships. Social networks have been shown to be crucial in influencing labor market transitions and their efficiency. Granovetter (1973) noted that most workers found their jobs through personal contracts and argued that social connections are the leading source of information about job opportunities. Montgomery (1991) examined the role of social connections by studying how they influence screening and matching. Relationships formed within and outside of one's workplace frequently help workers acquire new positions. This indicates the importance of network relationships in the social and economic operation of the labor market. The network structure connects different agents and defines the nature of interaction between them. Network formation depends on the strategic decision-making by its participants – by strategically deciding whom to for a relationships with participants can improve their future chances of landing a new job.

Empirical studies of networking in labor markets are hampered by data requirements. To properly examine the role played by social networks it is necessary to collect very detailed data where each agent's network relations are observed. To get full information on the network, such data are needed on every agent and on every relation connected to that agent. In extensive networks it is apparent that data requirements quickly preclude any empirical study. However, due to their particular nature academic labor markets are more conducive to studies of the role played by social networks. Unlike other labor markets, particularly those involving private companies, it is very difficult to easily observe employees and collect information on themselves as well as their social relations. Given their public character, universities provide plenty of information on their employees, particularly the academic staff. Various types of information on faculty are published by universities themselves in their catalogs or on university web pages. In addition, faculty themselves typically post their own resumes on their web pages. One can then observe where they obtained their training, their specialization, their publication record, and their employment history (provided such information is included in resumes). Given the wealth of readily available information, academic labor markets are good candidates to help assess the importance of networking in labor markets.

Networking is a very important determinant of finding an academic job upon completion of the Ph.D. degree. One of the standard advices new Ph.D. students receive is to talk to their advisors and determine who they know in universities which are advertising open positions and to ask them to contact those they know. This is a classic example of how networking is used in finding new jobs. In academic labor markets one can then examine the network created by interactions among universities. When Harvard University employs a graduate of Princeton University it creates a relationship or link between the two which may facilitate the flow of information in both directions which can be used in future hiring decisions. By collecting information of faculty employed by various universities and where each obtained their degrees, one can examine the education-employment network between universities.

The goal of this chapter is to describe the education-employment network in the academic labor market for economists. The first section of this chapter will describe the data collected on two hundred departments of economics throughout the world. Next, we will offer a preliminary study of this data aimed at discussing the main trends and providing the first glimpse of the network between universities. Finally, a summary of our findings will be given.

2.2 Data Description

In order to investigate the economics labor market, the information on the top two hundred departments and their faculty was collected. Tom Coupe's (2003) ranking of departments was used to select the departments to be studied. Coupe's (2003) ranking is used as it was the most recent available worldwide ranking when data collection began. As discussed below, Coupe (2003) arrived at his ranking by using more measures than most other available rankings. This chapter and the next use the terms department, university, and institution interchangeably and all refer to departments of economics located at these universities. Also, to clarify the terminology, the top two hundred universities ranked by Coupe (2003) are referred to as the ranked universities. Since the list of ranked universities is by no means extensive and it is possible those universities hire graduates of universities which are not ranked, those universities are collectively referred to as unranked.

In the past, most economics department rankings, such as the National Research Council (NRC) and the US News and World Report (USNWR), tended to be concentrated on departments' self-reported statistics, a subjective method. While the NRC and USNWR rankings are still widely used, over the last thirty years there have been several attempts by economists to provide more objective rankings of economics departments. These methods all attempt to evaluate departments' scholarly output using various measures. Some rankings focus only on departments in the US, while others rank economics departments throughout the world.

Scott and Mitias (1996) rank US economics departments based on faculty publications. They use the Herfindal Index to rate the quality of journals. Similarly, Dusansky and Vernon

(1998) rank the top fifty US departments based on their publication records using several ways of measuring the impact of journals including Leband-Piette's (1994) impact factor, which rates journals using the average citation per published article. Kalaitzidakis et al. (2003), Coupe (2003), and Heck et al. (2006) are recent examples of worldwide rankings based on scholarly output. Both Kalaitzidakis et al. (2003) and Coupe (2003) ranked two hundred departments, while Heck et al. (2006) ranked 186. While Kalaitzidakis et al.'s (2003) ranking method was based on journal citation analysis, which values frequently cited journals more, and focuses on 30 journals to weight and count the faculty's publication pages, Heck et al. (2006) focused only on the elite eight journals, also referred to as the 'Blue Ribbon' journals.[9] Coupe (2003) used 11 different publication-based methods to arrive at his ranking. He adopted many count-weighted approaches including focusing on publications in a limited number of journals (top ten journals), a bigger set of journals (seventy-one journals), and an even larger set of journals (258 journals). The final ranking is an average of all 11 methods. The strong point of Coupe's (2003) methodology stems from the idea that a high ranking should result from all 11 criteria.

Methodology aside, the correlation of department rankings between Coupe (2003) and others is very high, especially for the top one hundred departments. The correlation of worldwide ranked department between Coupe (2003) and Kalaitzidakis et al. (2003) is 0.78 overall and 0.73 in the top one hundred departments. The correlation between Coupe (2003) and Heck et al. (2006) is 0.80 in both cases. Moreover, for North American departments, the correlation between Coupe (2003) and Kalaitzidakis et al. (2003) and between Coupe (2003) and Heck et al. (2006) is very high, at 0.85 and 0.87, respectively.

---

[9] The Blue Ribbon Eight include: the *American Economic Review*, *Econometrica*, *International Economic Review*, *Journal of Economic Theory*, *Journal of Political Economy*, *Quarterly Journal of Economics*, *Review of Economic Studies*, and *Review of Economics and Statistics*.

Table 2.1 presents the correlation of rankings between Coupe (2003) and other recent rankings, including Roesster (2004) which is based on co-authorship and output evaluation. The correlation between Coupe (2003) and Roesster (2004) in both methods is high as well, 0.75 with co-authorship, and 0.80 with output ranking. As can be seen from all correlation, most rankings of departments are similar, not only in their ranking, but also in departments which are ranked. The use of Coupe's (2003) rankings, rather than a different one, will not result in radically different conclusions, particularly since a department's rank itself is not used in the analysis. Rather, Coupe's (2003) rank is used solely to select which departments are studied and is shown in Table 2.6 along with other information to be discussed at a later point.

Data were collected during the 2005-2006 academic year from information published on university websites. Data include tenured and tenure-track faculty (assistant, associate, and full professors) for each economics department. It includes faculty with a terminal degree in economics. Faculty members with degrees in other fields are excluded. In addition, information on economists in other parts of the university, such as business schools, agricultural economics departments, public policy departments, and others were omitted. While this resulted in some prominent business schools with many economists (Chicago, Northwestern, University of Pennsylvania) being omitted from the study, collecting data on all economists employed by a single university would be much more difficult. It would entail combing the rosters of every unit within a university which could potentially hire an economist. While in certain cases such information is somewhat easily available (some business schools do provide faculty breakdown based on fields of specialization), in others it is more difficult to obtain it. In addition, since the main goal is to evaluate the labor market for academic economists, studying both production and

Table 2.1: Correlation between Coupe (2003) and Other Rankings

		Correlation Coefficient with Coupe (2003)
Christian Roesster (2004) Network Rankings	Worldwide (200)	0.75
	100 upper rank	0.75
	100 bottom rank	0.29
	North America	0.85
	100 upper rank	0.82
	100 bottom rank	0.25
Christian Roesster (2004) Average Productivity Rankings	Worldwide (196)	0.80
	100 upper rank	0.76
	100 bottom rank	0.43
	North America	0.87
	100 upper rank	0.82
	100 bottom rank	0.66
Kalaitzidakis, Stengos, and Mamuneas (2003)	Worldwide (200)	0.78
	100 upper rank	0.73
	100 bottom rank	0.26
	North America	0.85
	100 upper rank	0.82
	100 bottom rank	0.49
Heck, Zaleski, and Dressler (2006)	Worldwide (186)	0.80
	100 upper rank	0.80
	100 bottom rank	0.31
	North America	0.87
	100 upper rank	0.85
	100 bottom rank	0.37
Scottand Mitias (1996) Wide 36 Journals (concentrate)	100 US Departments	0.87
Scottand Mitias (1996) Core 5 Journal (Stock)	80 US Departments	0.17
Dusansky and Vernon (1995)	50 US Departments	0.65
NRC Faculty Survey in 1993	50 US Departments	0.83
USNWR Overall Survey in 1996	50 US Departments	0.81

placement, most units outside economics departments produce few economics Ph.D. Thus, their omission should result in a small bias.

Information on each individual includes the university which granted the terminal degree, the current employer, and professorial rank. These three pieces of information are the minimum required for the analysis. For future use, where available the year of graduation, year of employment by current university, first employer and year, as well as fields of specialization were collected as well. The department which employed the individual when data were collected is referred to as ‘employer,’ while the department which the individual graduated with the terminal degree to as the ’grantor’ or ‘producer.’

The sample consists of two hundred employer universities, of which 126 are located in North America, 57 in Europe, 7 in Asia, and 4 in Australia. The total number of individuals in these departments is 5,530, of which the minimum required information (employer, grantor, and rank) is available for 5,081 individuals (91.88 percent), which is the size of the sample analyzed. The 449 economists for whom the required information is not available are all employed by universities in Europe, Australia, and Asia. Complete information is available for every academic economist employed in North America. A total of 321 universities granted terminal degrees to economists in the sample. Almost 98 percent of faculty with the minimum required information are Ph.D. degree holders in economics. The remaining 109 individuals do not have a Ph.D. as their terminal degree. With this in mind, we will generally refer to Ph.D. as the terminal degree for all individuals.

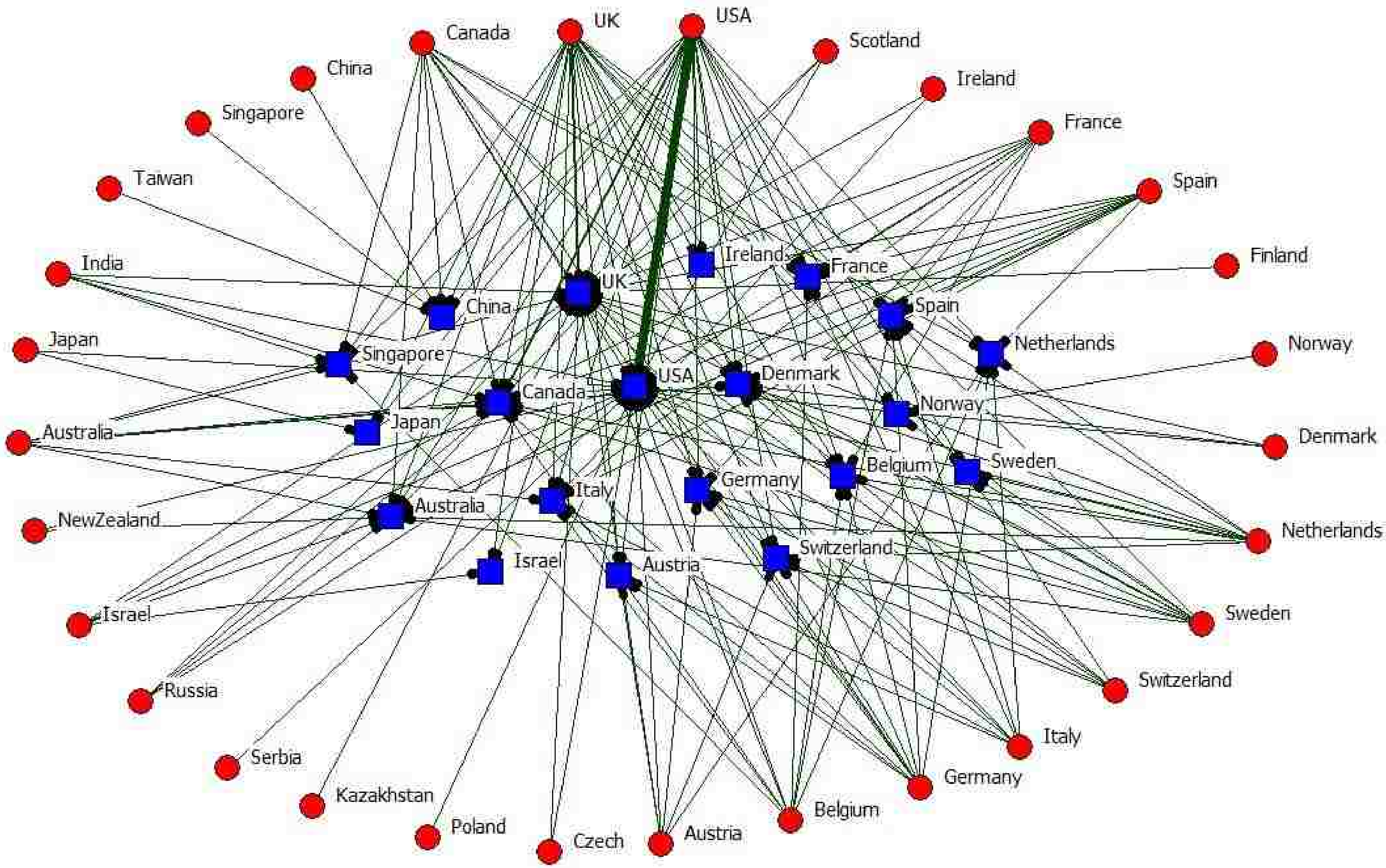
2.3 Preliminary Analysis

Table 2.2 presents the number of degrees granted and individuals hired by the countries in which the employer and grantor universities are located. The United States produces and hires by far

Table 2.2: Grantor and Employer Countries in Ranked Universities

Grantor Country	Number of Graduates	Percentage	Employer Country	Number of Hires	Percentage
Australia	78	1.54	Australia	166	3.27
Austria	19	0.37	Austria	20	0.39
Belgium	76	1.50	Belgium	52	1.02
Canada	240	4.72	Canada	485	9.55
China	1	0.02	China	94	1.85
Czech	2	0.04	Denmark	42	0.83
Denmark	38	0.75	France	120	2.36
Finland	1	0.02	Germany	79	1.55
France	141	2.78	Ireland	3	0.06
Germany	98	1.93	Israel	56	1.10
India	8	0.16	Italy	125	2.46
Ireland	4	0.08	Japan	109	2.15
Israel	27	0.53	Netherlands	114	2.24
Italy	63	1.24	Norway	31	0.61
Japan	64	1.26	Singapore	42	0.83
Kazakhstan	1	0.02	Spain	146	2.87
Netherlands	102	2.01	Sweden	49	0.96
NewZealand	2	0.04	Switzerland	33	0.65
Norway	25	0.49	UK	521	10.25
Poland	1	0.02	USA	2794	54.99
Russia	4	0.08	Total	5081	100.00
Scotland	3	0.06			
Serbia	1	0.02			
Singapore	5	0.10			
Spain	74	1.46			
Sweden	57	1.12			
Switzerland	21	0.41			
Taiwan	1	0.02			
UK	529	10.41			
USA	3395	66.82			
Total	5081	100.00			

the largest number of academic economists, producing a total of 3,395 economists in the sample (66.82 percent) and hiring 2,794 (54.99 percent). The UK is a far second with 529 economists produced and 521 hired. Canada is a close third. Figure 2.1 presents links between employer and



Grantor Countries Employer Countries

Figure 2.1: Interactions between Countries as Employers and Grantors

grantor countries[10]. Employers are located in 20 and grantors in 30 countries. Each employer and each grantor is represented by a node. If a country both hired and produced economists it is represented by two nodes. Hiring countries are concentrated in the middle of the figure, while grantors are on the edges. The direction of the arrow points in the direction placement of produced economists. An arrow pointing from France to Germany means that a French Ph.D. was hired in Germany. There is no arrow in the opposite direction, meaning no German trained scholar was hired in France.

Figure 2.2 shows interactions between continents, as both employers and grantors. The thicker line between North America and Europe indicates more interactions between these two continents compared to others. Information illustrated in Figure 2.2 is shown in Table 2.3 and 2.4.

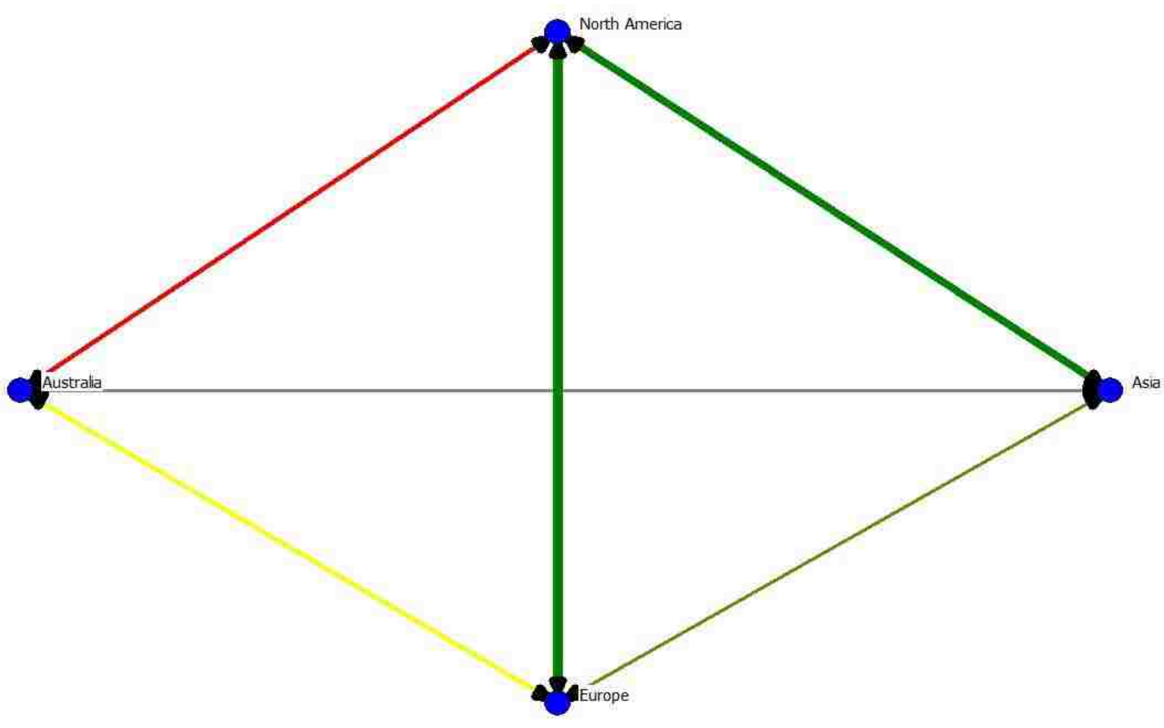


Figure 2.2: Interactions between Continents as Employers and Grantors

[10] This, and all sequent network figures as well as calculations were obtained by using the software package Ucinet 6.

Rows show the total number of degrees granted on each continent, while columns show the total number of economists employed on each continent. In both tables, 'N' indicates the actual number of economists. The only difference between the two tables is that Table 2.3 shows row percentages indicating the placement distribution across continents, while Table 2.4 shows the column percentages indicating the hiring distribution across continent. Take North America, for example. It produces a total of 3,635 economists, of whom 3,052 (83.96 percent) are hired by

Table 2.3: Distribution of Economists in Ranked Universities (Row Distribution)

			Employer				Total Granting
			Asia	Australia	Europe	North America	
Grantor	Asia	N =	88	1	4	13	106
		Row %	83.02	0.94	3.77	12.26	
	Australia	N =	4	60	6	10	80
		Row %	5.00	75.00	7.50	12.50	
	Europe	N =	13	28	1,015	204	1,260
		Row %	1.03	2.22	80.56	16.19	
	North America	N =	196	77	310	3,052	3,635
		Row %	5.39	2.12	8.53	83.96	
	Total	N =	301	166	1,335	3,279	5,081
		Row %	5.92	3.27	26.27	64.53	

Table 2.4: Distribution of Economists in Ranked Universities (Column Distribution)

			Employer				Total
			Asia	Australia	Europe	North America	
Grantor	Asia	N =	88	1	4	13	106
		Column %	29.24	0.60	0.30	0.40	2.09
	Australia	N =	4	60	6	10	80
		Column %	1.33	36.14	0.45	0.30	1.57
	Europe	N =	13	28	1,015	204	1,260
		Column %	4.32	16.87	76.03	6.22	24.80
	North America	N =	196	77	310	3,052	3,635
		Column %	65.12	46.39	23.22	93.08	71.54
	Total Employment	N =	301	166	1,335	3,279	5,081

North American universities, 310 (8.53 percent) are hired by European universities, 77 are hired by Australian universities, and 196 by Asian universities. The 3,052 North American trained economists hired by North American universities represent 93.08 percent of all economists hired by North American universities. North America hired 204 economists educated in Europe who represent 6.22 percent of all economists hired by North American Universities. Asia and Australia jointly account for less than one percent of academic economists in North America.

This chapter, as well as the next one, will focus on analyzing the North American labor market for several reasons. First, as already mentioned there is complete coverage of North America. Full information is available for every economist employed by the 126 universities in North America, while some information is missing for some 450 economists in the other parts of the world. This is in part due to some of these universities not providing as much information on the English language versions of their web pages as they do in their native languages. North American departments tend to provide more information on their faculty than do their counterparts in Europe, Asia, and Australia. Second, departments outside of North America have a much higher tendency to hire their own graduates than do the North American departments (though exceptions exist on both sides). This may be a result of smaller labor markets in which departments operate and may not truly reflect the networking aspects. Third, the academic labor market in North America is very centralized, facilitate primarily through the efforts of the American Economic Association. While universities and graduates outside of North America participate in the North American market, their participation is limited relative to their numbers (both employment and number of economists produced). Labor markets on other continents are not as centralized. In addition, geographic distances and borders between countries outside of North America are likely to create these fragmented markets and result in different institutional settings in these labor markets. However, to give a broader context and to assist further research, a description of the network of all ranked universities will also be presented.

In addition to these two samples, the North American and full sample, two more samples must be created. Since the goal is to analyze the labor market from a network point of view, it is necessary to create 'square' samples, where each department functions as both the employer and the grantor. Due to this restriction, the square samples will consist only of ranked departments worldwide and the relevant subsample of ranked departments within North America. For the few universities which do not grant Ph.D. degrees, such as Dartmouth College, granting activity is set to zero (i.e., they only employ economists, but do not place any).

Summary statistics for the various samples are provided in Table 2.5. The first column provides information for the entire network of all ranked universities. The faculty in ranked departments received their degrees from 321 departments. The second column provides information on the square version of the full network – keeping only ranked departments whose graduates were employed by ranked departments (all other economists and universities are dropped). Only 179 of the 200 top universities have granted degrees to 4,783 faculty members currently employed by the top 200 universities. They form 94.14 percent of all hires. Table 2.6 shows the hiring and granting information for all ranked departments. Almost all top 10 universities hire faculty with degrees from a ranked institution (UC Berkeley is the only exception).

The average number, the median, and the standard deviation of Ph.D. hires from all grantors are not significantly different than hires from ranked grantors only (column 2 of Table 2.5). The average faculty size of a ranked economics department is 25.41, and average number of faculty hired from ranked departments is 23.92. That means, on average, ranked departments

Table 2.5: Summary Statistics for Ph.D. Graduate Employment in Economics Departments

		Full Network	Full Square Network	North American Network	North American Square Network
		Ranked Employers and All Grantors	Ranked Employers and Ranked Grantors	North America Ranked Employers and All Grantors	North America Ranked Employers and North America Ranked Grantors
1	Number of Employing Universities	200	200	126	126
2	Number of Ph.D. Granting Universities	321	179	193	108
3	Total Number of Faculty	5,081	4,783	3,279	3,026
4	Average Faculty per Employing University	25.41	23.92	26.02	24.02
5	Median of Ph.D. Faculty per Employing University	24	22	25	22
6	Standard Deviation of Faculty per Employing University	12.05	11.42	10.63	9.81
7	Average Number of Placements per Ph.D. Grantor	15.83	26.72	16.99	28.02
8	Median Number of Placements per Ph.D. Grantor	3	12	3	9
9	Standard Deviation of Number of Placements per Ph.D. Grantor	35.48	44.61	36.50	45.66
10	Number Hired from Non-Ranked Ph.D. Grantors	298		71	
11	Number Hired from Non-North America Ranked Ph.D. Grantors			182	
12	Percentage of Hiring from Ranked Universities	94.14		92.28	

employed only 1.5 faculty members trained by a non-ranked department. The median number of faculty per employer does not differ much from the mean. However, the average, the median, and the standard deviation of the number of placements per granting department in the full network is much different than those in the square sample of all ranked departments. The network formed by ranked universities whose alumni have been hired by another ranked institution is called the *square ranked network*. While the average number of placements per grantor in the full network is 15.83, the average number of placements per grantor in the square network is 26.72 individuals. Such differences are due to the few economists trained by the many unranked departments (298 economists received their terminal degree from 121 unranked departments). This indicates there is an unequal throughput for producing Ph.D. graduates. In other words, the ranked universities have the ability to produce more Ph.D. holders. By the same token, differences in program capacity may help explain why the standard deviation in the square ranked network is higher than that in the full network.

Table 2.6 presents the role of each ranked university in granting and hiring of Ph.D. graduates. Harvard has granted degrees to 239 graduates who are presently employed by ranked universities. It itself employs 53 faculty members, all of whom are graduates of ranked universities, with 14 of them receiving their degree from Harvard. Self-hires account for 26.42 percent of Harvard's faculty. On average, 11.73 percent of employed economists are self-hires. Such high average self hiring percentage is largely due to universities outside North America, as shown below. Table 2.6 illustrates production capacity differences across universities alluded to above. Placement of graduates in ranked universities is difficult and is very unevenly distributed (and at least to some extent correlated with total capacity which is unobservable in this study since we do not observe graduates who find jobs outside the academia). Only Massachusetts

Table 2.6: Hiring of All Ranked Departments

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
1	Harvard	239	53	0	53	100.00	14	26.42
2	Chicago	218	26	0	26	100.00	4	15.38
3	Penn	114	32	0	32	100.00	2	6.25
4	Stanford	183	35	0	35	100.00	5	14.29
5	MIT	255	38	0	38	100.00	11	28.95
6	UC Berkeley	200	56	1	55	98.21	4	7.27
7	Northwestern	141	40	0	40	100.00	3	7.50
8	Yale	169	45	0	45	100.00	7	15.56
9	Michigan	78	57	0	57	100.00	1	1.75
10	Columbia	70	48	0	48	100.00	1	2.08
11	Princeton	173	52	1	51	98.08	9	17.65
12	UCLA	62	44	0	44	100.00	1	2.27
13	NYU	43	45	0	45	100.00	1	2.22
14	Cornell	70	34	3	31	91.18	1	3.23
15	LSE	115	50	1	49	98.00	2	4.08
16	Wisc Madison	125	31	0	31	100.00	1	3.23
17	Duke	46	32	0	32	100.00	2	6.25
18	Ohio State	23	33	0	33	100.00	0	0.00
19	Maryland	26	37	1	36	97.30	0	0.00
20	Rochester	86	20	0	20	100.00	1	5.00
21	UT Austin	16	35	0	35	100.00	1	2.86
22	Minnesota	135	26	1	25	96.15	2	8.00
23	UIUC	47	36	0	36	100.00	0	0.00
24	UC Davis	17	29	2	27	93.10	0	0.00

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
25	Toronto	38	61	4	57	93.44	3	5.26
26	Oxford	106	43	0	43	100.00	18	41.86
27	UBC	54	33	0	33	100.00	4	12.12
28	UCSD	58	37	1	36	97.30	0	0.00
29	USC	14	23	1	22	95.65	0	0.00
30	BU	24	36	2	34	94.44	0	0.00
31	Penn State	22	25	0	25	100.00	1	4.00
32	CMU	34	37	0	37	100.00	5	13.51
33	Cambridge	73	21	3	18	85.71	9	50.00
34	Florida	7	18	0	18	100.00	0	0.00
35	Mich State	29	41	1	40	97.56	1	2.50
36	Rutgers	10	33	0	33	100.00	1	3.03
37	U Washington	35	25	0	25	100.00	0	0.00
38	UNC	23	30	1	29	96.67	1	3.45
39	TAMU	15	29	0	29	100.00	1	3.45
40	Indiana	24	22	0	22	100.00	0	0.00
41	Iowa	22	21	1	20	95.24	1	5.00
42	Tel Aviv	5	20	0	20	100.00	2	10.00
43	UVA	33	27	0	27	100.00	0	0.00
44	UCL	24	37	4	33	89.19	4	12.12
45	Hebrew	18	23	0	23	100.00	5	21.74
46	Brown	56	27	0	27	100.00	0	0.00
47	Tilburg	26	22	3	19	86.36	6	31.58
48	Pitt	25	25	0	25	100.00	1	4.00

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
49	Warwick	26	28	1	27	96.43	5	18.52
50	Arizona	6	23	0	23	100.00	0	0.00
51	West Ontario	42	28	0	28	100.00	1	3.57
52	JHU	43	14	0	14	100.00	0	0.00
53	ANU	24	21	3	18	85.71	4	22.22
54	Vanderbilt	8	29	0	29	100.00	0	0.00
55	Queen's	43	32	0	32	100.00	4	12.50
56	WUSTL	24	21	1	20	95.24	0	0.00
57	Montreal	14	28	1	27	96.43	0	0.00
58	Georgetown	4	28	0	28	100.00	0	0.00
59	CO Boulder	12	29	1	28	96.55	0	0.00
60	UGA	1	16	0	16	100.00	0	0.00
61	VA Tech	16	13	1	12	92.31	0	0.00
62	Purdue	46	21	3	18	85.71	0	0.00
63	UC Irvine	4	25	0	25	100.00	0	0.00
64	BC	15	28	0	28	100.00	0	0.00
65	Iowa State	8	29	0	29	100.00	1	3.45
66	Amsterdam	22	35	1	34	97.14	14	41.18
67	NC State	10	26	1	25	96.15	1	4.00
68	Erasmus	15	17	0	17	100.00	7	41.18
69	Dartmouth	0	22	1	21	95.45	0	0.00
70	Louvain	35	22	6	16	72.73	13	81.25
71	U York	24	46	8	38	82.61	8	21.05
72	ASU	3	29	1	28	96.55	0	0.00

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
73	Toulouse	32	38	6	32	84.21	10	31.25
74	Essex	18	31	3	28	90.32	2	7.14
75	Stockholm	14	20	0	20	100.00	5	25.00
76	UCSB	13	29	1	28	96.55	0	0.00
77	LBS	5	7	1	6	85.71	0	0.00
78	Florida State	5	29	3	26	89.66	0	0.00
79	UNSW	12	35	2	33	94.29	7	21.21
80	Alberta	0	25	0	25	100.00	0	0.00
81	McMaster	15	27	0	27	100.00	1	3.70
82	Houston	2	22	0	22	100.00	0	0.00
83	Syracuse	11	26	0	26	100.00	3	11.54
84	UAB	24	41	7	34	82.93	9	26.47
85	Nottingham	11	43	9	34	79.07	5	14.71
86	HKUST	0	19	0	19	100.00	0	0.00
87	Bonn	23	19	7	12	63.16	7	58.33
88	York U	2	31	1	30	96.77	0	0.00
89	Cal Tech	19	16	0	16	100.00	0	0.00
90	LSU	3	14	1	13	92.86	1	7.69
91	Southampton	12	14	0	14	100.00	2	14.29
92	UConn	1	27	1	26	96.30	0	0.00
93	Georgia State	3	31	0	31	100.00	3	9.68
94	UKY	5	20	0	20	100.00	1	5.00
95	GWU	2	32	0	32	100.00	1	3.13
96	INSEE	0	10	2	8	80.00	0	0.00

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
97	SMU	4	18	0	18	100.00	0	0.00
98	Notre Dame	3	26	0	26	100.00	2	7.69
99	SSE	14	15	6	9	60.00	2	22.22
100	SFU	7	34	0	34	100.00	0	0.00
101	Oregon	5	19	0	19	100.00	0	0.00
102	GMU	9	29	1	28	96.55	2	7.14
103	Birkbeck	8	16	0	16	100.00	1	6.25
104	VUA	10	10	0	10	100.00	4	40.00
105	UMass	9	24	2	22	91.67	1	4.55
106	S Carolina	1	16	0	16	100.00	1	6.25
107	Paris I	51	41	10	31	75.61	30	96.77
108	Bristol	5	5	0	5	100.00	1	20.00
109	Melbourne	5	43	6	37	86.05	2	5.41
110	UIC	1	21	0	21	100.00	1	4.76
111	Copenhagen	31	42	7	35	83.33	26	74.29
112	McGill	8	34	2	32	94.12	2	6.25
113	Groningen	12	10	1	9	90.00	6	66.67
114	Ch UHK	0	19	1	18	94.74	0	0.00
115	ULB	17	10	0	10	100.00	10	100.00
116	Newcastle uT	1	3	1	2	66.67	0	0.00
117	Tulane	4	12	1	11	91.67	0	0.00
118	American	5	22	1	21	95.45	2	9.52
119	Mannheim	14	26	7	19	73.08	9	47.37
120	Auburn	1	12	1	11	91.67	0	0.00

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
121	UPF	17	56	8	48	85.71	1	2.08
122	Buffalo	6	20	2	18	90.00	1	5.56
123	Manchester	20	34	3	31	91.18	7	22.58
124	UCSC	4	23	0	23	100.00	1	4.35
125	Monash	16	32	7	25	78.13	11	44.00
126	Rice	12	21	0	21	100.00	1	4.76
127	Tennessee	1	16	0	16	100.00	0	0.00
128	Emory	0	16	0	16	100.00	0	0.00
129	NU Singapore	0	42	5	37	88.10	0	0.00
130	Laval	5	27	0	27	100.00	2	7.41
131	C3MU	5	49	10	39	79.59	0	0.00
132	Waterloo	3	23	1	22	95.65	3	13.64
133	Wayne State	1	13	0	13	100.00	0	0.00
134	Wisc Mil	3	22	0	22	100.00	0	0.00
135	Missouri	3	15	0	15	100.00	0	0.00
136	UC Riverside	4	19	2	17	89.47	0	0.00
137	Alabama	1	13	0	13	100.00	0	0.00
138	Quebec	2	28	1	27	96.43	0	0.00
139	Albany	2	19	0	19	100.00	0	0.00
140	Oslo	11	14	0	14	100.00	11	78.57
141	Miami FL	0	16	0	16	100.00	0	0.00
142	Maastricht	11	20	3	17	85.00	6	35.29
143	Delaware	0	26	0	26	100.00	0	0.00
144	Sydney	6	35	3	32	91.43	3	9.38

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
145	EHESS	25	20	4	16	80.00	5	31.25
146	Vienna	14	20	6	14	70.00	8	57.14
147	Munich	14	34	9	25	73.53	9	36.00
148	East Anglia	3	6	1	5	83.33	0	0.00
149	Geneva	2	7	0	7	100.00	0	0.00
150	INSEAD	0	11	0	11	100.00	0	0.00
151	Clemson	0	25	1	24	96.00	0	0.00
152	Birmingham	3	21	2	19	90.48	1	5.26
153	Guelph	2	24	2	22	91.67	1	4.55
154	Hitots	9	22	2	20	90.91	9	45.00
155	Tufts	1	22	0	22	100.00	0	0.00
156	BYU	0	21	0	21	100.00	0	0.00
157	Tokyo	32	55	3	52	94.55	25	48.08
158	CU Lon	3	7	0	7	100.00	2	28.57
159	Zurich	14	26	7	19	73.08	12	63.16
160	Stony Brook	13	13	0	13	100.00	0	0.00
161	Carleton	2	26	0	26	100.00	0	0.00
162	Reading	8	25	4	21	84.00	7	33.33
163	Academia S	0	35	7	28	80.00	0	0.00
164	KUL	20	20	1	19	95.00	10	52.63
165	Bar-Ilan	3	13	0	13	100.00	3	23.08
166	EUI	27	12	1	11	91.67	0	0.00
167	Bocconi	11	44	3	41	93.18	9	21.95
168	Utah	6	21	2	19	90.48	4	21.05

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
169	Brandeis	1	21	0	21	100.00	0	0.00
170	IUPUI	0	16	0	16	100.00	0	0.00
171	Exeter	4	15	1	14	93.33	1	7.14
172	Bologna	12	64	11	53	82.81	12	22.64
173	Wyoming	3	10	0	10	100.00	2	20.00
174	Nebraska	0	16	2	14	87.50	0	0.00
175	WVA	3	19	0	19	100.00	0	0.00
176	Kansas	3	21	0	21	100.00	0	0.00
177	NHH	14	17	1	16	94.12	14	87.50
178	Temple	1	24	3	21	87.50	0	0.00
179	Glasgow	5	16	2	14	87.50	1	7.14
180	SIUC	5	11	2	9	81.82	0	0.00
181	Kansas State	0	16	1	15	93.75	0	0.00
182	CUNY	9	58	1	57	98.28	4	7.02
183	Oklahoma	1	14	2	12	85.71	0	0.00
184	CWM	0	20	1	19	95.00	0	0.00
185	Strathclyde	5	15	1	14	93.33	2	14.29
186	Edinburgh	4	14	1	13	92.86	0	0.00
187	UHK	1	21	0	21	100.00	1	4.76
188	Wash State	4	12	0	12	100.00	0	0.00
189	Uppsala	12	14	5	9	64.29	7	77.78

Table 2.6 Continued

Coupe Rank	Employer	Number of Placed Graduates	Number of Employed Faculty	Number Hired from Non-Ranked Universities	Number Hired from Ranked Universities	Percentage Hired from Ranked Universities	Number of Self-Hires	Percentage of Self-Hires
190	Osaka	12	25	5	20	80.00	10	50.00
191	Tsukuba	1	7	0	7	100.00	0	0.00
192	UNM	1	11	0	11	100.00	1	9.09
193	UC Dublin	1	3	0	3	100.00	0	0.00
194	CO Denver	0	10	0	10	100.00	0	0.00
195	Rome LS	0	5	2	3	60.00	0	0.00
196	Concordia	2	24	3	21	87.50	0	0.00
197	SCU	0	13	0	13	100.00	0	0.00
198	QMUL	5	24	3	21	87.50	0	0.00
199	Montana State	1	13	0	13	100.00	1	7.69
200	URI	0	10	0	10	100.00	0	0.00
	Total	4,783	5,081	298	4,783	94.14	561	11.73

Institute of Technology, Harvard University, University of Chicago, and University of California at Berkeley have successfully placed more than 200 graduates each in other ranked university. Only an additional nine universities have placed more than 100 of their graduates in ranked universities. These top thirteen universities in terms of placement are responsible for training a full 45% of all economists hired by ranked universities in the square network (or 42% in the full network).

The third and fourth columns of Table 2.5 present the statistics of the North American sample. The number of all grantors whose graduates have been hired at ranked North American universities is 197, but only 108 of them are located in North America. The number of faculty members in ranked North American universities holding degrees from ranked North American universities, referred to as the *North American square ranked network*, is 3,026, while the total number of faculty members in North America is 3,279. While the average size of a ranked department in North America is 26.02, an average of 24.02 faculty were trained by ranked North American universities. Within North American universities, the employment capacity varies less than in the full network, leading to smaller standard deviations in hiring (10.63 and 9.81 for the square network). As with all ranked universities, average and median number of faculty in North America are not very different. The median faculty size is 25, while the median number of faculty hired from ranked North American universities is 22.

Since most faculty in North American universities received their degrees from other ranked North American universities, the average number of placed graduates per grantor is 16.99 for all grantors and 28.02 for North American ranked grantors. The standard deviation of placement by North American universities is as high as among all ranked universities because North American universities are the majority of all granting universities and their placement and production capacity varies greatly. Unsurprisingly, the large department can produce a large share of Ph.D. graduates. Harvard, for instance, has granted degrees to 196 Ph.D. graduates hired by top-ranking North American universities. The same number for Louisiana State University is only 3 (Table 2.7).

North American universities employ few Ph.D. graduates from non-ranked Ph.D. grantors (71 faculty members). Only 182 of their faculty received their degree from a ranked department outside of North America. Faculty hiring from North American grantors amounts to 92.28 percent of total faculty hiring (Table 2.5). Moreover, compared to all ranked universities, institutional self-hiring in North American ranked universities is much lower. Only 4.59 percent of faculty members in a North American university received their degree from their current employer as compared to 11.73 for all ranked universities.

Figures 2.3 and 2.4 illustrate the histogram of the years when degrees were granted for those economists with available data. These figures are presented for illustrative purposes only, showing the range of years over which current academic economists obtained their degrees. There are relatively few economists still active who obtained their degree before 1960. There is a rapid growth in the number of degrees granted in the 1960s, followed by a much slower growth rate of degrees granted from 1970s through mid-1990s. From 1997 on there is a much more rapid growth in the number of degrees granted. These trends can be explained by retirement patterns and replacement needs, growth of universities due to the larger number of students attending, and various other factors which are outside the scope of this analysis.

2.3.1 Tier Analysis

In order to capture the interactions between universities in different layers of the top 200, they have been divided into 8 tiers, each with 25 members as shown in Table 2.8. The 25 top-ranked

Table 2.7: Hiring of North American Ranked Departments

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
1	Harvard	239	196	53	51	2	0	96.23	14	27.45
2	Chicago	218	193	26	26	0	0	100.00	4	15.38
3	Penn	114	94	32	30	2	0	93.75	2	6.67
4	Stanford	183	157	35	32	3	0	91.43	5	15.63
5	MIT	255	214	38	35	3	0	92.11	11	31.43
6	UC Berkeley	200	172	56	51	4	1	91.07	4	7.84
7	Northwestern	141	117	40	38	2	0	95.00	3	7.89
8	Yale	169	141	45	43	2	0	95.56	7	16.28
9	Michigan	78	68	57	54	3	0	94.74	1	1.85
10	Columbia	70	57	48	42	6	0	87.50	1	2.38
11	Princeton	173	138	52	46	5	1	88.46	9	19.57
12	UCLA	62	51	44	42	2	0	95.45	1	2.38
13	NYU	43	25	45	41	4	0	91.11	1	2.44
14	Cornell	70	51	34	29	2	3	85.29	1	3.45
16	Wisc Madison	125	115	31	29	2	0	93.55	1	3.45
17	Duke	46	42	32	30	2	0	93.75	2	6.67
18	Ohio State	23	22	33	31	2	0	93.94	0	0.00
19	Maryland	26	22	37	36	0	1	97.30	0	0.00
20	Rochester	86	72	20	18	2	0	90.00	1	5.56
21	UT Austin	16	15	35	35	0	0	100.00	1	2.86
22	Minnesota	135	110	26	24	1	1	92.31	2	8.33
23	UIUC	47	41	36	33	3	0	91.67	0	0.00

Table 2.7 Continued

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
24	UC Davis	17	11	29	24	3	2	82.76	0	0.00
25	Toronto	38	33	61	50	7	4	81.97	3	6.00
27	UBC	54	41	33	32	1	0	96.97	4	12.50
28	UCSD	58	36	37	34	2	1	91.89	0	0.00
29	USC	14	7	23	17	5	1	73.91	0	0.00
30	BU	24	17	36	32	2	2	88.89	0	0.00
31	Penn State	22	14	25	21	4	0	84.00	1	4.76
32	CMU	34	30	37	37	0	0	100.00	5	13.51
34	Florida	7	7	18	17	1	0	94.44	0	0.00
35	Mich State	29	27	41	38	2	1	92.68	1	2.63
36	Rutgers	10	9	33	32	1	0	96.97	1	3.13
37	U Washington	35	32	25	24	1	0	96.00	0	0.00
38	UNC	23	23	30	28	1	1	93.33	1	3.57
39	TAMU	15	14	29	28	1	0	96.55	1	3.57
40	Indiana	24	23	22	21	1	0	95.45	0	0.00
41	Iowa	22	19	21	20	0	1	95.24	1	5.00
43	UVA	33	32	27	27	0	0	100.00	0	0.00
46	Brown	56	53	27	25	2	0	92.59	0	0.00
48	Pitt	25	23	25	22	3	0	88.00	1	4.55
50	Arizona	6	5	23	21	2	0	91.30	0	0.00
51	West Ontario	42	34	28	25	3	0	89.29	1	4.00

Table 2.7 Continued

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
52	JHU	43	36	14	14	0	0	100.00	0	0.00
54	Vanderbilt	8	6	29	27	2	0	93.10	0	0.00
55	Queen's	43	38	32	30	2	0	93.75	4	13.33
56	WUSTL	24	21	21	18	2	1	85.71	0	0.00
57	Montreal	14	14	28	22	5	1	78.57	0	0.00
58	Georgetown	4	1	28	25	3	0	89.29	0	0.00
59	CO Boulder	12	12	29	28	0	1	96.55	0	0.00
60	UGA	1	1	16	16	0	0	100.00	0	0.00
61	VA Tech	16	12	13	10	2	1	76.92	0	0.00
62	Purdue	46	41	21	18	0	3	85.71	0	0.00
63	UC Irvine	4	2	25	24	1	0	96.00	0	0.00
64	BC	15	10	28	24	4	0	85.71	0	0.00
65	Iowa State	8	8	29	27	2	0	93.10	1	3.70
67	NC State	10	10	26	24	1	1	92.31	1	4.17
69	Dartmouth	0	0	22	21	0	1	95.45	0	0.00
72	ASU	3	3	29	27	1	1	93.10	0	0.00
76	UCSB	13	12	29	27	1	1	93.10	0	0.00
78	Florida State	5	5	29	26	0	3	89.66	0	0.00
80	Alberta	0	0	25	22	3	0	88.00	0	0.00
81	McMaster	15	13	27	24	3	0	88.89	1	4.17
82	Houston	2	2	22	20	2	0	90.91	0	0.00

Table 2.7 Continued

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
83	Syracuse	11	10	26	26	0	0	100.00	3	11.54
88	York U	2	2	31	24	6	1	77.42	0	0.00
89	Cal Tech	19	18	16	15	1	0	93.75	0	0.00
90	LSU	3	3	14	13	0	1	92.86	1	7.69
92	UConn	1	1	27	26	0	1	96.30	0	0.00
93	Georgia State	3	3	31	31	0	0	100.00	3	9.68
94	UKY	5	4	20	20	0	0	100.00	1	5.00
95	GWU	2	1	32	32	0	0	100.00	1	3.13
97	SMU	4	4	18	15	3	0	83.33	0	0.00
98	Notre Dame	3	3	26	25	1	0	96.15	2	8.00
100	SFU	7	2	34	31	3	0	91.18	0	0.00
101	Oregon	5	5	19	16	3	0	84.21	0	0.00
102	GMU	9	7	29	27	1	1	93.10	2	7.41
105	UMass	9	9	24	21	1	2	87.50	1	4.76
106	S Carolina	1	1	16	15	1	0	93.75	1	6.67
110	UIC	1	1	21	21	0	0	100.00	1	4.76
112	McGill	8	7	34	26	6	2	76.47	2	7.69
117	Tulane	4	4	12	11	0	1	91.67	0	0.00
118	American	5	5	22	20	1	1	90.91	2	10.00
120	Auburn	1	1	12	11	0	1	91.67	0	0.00
122	Buffalo	6	6	20	18	0	2	90.00	1	5.56

Table 2.7 Continued

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
124	UCSC	4	4	23	23	0	0	100.00	1	4.35
126	Rice	12	9	21	19	2	0	90.48	1	5.26
127	Tennessee	1	1	16	16	0	0	100.00	0	0.00
128	Emory	0	0	16	16	0	0	100.00	0	0.00
130	Laval	5	5	27	21	6	0	77.78	2	9.52
132	Waterloo	3	3	23	22	0	1	95.65	3	13.64
133	Wayne State	1	1	13	13	0	0	100.00	0	0.00
134	Wisc Mil	3	2	22	22	0	0	100.00	0	0.00
135	Missouri	3	3	15	15	0	0	100.00	0	0.00
136	UC Riverside	4	0	19	17	0	2	89.47	0	0.00
137	Alabama	1	1	13	13	0	0	100.00	0	0.00
138	Quebec	2	2	28	22	5	1	78.57	0	0.00
139	Albany	2	1	19	19	0	0	100.00	0	0.00
141	Miami FL	0	0	16	16	0	0	100.00	0	0.00
143	Delaware	0	0	26	26	0	0	100.00	0	0.00
151	Clemson	0	0	25	24	0	1	96.00	0	0.00
153	Guelph	2	2	24	19	3	2	79.17	1	5.26
155	Tufts	1	0	22	20	2	0	90.91	0	0.00
156	BYU	0	0	21	21	0	0	100.00	0	0.00
160	Stony Brook	13	9	13	11	2	0	84.62	0	0.00
161	Carleton	2	2	26	24	2	0	92.31	0	0.00

Table 2.7 Continued

Coupe Rank	Employer	Number of Placed Graduates in all Ranked Universities	Number of Placed Graduates in North American Ranked Universities	Number of Faculty	Number of Hires from North America	Number of Hires from Ranked Non-North America	Number of Hires from Non-Ranked Universities	Percentage of Hires from North America	Self-Hires	Percentage of Self-Hires
168	Utah	6	5	21	19	0	2	90.48	4	21.05
169	Brandeis	1	0	21	21	0	0	100.00	0	0.00
170	IUPUI	0	0	16	15	1	0	93.75	0	0.00
173	Wyoming	3	3	10	9	1	0	90.00	2	22.22
174	Nebraska	0	0	16	14	0	2	87.50	0	0.00
175	WVA	3	3	19	19	0	0	100.00	0	0.00
176	Kansas	3	3	21	21	0	0	100.00	0	0.00
178	Temple	1	1	24	21	0	3	87.50	0	0.00
180	SIUC	5	4	11	9	0	2	81.82	0	0.00
181	Kan State	0	0	16	15	0	1	93.75	0	0.00
182	CUNY	9	8	58	56	1	1	96.55	4	7.14
183	Oklahoma	1	1	14	12	0	2	85.71	0	0.00
184	CWM	0	0	20	19	0	1	95.00	0	0.00
188	Wash State	4	4	12	12	0	0	100.00	0	0.00
192	UNM	1	1	11	11	0	0	100.00	1	9.09
194	CO Denver	0	0	10	10	0	0	100.00	0	0.00
196	Concordia	2	0	24	20	1	3	83.33	0	0.00
197	SCU	0	0	13	12	1	0	92.31	0	0.00
199	Montana State	1	1	13	13	0	0	100.00	1	7.69
200	URI	0	0	10	10	0	0	100.00	0	0.00
	Total	3,601	3,026	3,279	3,026	182	71	92.28	139	4.59

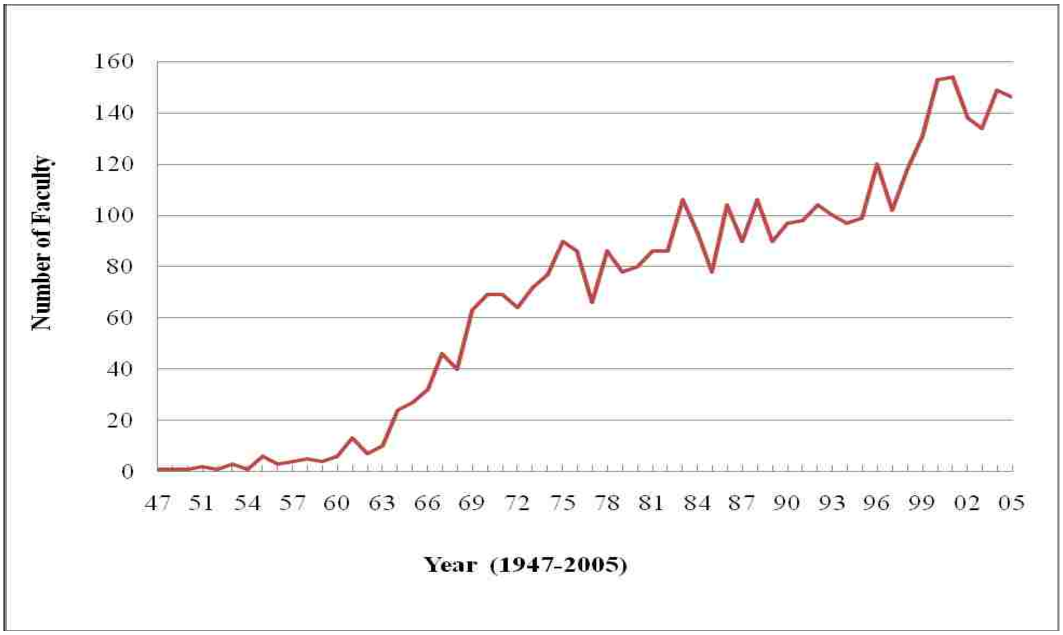


Figure 2.3: Histogram of Year of Graduation in All Ranked Universities

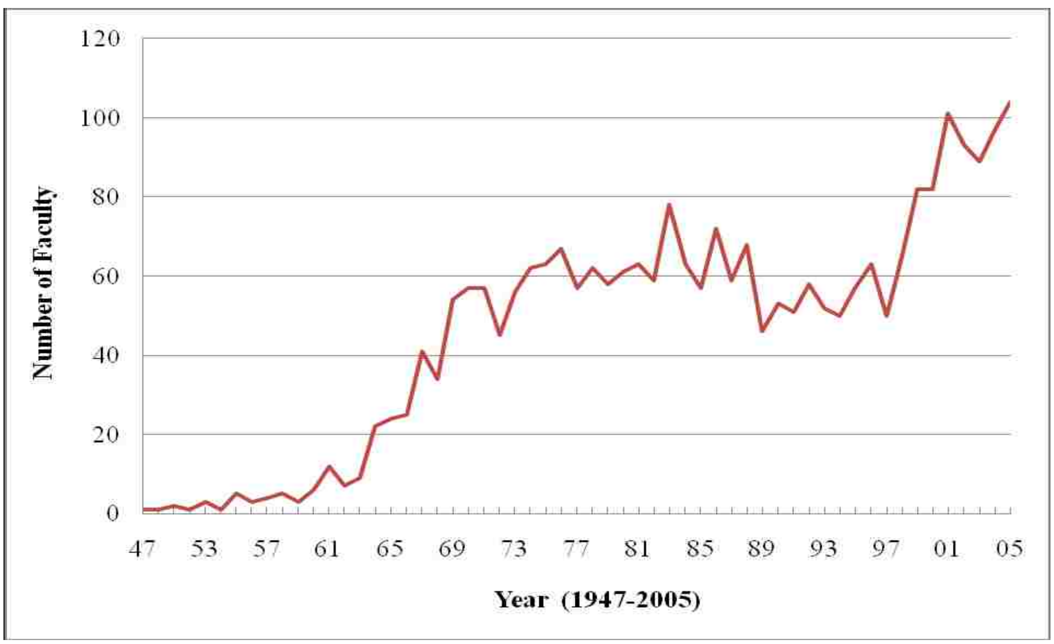


Figure 2.4: Histogram of Year of Graduation in North American Ranked Universities

universities form Group 1, those ranked 26 to 50 form Group 2, 51 to 75 form Group 3, etc. It may be said that moving from top to bottom tier reflects a decline in quality, since the ranking order is based on faculty publications. Ranked North American universities are divided into 5

Table 2.8: Eight Groups in Square Ranked Network

Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**	**Group 7**	**Group 8**
Harvard	Oxford	WOntario	UCSB	Oregon	Rice	Clemson	Kansas
Chicago	UBC	JHU	LBS	GMU	Tennessee	Birmingham	NHH
Penn	UCSD	ANU	FLSt	Birkbeck	Emory	Guelph	Temple
Stanford	USC	Vanderbilt	UNSW	VUAmsterdam	NUSingapore	Hitots	Glasgow
MIT	BU	Queen's	Alberta	UMass	Laval	Tufts	SIUC
UCBerkeley	PennSt	WUSTL	McMaster	SCarolina	C3MU	BYU	KanSt
Northwestern	CMU	Montreal	Houston	ParisI	Waterloo	Tokyo	CUNY
Yale	Cambridge	GTown	Syracuse	Bristol	WayneSt	CULon	Oklahoma
Michigan	Florida	COBoulder	UAB	Melbourne	WiscMil	Zurich	CWM
Columbia	MichSt	UGA	Nottingham	UIC	Missouri	StonyBrook	Strathclyde
Princeton	Rutgers	VATech	HKUST	Copenhagen	UCRiverside	Carleton	Edinburgh
UCLA	UWash	Purdue	Bonn	McGill	Alabama	Reading	UHK
NYU	UNC	UCIrvine	YorkU	Groningen	Quebec	AcademiaS	WashSt
Cornell	TAMU	BC	CalTech	ChUHK	Albany	KUL	Uppsala
LSE	Indiana	IowaSt	LSU	ULB	Oslo	Bar-Ilan	Osaka
WiscMad	Iowa	Amsterdam	Southampton	NewcastleuT	MiamiFL	EUI	Tsukuba
Duke	TelAviv	NCSt	UConn	Tulane	Maastricht	Bocconi	UNM
OhioSt	UVA	Erasmus	GASt	American	Delaware	Utah	UCDublin
Maryland	UCL	Dartmouth	UKY	Mannheim	Sydney	Brandeis	CODenver
Rochester	Hebrew	Louvain	GWU	Auburn	EHESS	IUPUI	RomeLS
UTAustin	Brown	UYork	INSEE	UPF	Vienna	Exeter	Concordia
Minnesota	Tilburg	ASU	SMU	Buffalo	Munich	Bologna	SCU
UIUC	Pitt	Toulouse	NotreDame	Manchester	EAnglia	Wyoming	QMUL
UCDavis	Warwick	Essex	SSE	UCSC	Geneva	Nebraska	MontSt
Toronto	Arizona	Stockholm	SFU	Monash	INSEAD	WVA	URI

groups as shown in Table 2.9. Each of the first four groups contains 25 universities, while the fifth has 26 members. Interactions and links between different tiers in both samples are presented in Figures 2.5 and 2.6. Both schemes are derived from valued-edge graphs. The edges are valued

Table 2.9: Five Groups in North American Square Ranked Network

Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**
Harvard	UCSD	UGA	Oregon	Clemson
Chicago	USC	VATech	GMU	Guelph
Penn	BU	Purdue	UMass	Tufts
Stanford	PennSt	UCIrvine	SCarolina	BYU
MIT	CMU	BC	UIC	StonyBrook
UCBerkeley	Florida	IowaSt	McGill	Carleton
Northwestern	MichSt	NCSt	Tulane	Utah
Yale	Rutgers	Dartmouth	American	Brandeis
Michigan	UWash	ASU	Auburn	IUPUI
Columbia	UNC	UCSB	Buffalo	Wyoming
Princeton	TAMU	FLSt	UCSC	Nebraska
UCLA	Indiana	Alberta	Rice	WVA
NYU	Iowa	McMaster	Tennessee	Kansas
Cornell	UVA	Houston	Emory	Temple
WiscMad	Brown	Syracuse	Laval	SIUC
Duke	Pitt	YorkU	Waterloo	KanSt
OhioSt	Arizona	CalTech	WayneSt	CUNY
Maryland	WOntario	LSU	WiscMil	Oklahoma
Rochester	JHU	UConn	Missouri	CWM
UTAustin	Vanderbilt	GASt	UCRiverside	WashSt
Minnesota	Queen's	UKY	Alabama	UNM
UIUC	WUSTL	GWU	Quebec	CODenver
UCDavis	Montreal	SMU	Albany	Concordia
Toronto	GTown	NotreDame	MiamiFL	SCU
UBC	COBoulder	SFU	Delaware	MontSt
				URI

in terms of the number of faculty hiring or granting between the vertices or groups. Thus, the thickness of lines, or edges, between groups reflects tie strength. The patterns of interactions between groups in both networks are quite similar. Although each group connects directly to

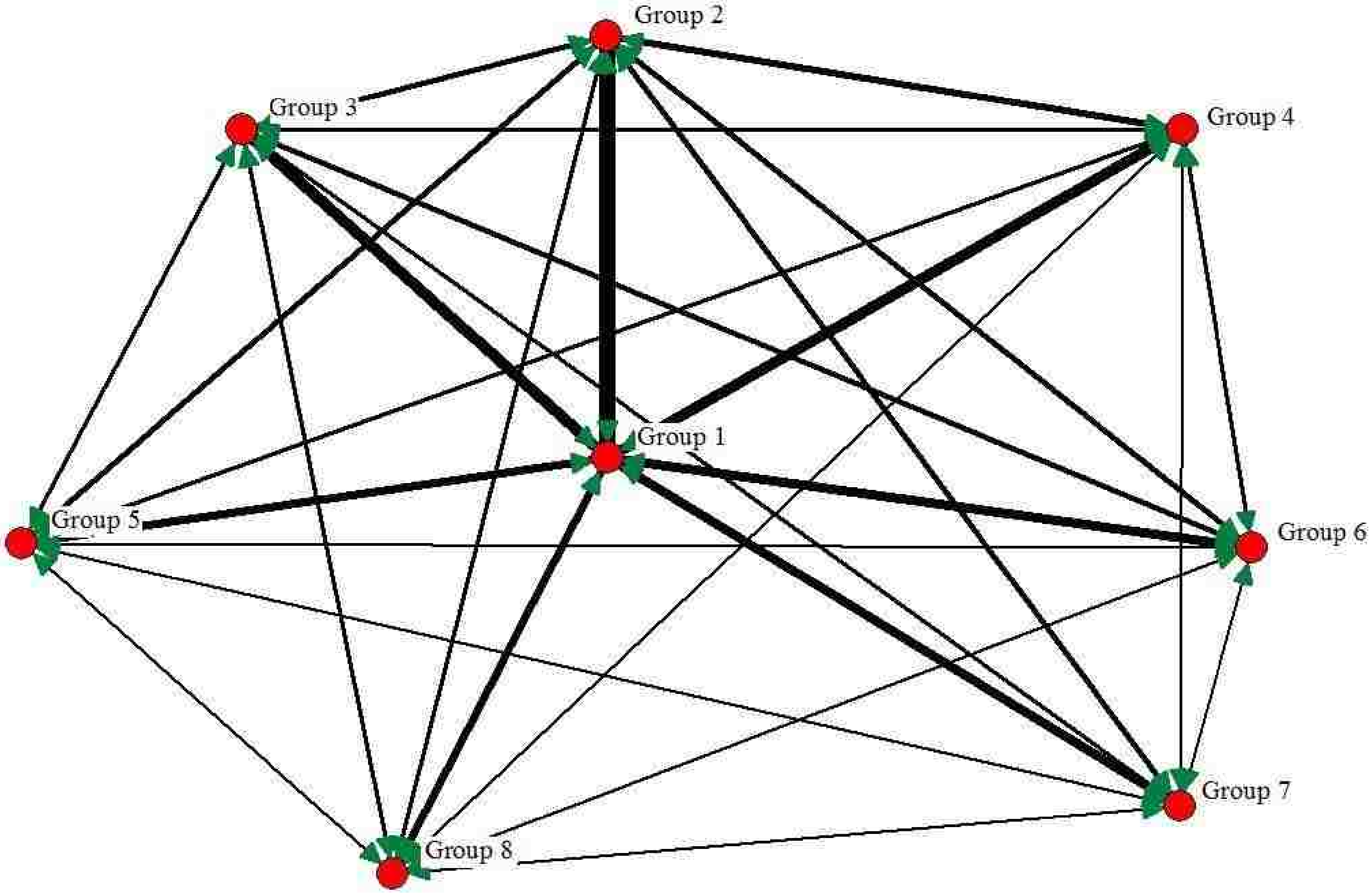


Figure 2.5: Group Interactions in the Square Ranked Network

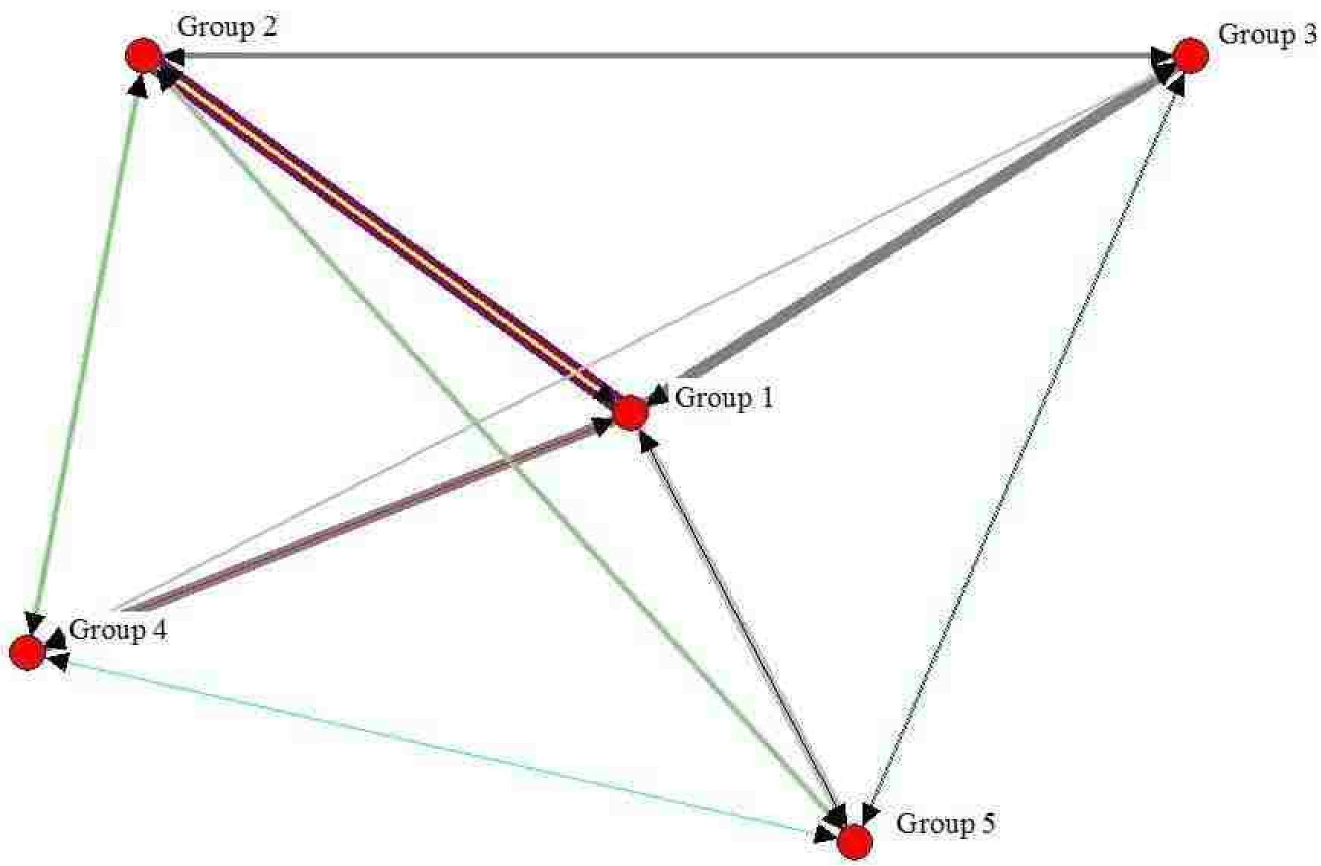


Figure 2.6: Group Interactions in the North American Square Ranked Network

each other, the strength of ties varies. Every group connects to Group 1, the highest-rated group, with strong ties. The strongest tie is between Group 1 and Group 2. The ties between groups, except with Group 1, are moderate to weak. The weakest ties are connections between low-level groups. For example, the tie between Group 5 and Group 8 in theSquare Ranked Network (Figure 2.5), and the tie between Group 4 and Group 5 in the North American ranked network (Figure 2.6), are both weak.

Tables 2.10 and 2.11 present the summary information for each group in both networks. These tables provide information on total and average faculty size as well as the distribution of hiring and placement of graduates relative to the rank of the group. 'Hiring (Placement) within' and refer to hiring (placement) within the same group, while 'Hiring (Placement) within Self' refers to hiring own graduates. Several patterns are clear. The rank of a group is inversely proportional to faculty size. The highest ranked group (group 1) has the largest number faculty.

Table 2.10: All Ranked Universities by Group

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Faculty size	995	706	648	610	579	532	593	418
Average size	39.8	28.24	25.92	24.4	23.16	21.28	23.72	16.72
% Hiring above		65.99	67.92	78.03	66.22	85.22	79.04	86.23
% Hiring within	86.34	22.97	19.80	11.60	27.03	11.29	20.04	13.77
% Self-hiring	7.75	9.45	11.46	7.73	22.97	9.86	18.01	10.65
% Hiring below	13.66	11.05	12.27	10.37	6.76	3.49	0.92	
% Placed above		10.14	14.88	21.94	23.85	44.53	31.64	39.08
% Placed within	31.50	20.55	25.37	33.67	53.85	42.97	61.58	60.92
% Placed within Self	2.83	8.45	14.68	22.45	45.77	37.50	55.37	47.13
% Granting below	68.50	69.31	59.75	44.39	22.31	12.50	6.78	

Table 2.11: North American Ranked Universities by Group

	Group 1	Group 2	Group 3	Group 4	Group 5
Faculty size	978	688	616	506	481
Average size	39.12	27.52	24.64	20.24	18.5
% Hiring above		77.25	85.04	90.77	95.40
% Hiring within	91.13	16.90	10.39	7.51	4.60
% Hiring within Self (Self-hiring)	8.65	2.69	2.46	3.65	2.91
% Hiring below	8.87	5.85	4.58	1.72	
% Placed above		12.38	19.41	33.33	55.32
% Placed within	37.40	20.08	34.71	44.87	44.68
% Placed within Self	3.55	3.19	8.24	21.79	27.66
% Placed below	62.60	67.54	45.88	21.79	

This tendency is clearest in the North American ranked university network. The high-level groups are much larger in terms of faculty than lower-level groups. The size ratio between the highest and lowest groups in both samples is 2-to-1. In the North American ranked university network, Group 5 has the smallest average faculty size.

Other patterns of hiring and granting are also clearer in the North American ranked university network (Table 2.11). Percentage of 'hiring above' in Table 2.11 indicates that every group hires intensively from groups ranked above it. As should be expected, the lower the group the more it hires from groups above it. This is not surprising as departments try to hire the best faculty possible, which should directly translate in hiring graduates from higher ranked departments. The same reasoning applies to patterns in 'hiring within' and 'hiring below.' The lower the group, the lower the percentage of faculty hired within the group and below the group. That Group 1 has the highest percentage of hires within the group is not surprising since it has no universities above itself to hire from. This pattern is not as clear-cut across the eight groups of all ranked universities. This is mainly because of the higher tendency of universities outside of North America to hires their own graduates. Group 5 stands out the most, followed by Group 7 and to some extent Group 3. As is seen in Table 2.10, universities in Group 5 hire by far the highest percentage of their faculty from their own graduates. Almost 23% of their faculty are their own graduates. In Group 7, 18% of faculty are self-hires. While Group 4 in the North American sample also breaks the trend for self hires, it does so by a much smaller margin.

Certain patterns are present in the distribution of placement of graduates. As the rank of the group decreases, the percentage of graduates placed to higher groups increases, while that of graduates placed to lower decreases. This is to a large extent due to the nature of the sample used. As the rank of the group decreases, the number of lower ranked universities observed where graduates can be placed decreases, while the number of higher ranked universities where graduates can be placed increases. Given that all possible academic placements are not observed, the limits imposed by what is possibly observed in the sample results in the observed patterns.

Percentages of 'placement below' indicate that higher groups grant faculty to lower groups more than lower group grant faculty to higher ones. For example, Group 4 hires 90 percent of its faculty from higher-ranked universities, while it is able to place 22% of its graduates to universities in Group 5. In addition, comparing placements above and below, most universities place the majority of their graduates in lower ranked universities. This pattern is somewhat reduced in the lower half of ranked universities, but it is still the case that they place more graduates at their own level or below than in universities ranked above them. These findings are in line with Moore and Newman (1977)'s "downstream pattern," which means that most new Ph.D. graduates are likely to find a job at a lower-level university than their grantor.

Table 2.11 provides two more interesting facts. One is that lower groups place more graduates to other universities within the same group than higher groups do. For example, Group 5 in the North American ranked network places 45% of its graduates within its own group, while Groups 1 and 2 place only 37% and 20% within their own groups. What is more, universities in lower groups more frequently place their graduates to themselves than the universities in higher-level groups. Group 5 in Table 2.11 has the highest percentage of placement to self. The percentage of self-hiring to total Ph.D. granting of group 5 is 27.66 percent compared 3.55 in group 1. Both of these results are due to the low production capacity of the lowest ranked universities. Given their low rank and the few graduates they are able to place to ranked universities, they are much more likely to be able to place them within their own group, or to hire them themselves, than to place them in higher ranked departments.

2.3.2 Trivial Pursuits

A university level analysis of the data presented in Tables 2.6 and 2.7 would just involve a repetition of the numbers shown there. Hence we provide an idiosyncratic collection of facts culled from those tables.

In our sample MIT has placed the largest number of students in the ranked universities (255) followed by Harvard at 233 placements. The first non-North American institution is London School of Economics with 115 and placements. They are followed by Oxford with a 106 placements. In the top 10 programs Michigan (rank 9) and Columbia (rank 10) have each placed less than a 100 students while the other 8 have all placed over 100 students. Focusing on continental Europe we institutions with highest placements are Paris I (51), Louvain (35), Toulouse (32), Copenhagen (31), European University Institute (27), Tilburg (26) and Bonn (23). In Australia, Australian National University has the highest number of placements with 24. In Asia, Tokyo University with 32 placements is the placement leader. Typically the number of placements decreases with the rank, but occasionally there are some exceptions like Queen Mary in London which despite its rank of 198 has placed 5 students in top ranked programs. Similarly from Table 2.7 we see that Stony Brook and CUNY are exceptions with 13 and 9 placements respectively. Explanations for these differences can range from the size of the program to locations to reputation of the institutions themselves.

Our data on university faculty sizes is based on roster data available from the department's webpage. Note that this number can be misleading since in many universities economists may be spread over different departments. The single largest department in our sample is University of Bologna with faculty size of 64 followed closely by Toronto with 61 members. While the smallest departments in our sample are from Newcastle upon Tyne and

University College Dublin with three members each, this conclusion is somewhat misleading as this number reports the number of faculty for whom the minimum required information is available. Thus, Newcastle upon Tyne and University College Dublin have mote than three faculty members, but do not report all needed information on the rest of the faculty. This is not an issue with North American departments as all needed information is available. In the North American sample there is a tie for this spot with both University of Colorado at Denver (rank 194) and University of Rhode Island (rank 200) having 10 faculty members each.

When we examine which ranked universities hire from non-ranked ones we find the presence of a language effect, which is likely a consequence of the segmented labor market outside of North America. University of Bologna has 11 members from non-ranked universities followed closely by Universidad Carlos III de Madrid and Paris I with 10 members each. Among universities where English is the sole language of instruction, this distinction goes to University of Nottingham with 9 of its 43 members being from non-ranked universities. In the North American sample Toronto (rank 25) has 4 of its 61 members hired from universities that are not ranked. Cornell (rank 14), Purdue (rank 62), Florida State (rank 78), Temple (rank 178) and Concordia (rank 196) each have 3 faculty members who obtained their doctoral degree from a non-ranked university. Among the top 10 schools only UC Berkley has a hire from a non-ranked university - Université Libre de Bruxelles. Princeton's one non-ranked hire is from the University of St. Gallen, while Cornell's non-ranked hires are from Heidelberg (2) and University of Aarhus (1).

In the North American sample we also examine how often universities hire from ranked schools outside of North America. In terms of having an "international flavor" Canadian universities seem to be doing better than their US counterparts. Of the US Universities, USC has

26% and Virginia Tech 23% of their hires from ranked universities outside North America. Six Canadian schools have more than twenty percent of their faculty members from ranked universities outside North America – McGill (24%), York (77%), Laval (78%), Montreal (79%), Quebec (79%) and Guelph (79%).

Self-hiring is quite predominant in the top 10 universities with MIT having 29% of its own doctoral students on its faculty. Harvard comes next with 26% while Michigan (2%) and Columbia (2%) have the lowest percentages. Among the other American schools Carnegie-Mellon has the highest number of self-hires at 14%, closely followed by Syracuse with 11%. Other universities with high self-hires are American University and Georgia State at 10% each. Among Canadian universities Waterloo has the largest number of self-hires (14%) followed by Queen's at 13% and UBC at 12%. In universities outside the US self-hires can be very high ranging from 42% at Oxford and 50% at Cambridge to over 90 percent in a few institutions. Paris I, which is the leading continental Europe university in placement, placed 30 of its 51 graduates with itself, resulting in 75% of its faculty being self-hires. These differences with North America can be attributed to language barriers, reputation of institutions as well as to institutional factors such as segmented nature of academic labor markets outside North America.

2.4. Summary

The goal of this study is to examine the labor market for academic economists from a social network perspective. Studying social networks in labor markets present empirical challenges due to their stringent data requirements. This study solves this problem by studying the labor market for academic economists, examining the links created between universities when they hire faculty to staff their departments. Since universities readily publish information about their faculty and faculty do so themselves, it is easy to observe where each economist obtained his or her terminal degree and which university employs them. If MIT employs a graduate of Princeton University, it creates a connection between the two universities which can facilitate the flow of information used for future hiring decisions of both universities. Hence, unlike other labor markets, the academic labor market is an example of a market where many interactions or links are observed. By studying the faculty of MIT and Princeton and where each was trained, one can start compiling information on the whole network. The more departments one examines, the more complete the picture of the network.

This study examines the network created by hiring decisions of the top two hundred economics departments in the world as ranked by Coupe (2003), and focuses more closely on the subset of these located in North America (126 universities). Information on economists employed by these universities was collected from university websites. This chapter described the data end examined the basic patterns of hiring and placements between universities as well as continents where universities are located. North American universities, particularly the US, dominated the network. They both hire the most and produce the most academic economists. While departments are somewhat uniform in terms of the number of faculty they higher, they vary much more in their ability to produce and place their graduates in other ranked departments. Only thirteen universities account for over 40% of all academic economists in the network of two hundred universities.

Universities in Australia, Europe, and Asia are much more likely to employ their own graduates than are North American ones. This is likely due to higher fragmentation of academic labor markets outside of North America which is dictated by smaller country size as well as geographic distances between universities which train academic economists. In all ranked universities 11.73 percent of faculty received their degrees from the same university which employs them, while in North American universities only the corresponding figure is 4.59 percent. In the US, employment of own graduates is most prominent in the highest ranked universities as well as the lowest ones. Highest ranked universities produce the best economists, so it is not surprising they tend to hire more from themselves as there is not better product available. Lowest ranked universities, on the other hand, might use the same argument, but in their case it is likely they are unable to attract as many candidates from better departments and are forced to employ their own graduates.

Not surprisingly the lower ranked universities tend to hire more faculty from universities which are ranked above them. The lower the rank of a university, the smaller is the size of its faculty, the more they hire from hire ranked universities, and the less they hire from universities at their own level or below.

Link between higher ranked departments are stronger than those between higher and lower ranked departments. This is because higher ranked departments tend to hire from themselves and not from lower ranked departments. In a sense, links between two high ranked departments are bidirectional or reciprocal, while those between a high and low ranked department are one directional. The next chapter will provide a more thorough description of the network created by university hiring decisions.

There are three main conclusions of this chapter. When universities are divided in groups of 25 based on their ranking, every group tends to hire graduates from groups ranked above it. On the other hand, the top-ranked grantors place their graduates mostly in groups ranked below them. This corroborates the "downstream pattern" found by Moore and Newman (1977). Second, the low-level groups hire fewer graduates from their own group but have high potential to place their own graduates to other members of the same group. This is due to the few graduates they are able to place in the observed departments. While they hire many faculties from higher ranked departments, the few they are able to place are hired by universities in their own group. Third, the lower the level of a group, the more likely it is to hire its own graduates.

CHAPTER 3: SMALL WORLD IN THE LABOR NETWORK OF ACADEMIC ECONOMISTS

3.1 Introduction

An individual's connectedness within the social network can influence their job search and improve their job mobility. Interactions between agents in the social network follow relatively stable patterns. Knowledge about these interaction patterns can allow an individual to manipulate the network to his or her advantage. This individual can then choose to build connections for efficient input of information from every other member in the network. Social relationships between universities denoted by nodes, represented by links in the network, have been formed through sustained patterns of production and placement of Ph.D. graduates in the economics academic market. Every group of universities representing a similar quality level in the economics academic market can connect to every other and tends to place its graduates in a "downstream pattern" identified by Moore and Newman (1977). Most of the higher-ranking Ph.D. grantors place their products in lower-ranking groups. On the other hand, every group is eager to hire Ph.D. faculty from a higher quality group.

One surprising characteristic of the academic market for economists is that any ranked university can reach any other university through a maximum of seven links. This means that even when two different universities cannot directly connect to each other, they are indirectly connected through only a few links. Interactions between persons or institutions can spread information far and fast over a network with high clustering and short paths. Within that kind of network, everyone can connect to everyone through very few intermediaries. The economics academic market seems to be a good candidate for the so-called small-world network. This chapter aims to investigate whether or not the structure and properties of the economics academic network can be formalized into the claim that it is indeed a small-world network.

Since the structure of a social network determines, to a degree, economic interactions in the labor market, in addition to knowing the role of each agent and the number of participating agents, and the total network diameter, other structural characteristics of the network should be investigated. These include the density of relationships between members in the network, the inequality or reciprocity of connections, the central or most influential vertices, etc. Knowledge of these characteristics can be useful to agents who desire to maximize their influence.

Before we proceed, it would help to provide an overview of the existing literature that has applied small-world network theory to the world of economics. After that, we will introduce the model and tools to test for small-world phenomenon in a social network. Finally, we will present the empirical results of applying this model to the economics academic network.

3.2 Application in Economics

In recent years there has been an explosion in research on the small-world phenomenon. However, the literature in economics on this topic is relatively small. Goyal et al. (2004) investigated what they perceived to be an emerging small-world network of increasing collaboration and distant co-authorships among economists. Goyal et al. (2005) tested the hypothesis of the strength of weak-ties in a small-world network of economics scholars. Even though they found evidence for the small world phenomenon, they refuted the hypothesis of the strength of weak-ties in the economics networks. Instead, they contended that the collaboration network of economists exhibits different properties and structure when compared to physics or medical science networks. First, the largest component in both the physics and medical science networks cover almost the whole population, while the largest component among economists covers around a half of the population. Second, the characteristic distance in both physics and medical networks are very small compared to the characteristic path among economists. Trevio (2006) ranked economics department using data from the Google engine. Using a Multidimension Scaling (MDS), he also divided departments into clusters. Next we briefly summarize the main finding of each of these papers related to small-world network.

Goyal et al. (2004) constructed distinct networks of collaboration among world economists who published in journals during three periods: 1970-1979, 1980-1989 and 1990-1999. Six data sets were compiled, three from the set of all journals covered in Econlit and three from the set of journals by the Tinbergen Institute Amsterdam-Rotterdam (TI). The researchers represented individual economists as vertices in the network. Two persons were said to be connected to each other if they had published at least one paper together. Goyal et al. (2004) then examined the small-world properties in that collaboration network. The results from those six data sets indicated the presence of a small-world phenomenon in the collaboration network. Giant components not only existed, but had grown substantially from fifteen percent in the first period to forty percent in the third period. All six networks were connected. The number of authors was very large, while average degrees of connection overall, and in the giant component, were low, although they had increased over time.

These findings confirmed the sparseness condition ([redacted]). The average distance in the actual giant components had been very small throughout the three periods, smaller than the average in simulated random networks of corresponding size. Although the giant components grew during those three periods, Goyal et al. concluded that they had become significantly smaller in terms of path lengths. This seemed to comply with the condition of length that is small and logarithmically decreasing with respect to the population size ($L(G)$ is of order $\ln(n)$). When they simulated the Erdős-Renyi random networks with size and average degree that were identical to the TI List networks, researchers were able to confirm that the clustering coefficients in the actual networks were much higher than in their random counterparts $(C_{actual} \gg C_{random})$ throughout all three periods. Thus, they concluded that collaborating economists are increasingly forming a small-world network.

Going further, Goyal et al. employed the Lorenz Curve to investigate the degree distribution. They found a fat tail, indicating inequality in the degree of collaboration. Only twenty percent of the most-linked authors accounted for about sixty percent of all links. That meant that a handful of economists were much more connected than the majority. However, the Lorenz Curve also indicated a decrease in inequality over time. When the researchers made a Pareto plot of the degree distribution, the networks decayed under a power law instead of according to a binomial distribution. The author with the most links in the 1970's (25), had 35 in the 80's, and 54 in the 90's. This individual was one of the *stars*, central players which are very important as connectors in the network. The economists' world was spanned by a few inter-linked stars.

Next, Goyal et al. tried to come up with an explanation for the emerging economics small world. If one of the main reasons was an increase in inequality, then all calculated Gini coefficients must increase. Indeed, they found an increase in Gini coefficients for the giant component, which indicated an increase in inequality. By contrast, the Gini coefficient for the overall networks had decreased. Therefore, the researchers concluded that an increase in inequality is not the reason for the emerging small-world properties in the network of economists. One of the possible explanations came from analyzing every period's fat tails in the Pareto plot. It indicated an increase in the average degree of the fixed, inter-linked stars. These stars brought many unconnected economists together, which led to an increased size of the giant component. Another possible explanation offered by several researchers is an increase in distant collaborations due to the drop in communication and travelling costs.

Goyal et al. tried to estimate the relative importance of those two possible explanations by using a two-step procedure, controlling the size and creating remote links, in order to compare the 1970's to the 1980's and the 1980's to the 1990's. Remote link adjustment had a small impact in the form of a slight increase in the giant component. Keeping in mind that the average degree had increased over time, and that the average distance had decreased, the researchers concluded that the main driving force behind the emergence of small-world dynamics in the economics network was not only due to distant collaborations, but to an increase in the average degree at all levels of the collaboration structure.

Goyal et al. (2005) tested Granovetter's argument about the strength of weak ties in the economics collaboration network. Weak ties are significant in the sense that they can build new and shorter paths connecting pairs of vertices. Removing weak ties from the network would break the shortest path and increase the average path length more than removing strong ties. In their 2005 study, Goyal et al. examined the transitivity property of strong-ties and the significance of weak ties. Only one part of the transitivity property of strong ties was supported by findings, while the other parts of the hypothesis were rejected.

They measured the strength of a tie by counting the number of articles over a decade that was co-authored by a pair of associated economists. When an article was written by three co-authors, a close triad was formed. First, researchers examined the transitivity properties by testing cases of triad incompletion and calculating the high probability of triad completion in the presence of two strong ties in a connected triple. They first looked at situations in the social network where A has strong ties with both B and C, but C and B do not have, and do not form, a direct tie. Triad completions in connected triples having two strong ties were lower than 0.5 at every strong tie threshold. Based on these findings, Goyal et al. rejected the hypothesis of the transitivity property of strong ties. However, the logistic regression of random subsamples yielded a result that seemed to support the notion of transitivity of strong ties. The probability of completion in the case of two strong ties in a connected triple increased with respect to the average strength of those two strong ties but decreased with respect to the difference in the strength of the two ties. Thus, regression results supported the transitivity of strong ties in the economics collaboration networks.

The researchers went on to examine the significance of weak ties by arbitrarily removing those that would break the shortest path between actors and would increase the average path length. For the hypothesis test, they calculated the link betweenness for a pair of vertices in the giant component and ran regressions of the random subsample in the three periods. They found that the strength of a tie has a significantly positive impact on link betweenness. In other words, in the collaborating economists' network, strong ties caused higher link betweenness. This undermined the hypothesis that arbitrarily removing weak ties would break the shortest path. Strong ties were found to be crucial in connecting both close and remote actors in the collaborative network.

Next, the researchers simulated and compared the average distances and average size of the giant component after randomly deleting 50 strong, and then 50 weak ties. One result was that the size of giant component was larger when deleting strong ties than when weak ties were deleted. This supported the strength of weak ties. By contrast, another result showed that deleted strong ties led to larger distances than deleted weak ties, which did not support the hypothesis of the strength of weak ties. Although the hypothesis that arbitrarily removing weak ties would break the shortest path was refuted, there remained a contradiction in the empirical tests presented in their paper.

Goyal et. al (2005) attempted to account for these conflicting results by arguing that the collaborative economics network has different structure than the networks in Granovetter's theory. For one, significant inequality indicates that the economics network is connected by a small set of inter-linked stars, which have a much higher than average degree. The star connects economists with much lower degrees. Authors who are members of the star's neighborhood are unlikely to connect to each other directly. The star has high betweenness centrality. He or she connects to other individuals who play a star role. The link between two stars of a high degree has higher link betweenness and is stronger than other links. Hence, the structure of the economics network is characterized by a core-periphery dynamic in which actors within a clique are tied to the core of the clique with weak ties, which form local bridges of a less transitive nature. However, each clique connects to another through the strong tie between stars, which has more transitivity and high link betweenness. This kind of network, then, would satisfy the transitivity property but would not support the overall significance of weak ties.

The economics network structure is different from the ones in Granovetter's theory, which are characterized as *island networks*. The actors within the island, or clique, are tied to each other with strong ties while each island connects to others with weak ties. In that case, the strong-ties within an island are more transitive and the link betweenness of weak ties, spanning the distance between cliques, is higher. Therefore, in island networks, the weak tie connecting entire communities as a bridge is more crucial. By contrast, in the network of collaborating economists, as an example of a core-periphery network, a weak tie connecting a single peripheral player to another player in the periphery is not as important as a strong tie connecting different cliques.

3.3 Model and Data

For the purpose of investigating the small-world phenomenon in the economics academic labor market network, the model proposed by Watts and Strogatz (1998) and Watts (1999) is adopted. Data about the economics academic labor market network was as described in the previous chapter. This chapter will test the economics network to see whether it satisfies the following small-world properties (modified from Goyal et al., 2004):

1. The number of vertices needs to be far greater than the average degree. for the $n >> k >> \ln(n) >> 1$, condition of sparseness to be met and $k >> 1$ to guarantee that the network is connected.
2. The network must be connected or have a largest component for the characteristic path length to be measured.
3. The characteristic path length must be almost as small as the characteristic path length in a corresponding random network, $L \approx L_{random} \sim \dfrac{\ln(n)}{\ln(k)}$.
4. The clustering coefficient must be much greater than that in a corresponding random network: $C >> C_{random} \sim \dfrac{k}{n}$.

To compute the characteristic path length, clustering coefficient, and many structural characteristic of a network, as well as to construct random networks, the social network analysis program Ucinet 6 was utilized.

This chapter will also construct a *Lorenz curve*, which gives a rough measure of the equality of the distribution of in-degree and out-degree of vertices in the network. This measure is known as the *Gini coefficient*. Let the number of vertices in graph G be n. The set of ordered vertices in graph G is S, such that $i < j$ if and only if the degree of vertex i, k_i, is less than the degree of vertex j, k_j, and $k_i < k_j$, for the vertices $i, j \in S$. The number of vertices in S is denoted by n_S ($|S(G)| = n_S$). Total degree in S is denoted by K(n_S), $K(n_S) = \sum_{i=1}^{n_S} k_i$ and the average degree in S is denoted by $k_S = \frac{\sum_{i=1}^{n_S} k_i}{n_S}$.

The proportion of all degrees attributable to each vertex degree is given by: $\frac{k_i}{K(n_S)}$. At any h, $h \in n_S$, the degree accumulation at h is denoted by $k(h)$, $k(h) = \sum_{i=1}^{h} \frac{k_i}{K(n_S)}$. Since the proportion of each vertex in S is $1/n_S$, the accumulation of vertices at h is denoted by $S(h)$, $S(h) = \sum_{i=1}^{h} \frac{n_i}{n_S}$. Then, the Lorenz curve for S is drawn by connecting the points ($S(h)$, $k(h)$), where $k(h)$ and $S(h) \in [0,1]$, for $h = 0, 1, 2, \ldots, n_S$. The Lorenz curve maps the cumulative degree share on the vertical axis against the distribution of universities on the horizontal axis.

The Gini coefficient can be easily viewed by the area between the 45 degree line and the Lorenz curve over the triangle area under the 45 degree line. If each individual university had the same degree, or perfect equality, the Lorenz curve of the graph would be straight and the Gini coefficient would be zero. The Gini coefficient of the network can be measured[11] as follows:

$Gini(S)$ = 1- (area under Lorenz curve). Or,

$$Gini(S) = 1 - \sum_{h=1}^{n_S} \left[\left\{ \sum_{i=1}^{h} \frac{n_i}{n_S} - \sum_{i=1}^{h-1} \frac{n_i}{n_S} \right\} \cdot \left\{ \sum_{i=1}^{h} \frac{k_i}{K(n_S)} + \sum_{i=1}^{h-1} \frac{k_i}{K(n_S)} \right\} \right]$$

[11] The Gini coefficient is modified from http://www.techwranglers.net/dtree/gini.html. And, it can be reduced to the form as in Goyal et al. (2004, p. 5).

$$Gini(S) = 1 - \sum_{h=1}^{n_S} \left[\frac{1}{n_S} \cdot \left\{ \sum_{i=1}^{h} \frac{k_i}{K(n_S)} + \sum_{i=1}^{h-1} \frac{k_i}{K(n_S)} \right\} \right]$$

$$Gini(S) = 1 - \frac{1}{n_S} \sum_{h=1}^{n_S} \left\{ \sum_{1=1}^{h} \frac{k_i}{K(n_S)} + \sum_{i=1}^{h-1} \frac{k_i}{K(n_S)} \right\}$$

$$Gini(S) = 1 - \frac{1}{n_S} \sum_{h=1}^{n_S} \{k(h) + k(h-1)\}$$

If only one university possessed the total degree *K(n)* in the network, Gini coefficient is equal to one due to the Lorenz curve drawn by the line passing through the points (0, 0), (1, 0) and (1, 1). This would mean perfectly inequality of the degree distribution in the network. A high Gini coefficient therefore means high inequality in the degree distribution in the network.

It has been argued that the Gini coefficient has disadvantages when representing the inequality in a subgroup of the population, or a subgraph. That is, the Gini coefficient of a network cannot be summarized from Gini coefficients calculated from its subgroups. In addition to the Gini coefficient, the *Theil Index* of inequality is computed as well, which is summation across different subgroups. While the Gini coefficient includes everything that affects the degree distribution, the Theil Index allows for the decomposition of inequality within and between subgroups. The Theil Index can distinguish whether the inequality is due to any subgroup or not. It is the weighted average of degree relative to its mean.

The Theil Index (as cited in Akita et al., 1999 and Collier, 1999) is computed as follows. Let the number of vertices in graph *G* be *n*. The total degree in the network is denoted by *K(n)*, $K(n) = \sum_{i=1}^{n} k_i$, and the mean degree of all vertices in graph G, denoted by k, is defined as $k = \frac{\sum_{i \in V(G)} k_i}{n}$ Let graph G be divided into *g* subgraphs. The Theil Index of the overall graph, denoted by T, is given by:

$$T=\sum_{i=1}^{n}\left[\frac{k_i}{K(n)}\cdot\ln\left(\frac{k_i}{k}\right)\right]$$

Under perfect equality, everyone has the mean degree. Therefore, T=0 means no inequality. In perfect inequality, when one university has the total degree in the network and every other university has no degrees, T=ln(n). While there is no upper limit for inequality when given as a Gini coefficient, a higher value of the Theil Index corresponds to a higher level of inequality.

Let the set of vertices in graph G be Z(G), Z(G)=$\{Z_1, Z_2, \ldots, Z_g\}$, g>1 and there is no overlap between different subgraphs. The number of vertices in subgraphs $Z_1, Z_2, \ldots, Z_g$ are n_1, $n_2, .., n_Z$, respectively, where $n_1+n_2+\ldots+n_g=n$. The total degrees in subgraphs $Z_1, Z_2, \ldots, Z_g$ are $K(n_1), K(n_2), \ldots, K(n_g)$, respectively. The shares of these subgraphs in graph G are $s^1, s^2, \ldots, s^g$, where $s^1+s^2+\ldots+s^g=1$. The average degrees in subgraphs $Z_1, Z_2, \ldots, Z_g$ are $k^1, k^2, \ldots, k^g$, respectively. The Theil Index for any subgraph b is, given by:

$$T_b=\sum_{i=1}^{n_b}\left[\frac{k_i^b}{K(n_b)}\cdot\ln\left(\frac{k_i^b}{k^b}\right)\right].$$

With g subgraphs, the Theil Index can be decomposed into Theil-within-group, denoted as T^{within}, and Theil-between-groups, denoted by T^{between}. The Theil-within-group is the inequality index relative to a group average. It is the weighted average of Theil indices between subgroups in the graph, where the weight is the degree share of each group in the total degree. The Theil-between-group is the Theil Inequality index relative to the entire graph average. Then, the Theil Index decomposition of inequality can be computed as follows:

$$T=T^{within}+T^{between}$$

$$T^{within}=\sum_{b=1}^{g}s^b T_b \quad \text{and} \quad T^{between}=\sum_{b=1}^{g}\left\{s^b\cdot\ln\left(\frac{k^b}{k}\right)\right\}$$

3.4 Empirical Results

Data used in this study was collected in April-July 2006 as described in Chapter 2. Due to the construction of the network, data is transformed into a square matrix, where each university functions as both the grantor and employer. This implies that some unranked universities who were able to place their graduates into ranked universities were dropped as their hiring decisions were not observed. The economics academic network is described by the grantor-employment matrix A, where A_{ij} is the number of faculty graduating from university *i* and employed by university *j*. It is a weighted-directed-network, in which universities are the vertices and the number of Ph.D. graduates placed and hired form the edges, or links, between them. The network matrix constructed from the full sample of 200 ranked universities is referred to as the "*square ranked network*." Figure 3.1 shows the square ranked network. The arrow lines show the source and target of exchange (self-hires are not shown). The arrows indicate transfer from the grantor to the employer. The thickness of each line indicates tie strength. The figure shows that all universities in the network are connected. There is only one single component in the network and no isolated universities. Figure 3.2 shows Harvard University's neighborhood, its so-called *ego network*, which reveals a number of important features of the square ranked network. For instance, Harvard has placed eight Ph.D. graduates in MIT and four in Oxford. It employs thirteen Ph.D. graduates from MIT and one from Oxford. Although Harvard has bilateral relationships with both MIT and Oxford, the relationship with MIT is stronger. Single-direction relationships are exemplified by the fact that Harvard has placed twelve graduates at Columbia, seven at Boston University, and one at UC Riverside, but does not employ a single graduate of any of these institutions. Harvard, then, has stronger ties with Columbia than with Boston University, which is a stronger tie than that with UC Riverside.

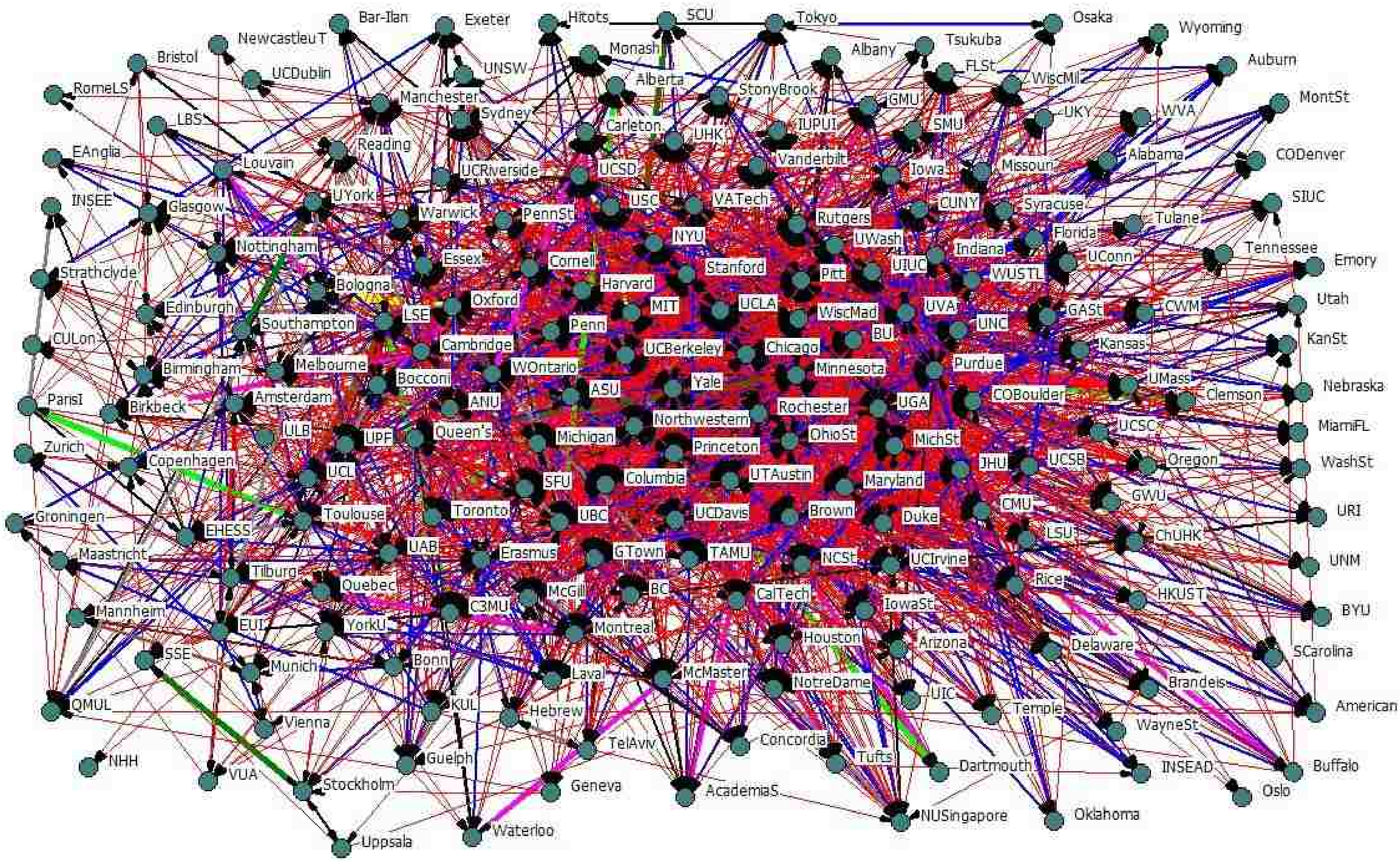


Figure 3.1: Connections between Universities in Square Ranked Network

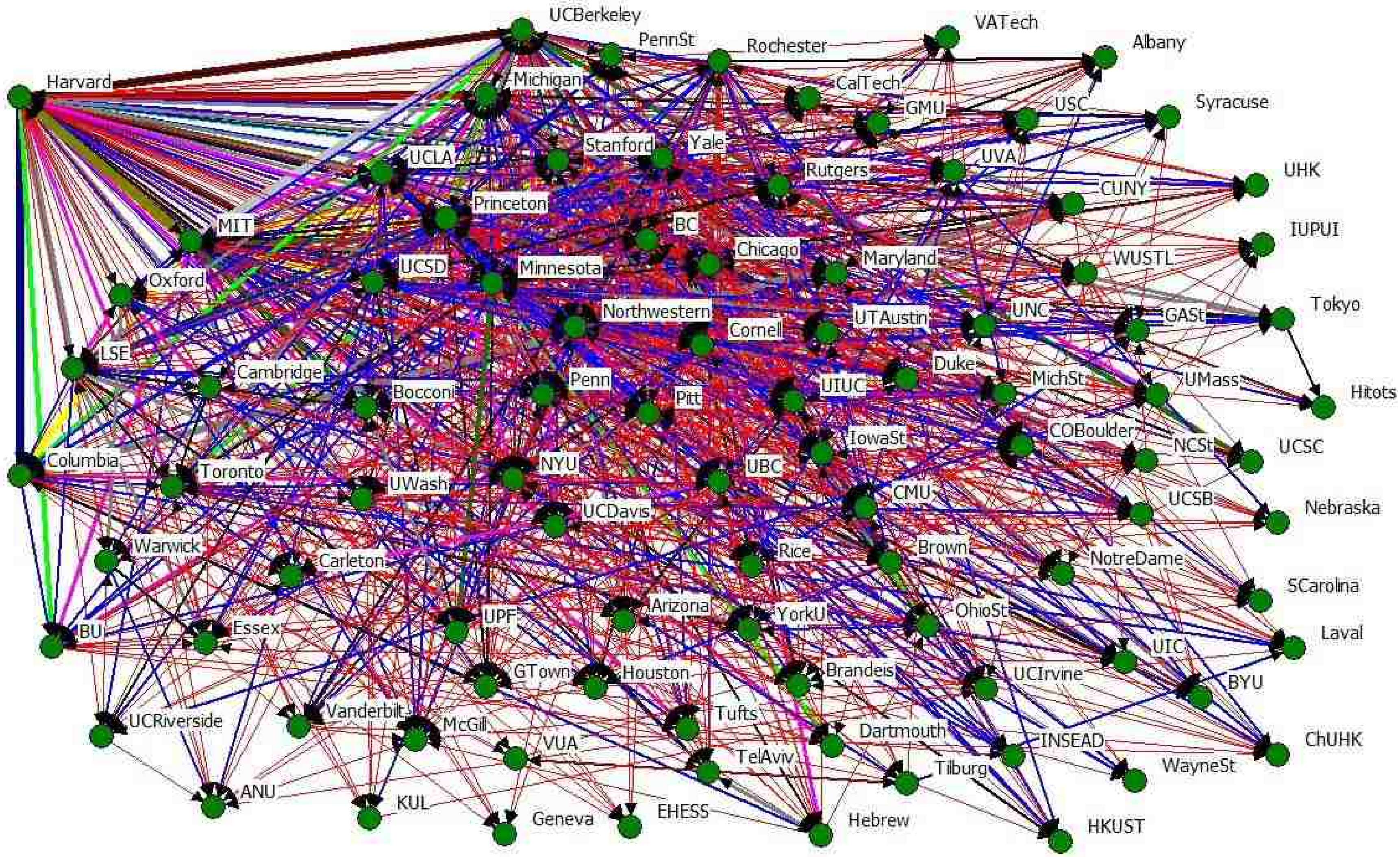


Figure 3.2: Harvard's Ego Network within the Square Ranked Network

When the network is divided into eight groups where each group approximates a quality level, with each level consisting of twenty-five universities, then Harvard, MIT, and Columbia are all in Group 1, the highest group. Oxford and Boston University fall into Group 2, which is ranked lower than Harvard. UC Riverside falls into Group 6, which is far below Group 1. The examples above indicate that the strongest relationships are formed within the same quality group, while relationships with lower levels are weaker. This pattern speaks to a hierarchy of placement and employment in the economics academic market.

Figure 3.3 presents Oxford University's ego network. It provides an example of a European institution which hires its own products at a much higher rate. Oxford hires about 42 percent of its faculty from within the square ranked network. Although it interacts with universities inside and outside Europe, there is a sense of stronger ties within Europe. For example, Oxford University has placed five graduates at University of College London (UCL), five at London School of Economics (LSE), and nine at Bologna, while placing only two at MIT, one at Harvard, and two at UC Berkeley. On the other hand, Oxford has hired five graduates of Cambridge, four of LSE, and four of UC Berkeley.

The academic labor market in North America seems more uniform because universities on other continents have a much greater percentages of self-hiring, which causes high value in diagonals. The diagonal values in the valued graph are diagonal elements of the grantor-employer matrix and refer to the number of self-hiring of each university. Chapter 2 discussed about self-hiring across universities. While within the North America, the single highest self-hiring rate is less than 30% of total faculty, non-North American universities have much higher rates of self-hiring. Oxford hires 42% of its faculty from itself, while a half of faculty at Cambridge are graduates of Cambridge. This chapter, as the previous one, will focus on North

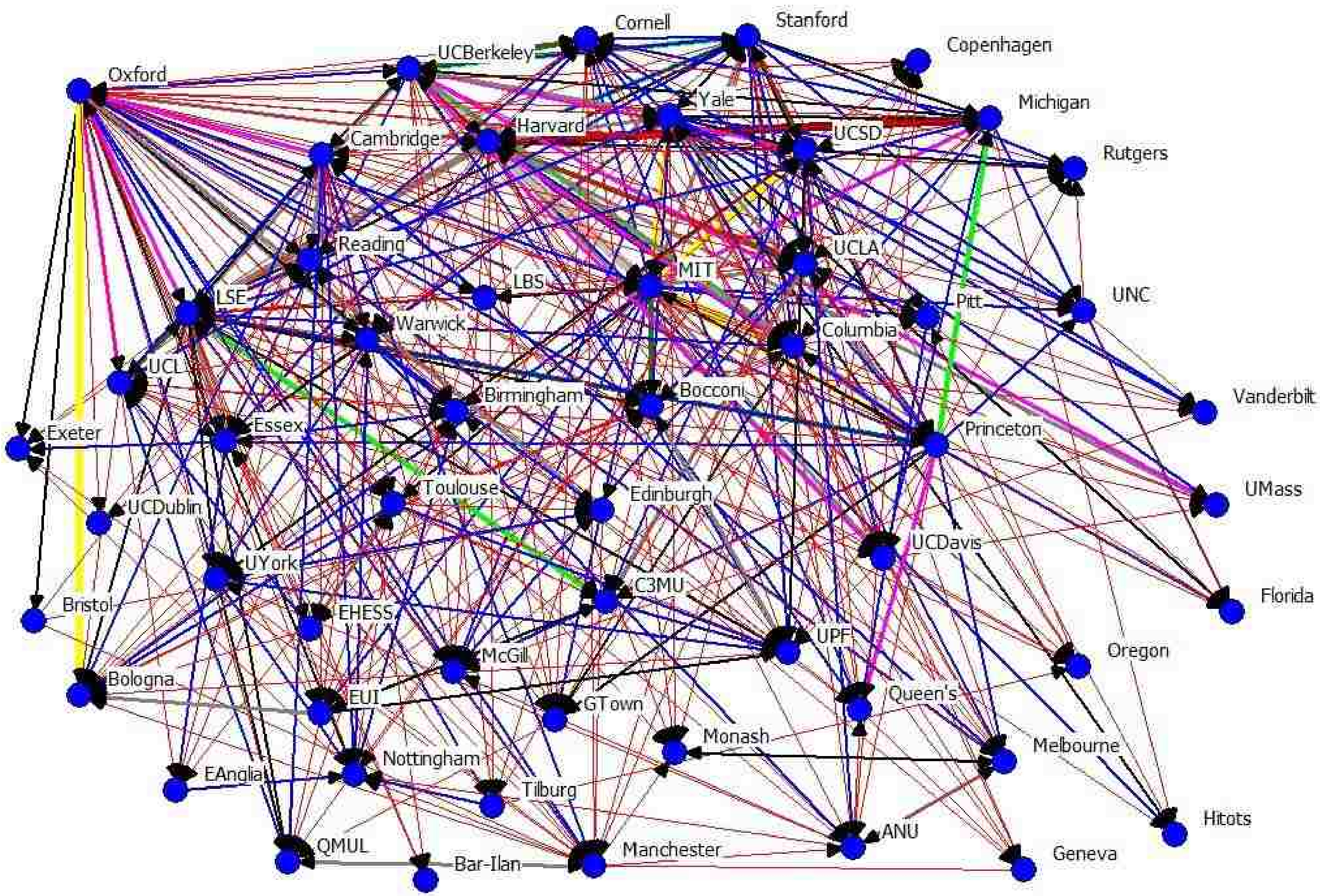


Figure 3.3: Oxford's Ego Network within the Square Ranked Network

America's ranked university network composed of 126 US and Canadian universities. Figure 3.4 shows the interaction between the 126 universities in the North American Square Ranked Network. Like the overall square ranked network, the North American square ranked network is connected and has only one component. The relationship between different universities in the North American Square Ranked Network seems denser than in the square ranked network. The reciprocal interactions are more present in the North American Square Ranked Network.

In order to analyze and test for small-world properties in the academic network, its matrix was transformed to represent as an unweighted-directed-unlooped network. The transformed matrix Y consists of element $Y_{ij}=1$, if $A_{ij}>1$, $i \neq j$, $Y_{ij}=0$ if $A_{ij}=0$, $i \neq j$ and $Y_{ii}=0$. This representation helps distinguish between in- and out-degree. Out-degree captures placement of graduates. In-degree captures employment of graduates from one institution by another. The total number of faculty in the valued networks is 4,783 individuals in the square ranked network and 3,026 individuals in the North American Square Ranked Network (see Table 3.1 and 3.2). After transforming networks from valued into unweighted networks, the total in-degree and out-degree connections are 2,646 and 1,739 in the ranked university network and the North American ranked university network, respectively.

In the valued or weighted network the row summation is the number of graduates placed by a university, in the unweighted network it is the number of universities where one university was able to place its graduates – the number of degrees is less than the number of faculty in the networks. For example, Harvard placed 239 graduates in the overall square ranked network, representing a total of 90 universities (excluding Harvard itself). Harvard's out-degree in the valued network is 239 and only 90 in the unweighted network. Harvard's in-degree in the valued network is 53 (faculty size), but only 15 in the unweighted network (total number of different

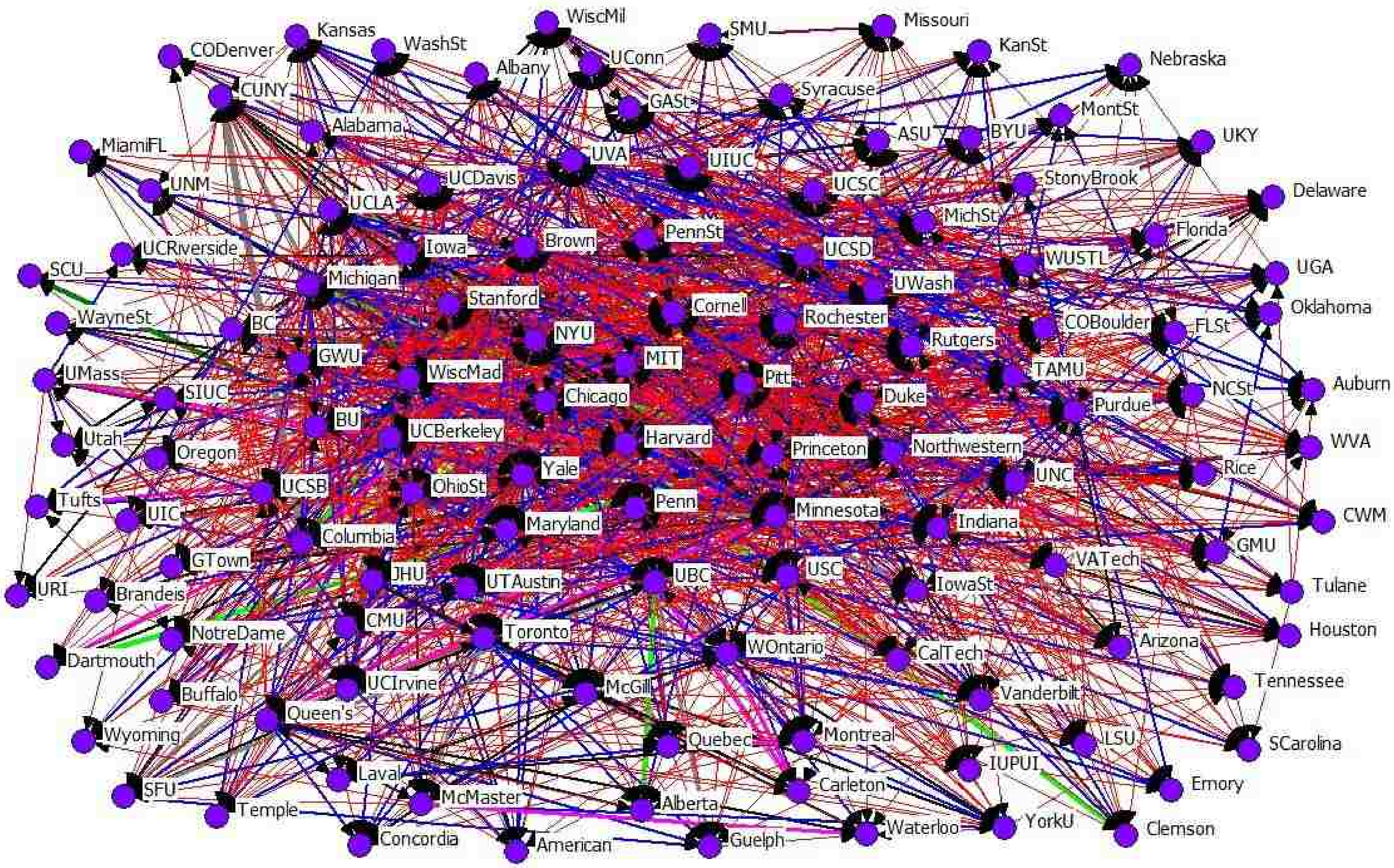


Figure 3.4: Connections between Universities in North American Square Ranked Network

Table 3.1: Real-World and Random Network of Ranked Universities Compared

	Real-World Network	Random Network	
	Square Ranked Network	Erdos and Renyi's model	Average of 100 Random Graphs
Total Universities	200	200	200
Total Number of Faculty	4783		
Total In-Degree	2646		2667
Average In-Degree	13.23	13.23	13.34
Standard Deviation of In-Degree	5.4		3.53
Total Out-Degree	2646		2667
Average Out-Degree	13.23		13.34
Standard Deviation of Out-Degree	21.13		4.14
Density	0.07		0.07
Standard Deviation of Density	0.24		
Characteristic Path Length	2.93	2.05	2.32
Distance-Weighted Fragmentation	0.69		0.53
Clustering Coefficient	0.24	0.07	0.07
Weighted Clustering Coefficient	0.17		0.07

Note: The square ranked network was transformed into an unweighted directed un-looped network. 100 Random graphs were generated, with the same density and size as the ranked university network. In-Degree for random network is the average employment value calculated from the set. Out-Degree value is average placement.

universities able to place graduates at Harvard). Even though the number of in-degree and out-degree for each university are different, in the square matrix the summation of these parameters are the same.

3.4.1 Small-World Properties Examined

The small-world properties will be examined in the order they were introduced earlier. Tables 3.1 and 3.2 show the characteristics of the transformed square ranked network and the transformed North American square ranked network, respectively.

There are 200 total vertices in the square ranked network and 126 vertices in the North American square ranked network. The average in-degree and out-degree is 13.23 in the former and 13.80 in the latter. The average degree in the square ranked network is greater than 2.30, which is the logarithmic value of the number of vertices in the network ($ln(200) \approx 2.30$). The

Table 3.2: Real-World and Random Network in North America Compared

	Real-World Network	Random Network	
	North American Square Ranked Network	Erdos and Renyi's Model	Average of 100 Random Graphs
Total Universities	126	126	126
Total Number of Faculty	3026		
Total In-Degree	1739		1732
Average In-Degree	13.80	13.80	13.80
Standard Deviation of In-Degree	3.6		3.45
Total Out-degree	1739		1732
Average Out-degree	13.80		13.80
Standard Deviation of Out-Degree	19.76		4.05
Density	0.11		0.11
Standard Deviation of Density	0.31		
Characteristic Path Length	2.80	1.84	2.09
Distance-Weighted Fragmentation	0.67		0.48
Clustering Coefficient	0.27	0.11	0.11
Weighted Clustering Coefficient	0.20		0.11

Note : The North America square ranked valued network was transformed into an unweighted directed un-looped network. 100 Random graphs were generated with the same density and size. In-Degree value for random network is the calculated average employment across all 100 random graphs. Out-Degree is average placement.

average degree in the North American square ranked network is also greater than the logarithmic value of the number of vertices in the network (ln(126) ≈ 2.00). Both connected networks have a number of vertices far greater than the average degree. This number is also greater than their logarithmic values ($n>>k>>ln(n)>>1$). The first condition for a small world network is therefore satisfied in both the square ranked network and the North American square ranked network.

Due to the fact that both the square ranked network and the North American square ranked network have only one single component, the characteristic path length pertains to each graph as a whole. The characteristic path length, which is the average of the shortest distance from one university to any other, is 2.80 in the North American square ranked network, a little shorter than the 2.93 of the overall square ranked network. In other words, an economics department in one North American university can reach any other on the same continent with a slightly shorter chain than in the global ranked university network. Most top 10 ranked universities connect directly to each other. The maximum shortest path length between them, in both networks, is only 2. For example, the distance between Harvard and Chicago, University of Pennsylvania, and MIT is one. The distance between Chicago and MIT to Harvard is also one; from University of Pennsylvania to Harvard is two.

The distance between lower ranked universities to the higher ones or within their same group of quality is longer. For example, the distance from University of Southern California (USC), located in Group 2, to Harvard is five, which is higher than distance to any other university in Group 2 but University of North Carolina (UNC) and Vanderbuilt. Since some universities are unable to place graduates at ranked universities, some distances in the network cannot be computed. Those universities, then, are considered to be somewhat disconnected. The distance weighted fragmentation, shown in Tables 3.1 and 3.2, calculates the distance weighted by the number of connected vertices. The distance weighted fragmentation in the North American square ranked network is 0.67, quite a bit smaller than 0.96 in the square ranked network. The distance weighted fragmentation of these real networks is quite close to random networks. Hence, in both networks the second condition for small-worlds is satisfied.

Unsurprisingly, the density value in the North American ranked network (0.11) is greater than the value in the ranked university network (0.07), of which it is a subset. Within the North American ranked university network, the chance of sharing information or exchanging Ph.D. graduates is higher than in the ranked university network. The standard deviation values of density in both networks differ only slightly.

The clustering coefficient in either network is higher than its density. This is because the clustering coefficient is measured with the local neighborhood, which is of much higher density than that of the overall network. However, the clustering coefficient of the square ranked network is 0.24, which is slightly smaller than the 0.27 of the North American square ranked network. In other words, the neighborhood of one university in the square ranked network relates to each other less than in the North American square ranked network. This fact might indicate that universities in North America rely on connections in the network more than universities outside of that continent. The North American square ranked network has a size weighted clustering coefficient to of 0.20, compared with 0.17 for the square ranked network.

The characteristic path length and clustering coefficients of the respective Erdos and Renyi random networks are also given in Tables 3.1 and 3.2. The characteristic path length calculated for a random network with the same number of vertices and average degree as the square ranked network is 2.05 and 1.84 for the North American square ranked network. The characteristic path length in both real-world networks is not much greater than that in the Erdos and Renyi's random networks (2.93 and 2.80). Thus, it can be concluded that the third condition for a small-world phenomenon is satisfied by both networks.

The clustering coefficient of an Erdos and Renyi random network with the same number of vertices and average degree is 0.07 for the ranked university network and 0.11 for the North American ranked network. Both clustering coefficients in the real-world networks are much higher than those of their random counterparts. Thus, the last condition for a small-world network is also satisfied.

In order to make these above comparisons more reliable, one hundred Erdos and Renyi random networks with the same number of vertices and density values were constructed by using

Ucinet6. The average characteristics of these random networks are presented in Tables 3.1 and 3.2. In Table 3.1, the average value of the characteristic path length and clustering coefficient of random networks are 2.32 and 0.07, respectively. In Table 3.2, the average value of the characteristic path length and clustering coefficient of random networks are 2.09 and 0.11, respectively. The third and fourth conditions are confirmed by comparisons with multiple random networks. Since they satisfy all conditions, both real-world networks can be said to exhibit properties of the small-world phenomenon.

3.4.2 Inequality and Centrality of Small-World Networks

Having established the small-world properties of the economics academic network, we can proceed to compare the entire network and its subgroups. The dividing principle for these subgroups is level of quality as determined by university rankings. Using this principle, eight groups of ranked universities and five groups of North American universities are constructed as described in Chapter 2.

Table 3.3 presents a comparison between the eight groups and the entire ranked university network. Even though the number of universities in each group is equal, the total and average in- and out-degree are concentrated in the higher-quality groups, particularly Group 1. While the average in- and out-degree in the entire network is 13.23, the highest group average in- and out-degree of 11.44 belongs to Group 1. The second highest, belonging to Group 2, is only 2.50. A similar comparison in the North American square ranked network, presented in Table 3.4, outlines the same trend. It would seem, then, that some degree of inequality exists in both networks.

The in-degree and out-degree distributions of both ranked university and North American ranked university networks are given as log-log plots in Figures 3.5-3.8. If the log-log plot of

Table 3.3: Compare Subgroups in Square Ranked Network

	Square Ranked Network	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
Number of Faculty in Networks	4783	995	706	648	610	579	532	593	418
Total In-Degree	2646	286	63	45	17	17	5	6	9
Total Out-Degree	2646	286	63	45	17	17	5	6	9
Number of Universities	200	25	25	25	25	25	25	25	25
Average In-degree	13.23	11.4	2.5	1.8	0.68	0.68	0.2	0.24	0.36
Standard Deviation of In-Degree	5.4	2.79	1.65	1.47	1.09	0.84	0.49	0.59	0.63
Average Out-Degree	13.23	11.4	2.5	1.8	0.68	0.68	0.2	0.24	0.36
Standard Deviation of Out-Degree	21.16	7.8	2.52	1.57	1.09	0.676	0.49	0.51	0.63
Density	0.066	0.48	0.105	0.075	0.028	0.028	0.008	0.01	0.02
Standard Deviation of Density	0.2491	0.5	0.307	0.263	0.166	0.165	0.091	0.09	0.12
Characteristic Path Length	2.93	1.5	3.189	2.907	1.19	1.838	1	1.143	1.357
Clustering Coefficient	0.24	0.55	0.136	0.059	0.043	0.131	0	0.389	0
Diameter	7	1	10	9	2	4	1	2	2

Note: Prior to these calculations, the ranked university valued network and subgroups were transformed into an unweighted directed un-looped network.

Table 3.4: Compare Subgroups in the North American Square Ranked Network

	North American Square Ranked Network	Group 1	Group 2	Group 3	Group 4	Group 5
Number of Faculty in Networks	3026	978	688	616	506	481
Total In-Degree	1739	280	72	39	15	8
Total Out-Degree	1739	280	72	39	15	8
Number of Universities	126	25	25	25	25	26
Average In-Degree	13.8	11.2	2.88	1.56	0.6	0.31
Standard Deviation Of In-Degree	3.6	2.53	1.95	1.55	0.63	0.606
Average Out-degree	13.8	11.2	2.88	1.56	0.6	0.31
Standard Deviation Of Out-Degree	19.76	7.84	2.94	1.92	0.75	0.54
Density	0.11	0.4667	0.12	0.065	0.025	0.0123
Standard Deviation of Density	0.3134	0.4989	0.325	0.2465	0.156	0.1103
Characteristic Path Length	2.8	1.578	2.629	1.963	1.25	1.417
Clustering Coefficient	0.267	0.54	0.211	0.128	0	0
Diameter	8	4	7	5	2	3

Note: Prior to these calculations, the North America ranked university valued network and subgroups were transformed into an unweighted network.

the degree distribution is linear, it follows a power law distribution. The log-log plot of in-degree distribution in both networks, given in Figures 3.5 and 3.6, is not quite linear. In both figures, the tail decays faster than under the power law, and seems to fit the description of a Gaussian distribution. Because of this characteristic, both networks should be classified as *single-scale*. However, the log-log plot of out-degree distribution for both networks (Figures 3.7 and 3.8) initially seems to have a linear trend but is then truncated. Therefore, due to their out-degree distribution, both networks are instead classified as *broad-scale*.[12]

[12] For the classification of small-world networks as: scale-free, broad-scale and single-scaled by Amaral et. al. (2000) please refer to Chapter 2.

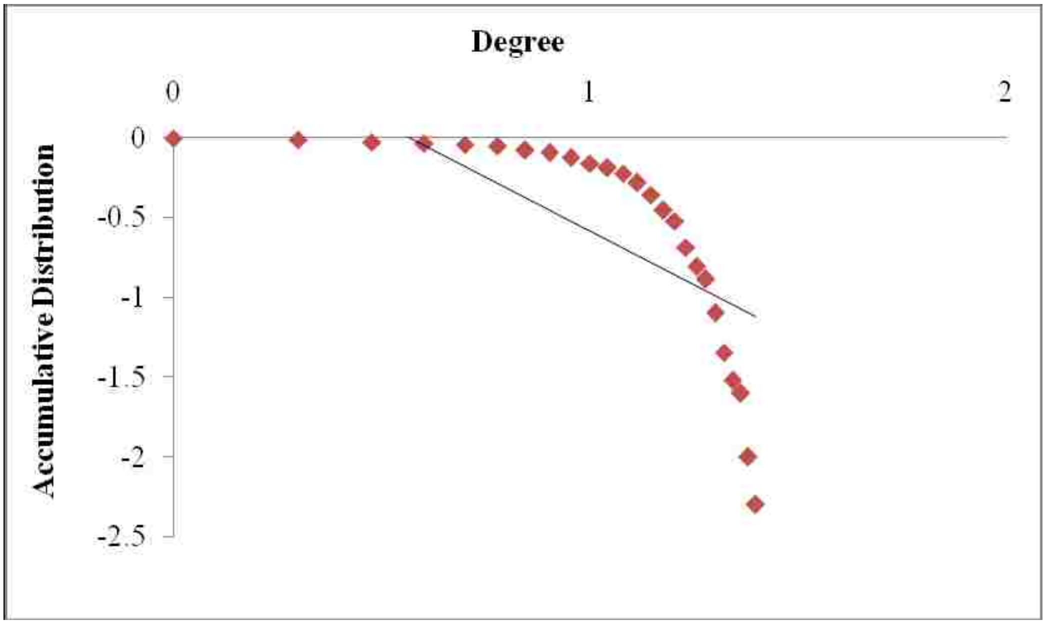


Figure 3.5: Log-Log Plot (P(X>a)) of In-Degree Distribution for Square Ranked Network

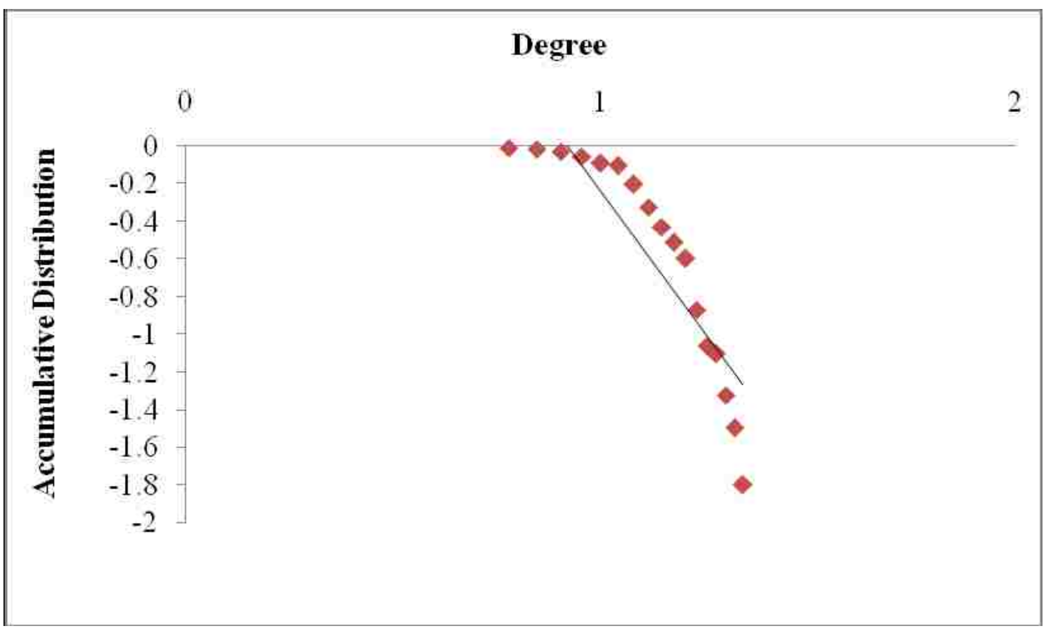


Figure 3.6: Log-Log Plot (P(X>a)) of In-Degree Distribution for the North American Square Ranked Network

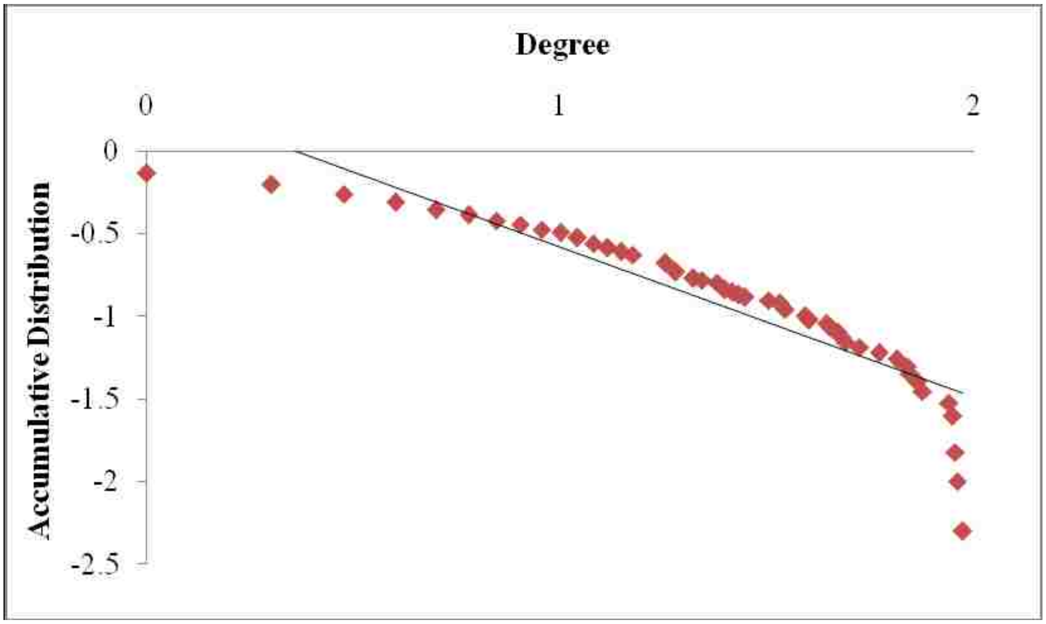


Figure 3.7: Log-Log Plot (P(X>a)) of Out-Degree Distribution for the Square Ranked Network

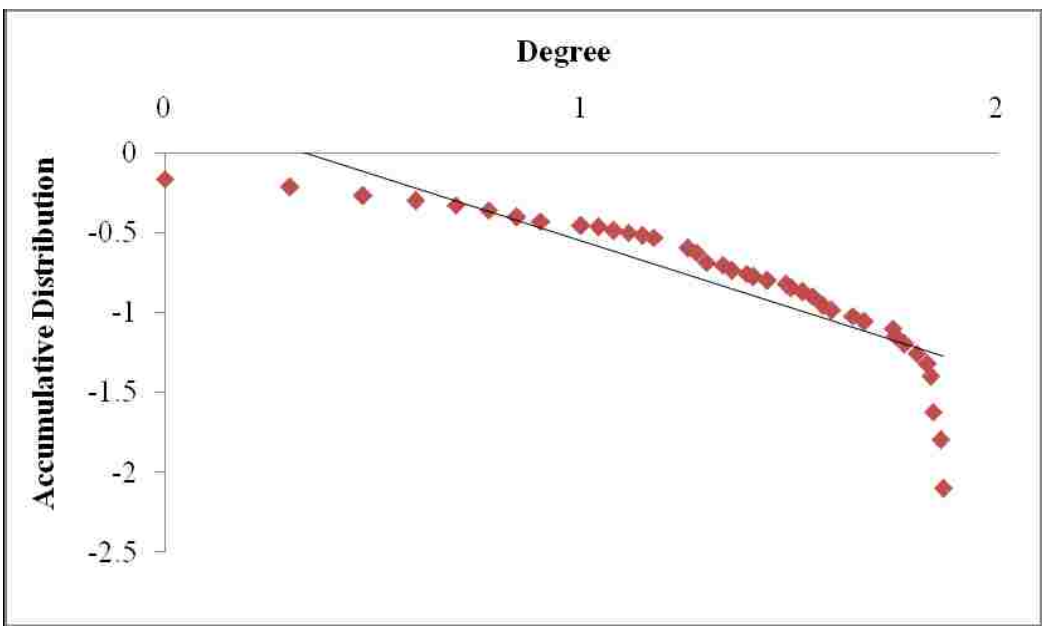


Figure 3.8: Log-Log Plot (P(X>a)) of Out-Degree Distribution for the North American Square Ranked Network

Only one percent of universities in both networks have an out-degree exceeding 40. Judging from the degree distribution, this fact does not point to an inequality of employment, but it does spell inequality of placement among institutions. This inequality can be examined further by drawing a Lorenz curve and measuring the Gini coefficient and Theil index.

The Lorenz curves are also drawn for the in-degree of the ranked university network and North America ranked university network in Figures 3.9 and Figure 3.10, respectively. Both curves are quite close to the perfect equality line. Moreover, the Gini coefficients calculated for the in-degree of both networks are very small, at about 0.23 and 0.15, respectively (see Table 3.5). This means that the universities in the ranked university network, and even more so in the North American university network, are quite equal in terms of in-degree. In other words, when hiring alone is considered, relationships between universities in the networks are more or less equal.

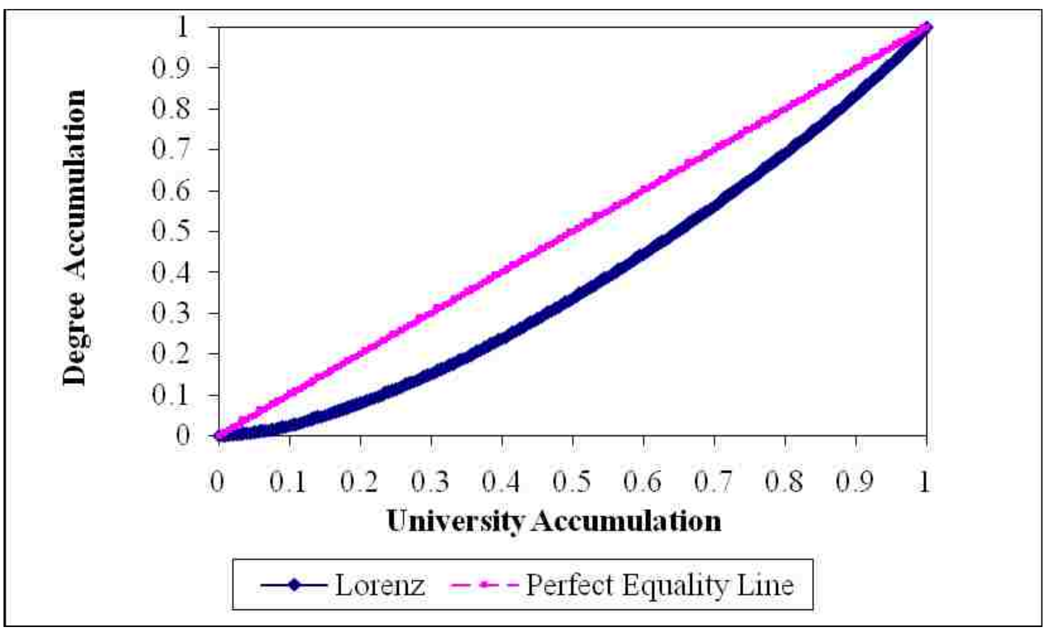


Figure 3.9: Lorenz Curve of In-Degree for the Square Ranked Network

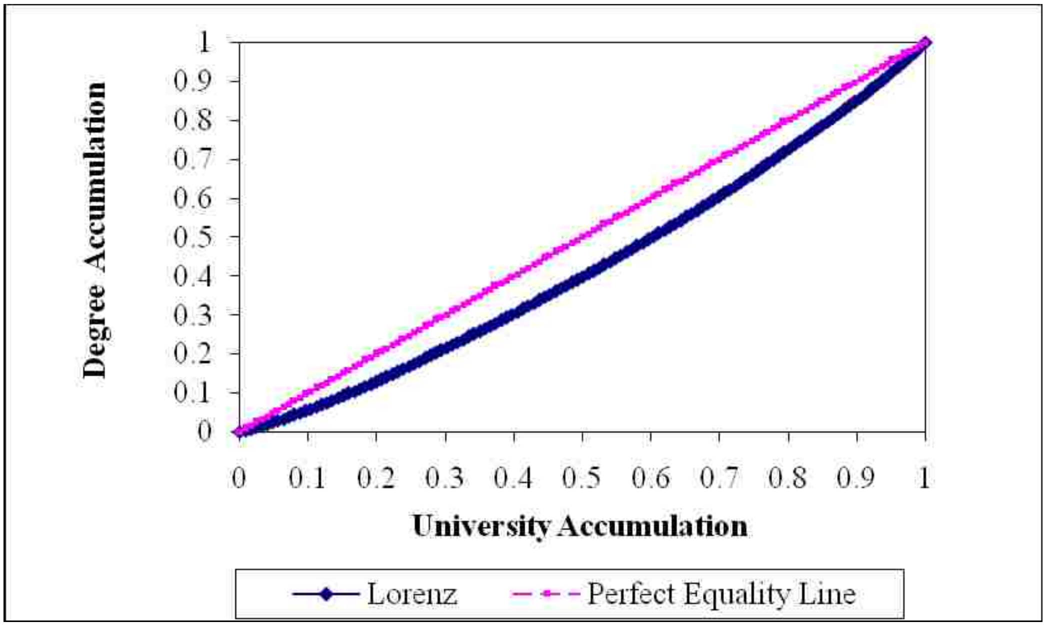


Figure 3.10: Lorenz Curve of In-Degree for the North American Square Ranked Network

Table 3.5: Summary of Inequality Indices

	Square Ranked Network		North American Square Ranked Network	
	In-Degree	Out-Degree	In-Degree	Out-Degree
Gini Coefficient	0.231	0.689	0.146	0.675
Theil Index (overall)	0.096	0.972	0.029	0.828
Within-Theil	0.084	0.296	0.029	0.208
Between-Theil	0.012	0.676	0.005	0.619
Theil Group 1	0.015	0.126	0.018	0.133
Theil Group 2	0.034	0.612	0.028	0.172
Theil Group 3	0.063	0.268	0.028	0.655
Theil Group 4	0.084	0.462	0.023	0.485
Theil Group 5	0.224	0.298	0.053	0.485
Theil Group 6	0.108	0.570		
Theil Group 7	0.060	0.936		
Theil Group 8	0.140	0.658		

On the other hand, the Lorenz curves of out-degree for both ranked university network and North American ranked university network, given in Figures 3.11 and 3.12, diverge

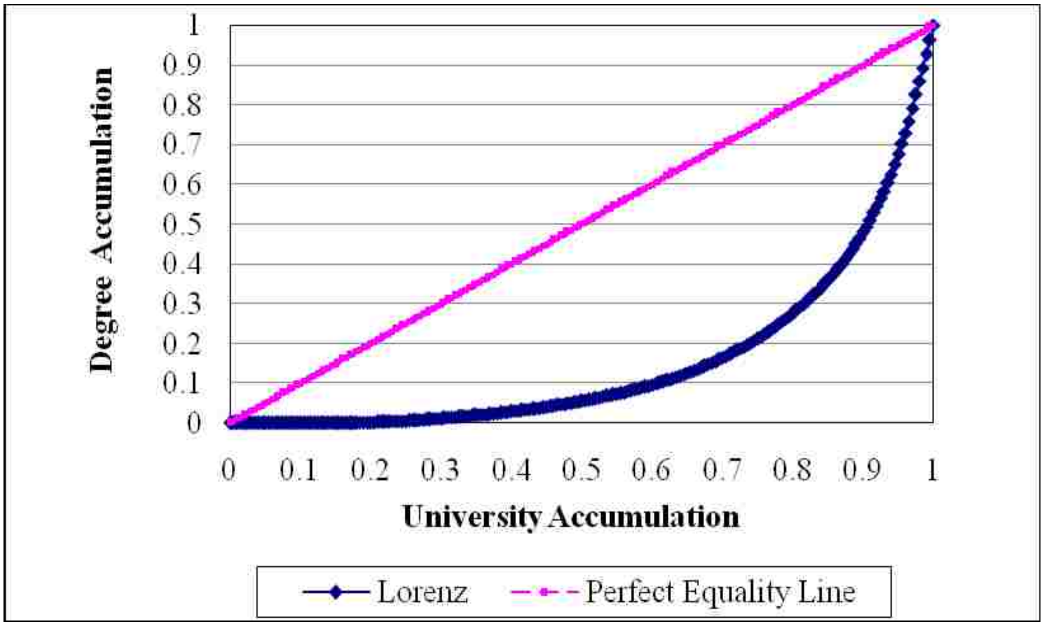


Figure 3.11: Lorenz Curve of Out-Degree of Square Ranked Network

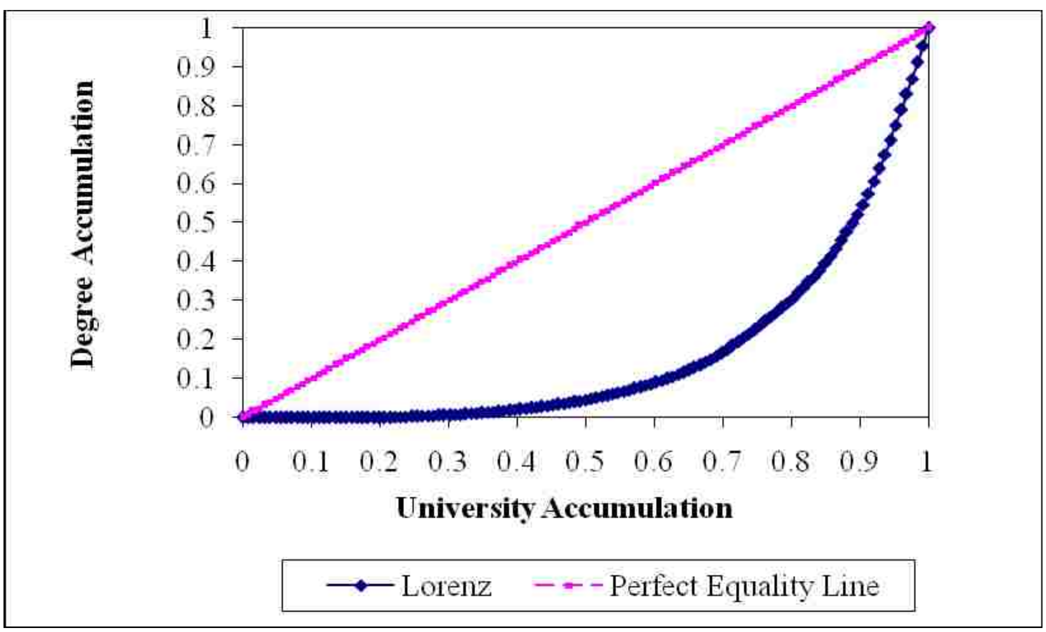


Figure 3.12: Lorenz Curve of Out-Degree for the North American Square Ranked Network

significantly from the perfect equality line. The Gini coefficient of out-degree for the former is about 0.69; for the latter it measures about 0.68 (see Table 3.5). Both the Lorenz curve and Gini coefficients indicate a high level of inequality in out-degree in both real-world networks. In other words, when placement of graduates is considered, relationships between universities in the networks are unequal.

The Theil Index, presented in Table 3.5, supports similar conclusions about in-degree equality and out-degree inequality. The overall Theil indices of in-degree for both the overall and North American square ranked networks are very small, which indicates equality among employers. But the Theil indices of out-degree indicate some inequality in placement of graduates. However, this inequality stems not from imbalance within the group, but from imbalance between groups. The inequality of out-degree in the ranked university network is highest in Group 7, which is of the second lowest quality. In the North American square ranked network, Group 3 exhibits greatest inequality. In this group, the highest out-degree belongs to Purdue, which can place its Ph.D. graduates to 32 universities. There are only three other universities which have out-degree greater than 10. Many can place to very few universities and four universities cannot place their Ph.D. graduates into the network at all. Group 1 in either network does not have any inequality in either placement or employment (see Figures 3.13 through 3.20).

When the networks are observed closely by group, the following features come to light: Group 1 and Group 2, the two highest groups, have higher densities than the entire network (see Table 3.3 and 3.4). For example, while the densities of Group 1 and Group 2 in North America

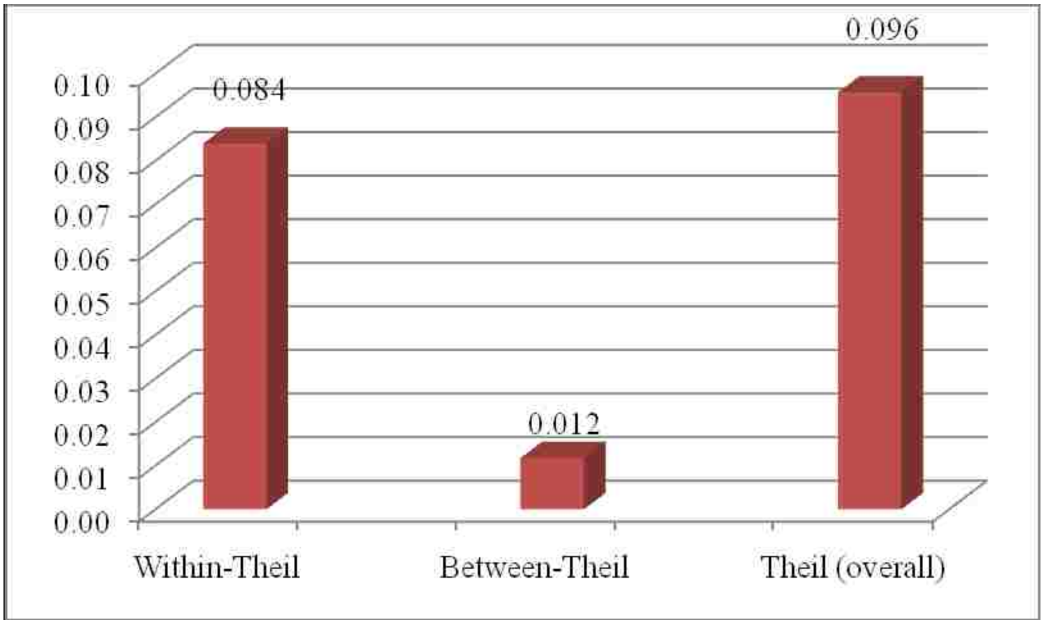


Figure 3.13: Theil Index of In-Degree for the Square Ranked Network

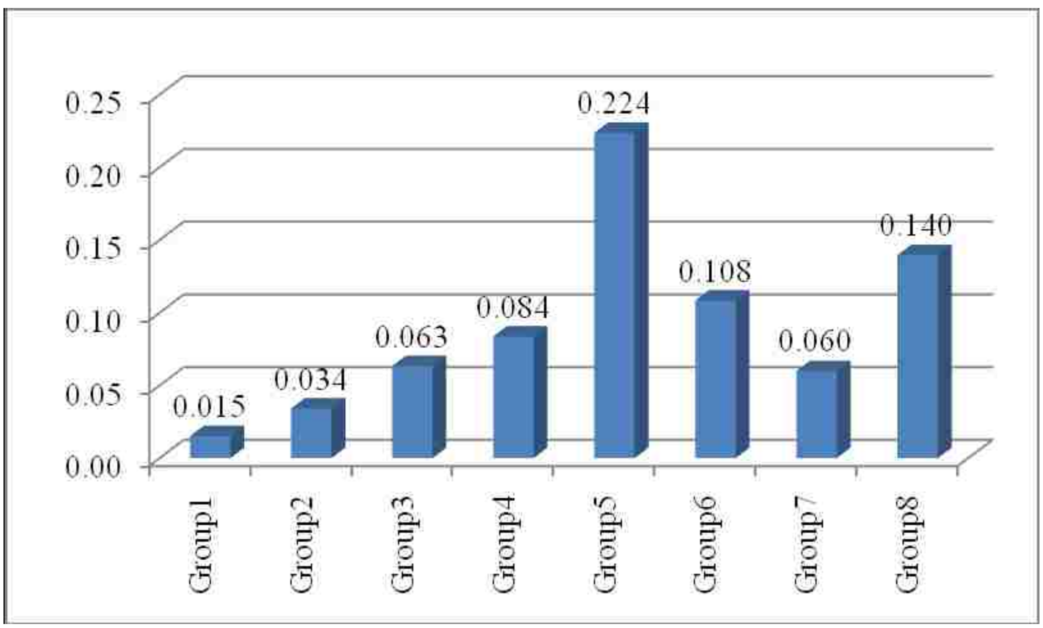


Figure 3.14: Theil Index of In-Degree of All Eight Groups in the Square Ranked Network

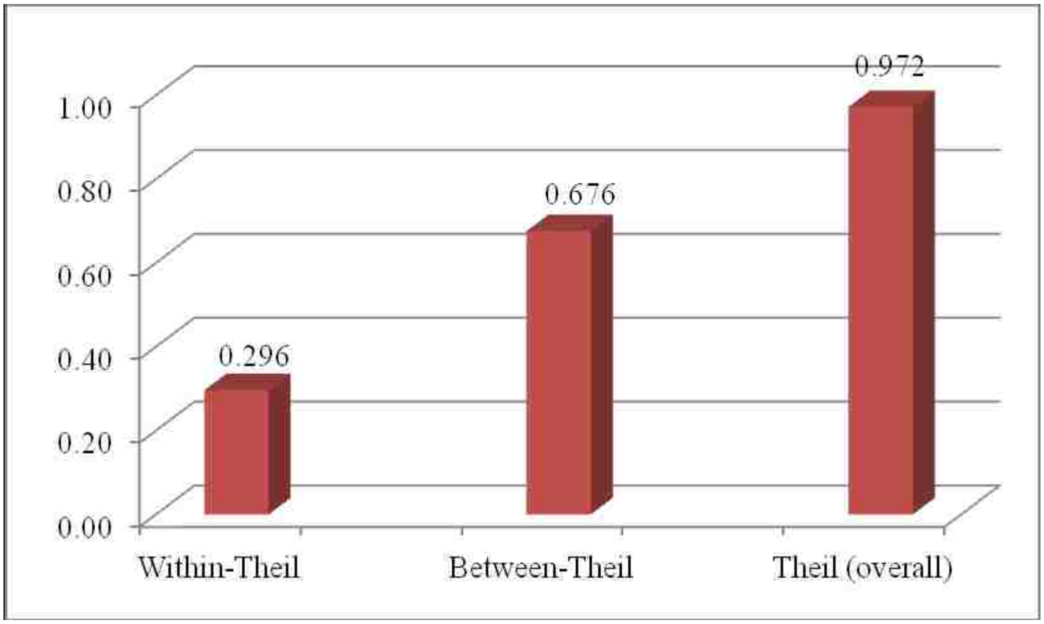


Figure 3.15: Theil Index of Out-Degree in the Square Ranked Network

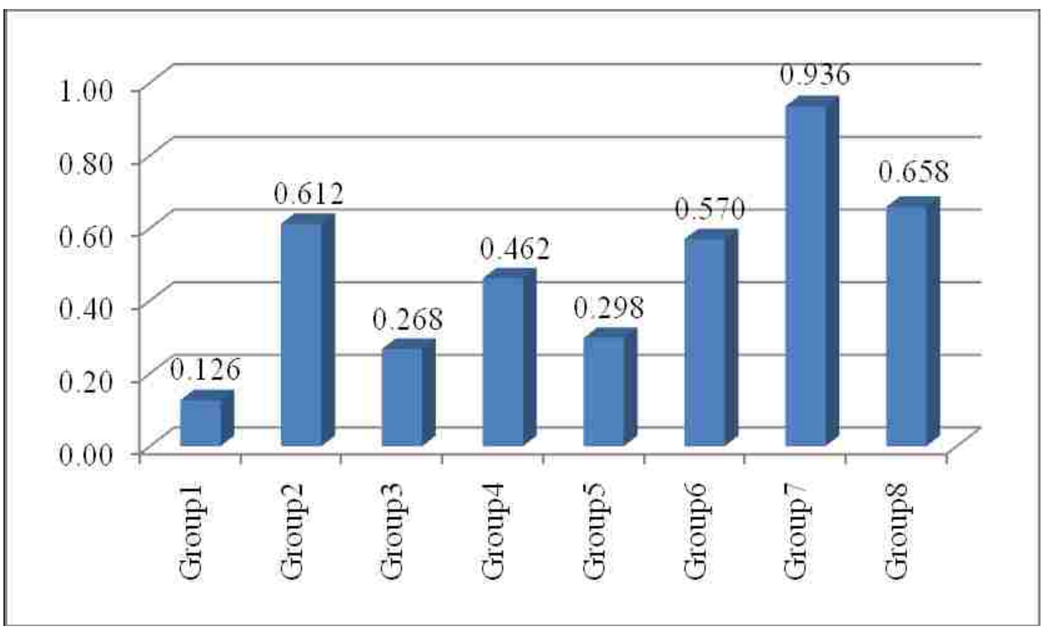


Figure 3.16: Theil Index of Out-Degree of All Eight Groups in the Square Ranked Network

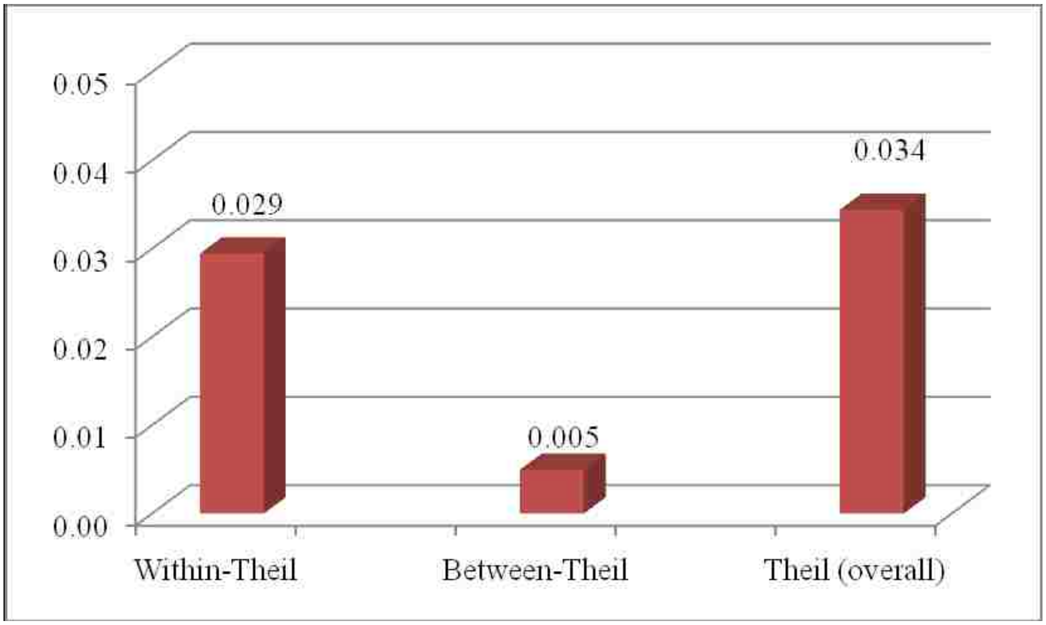


Figure 3.17: Theil Index of In-Degree of the North American Square Ranked Network

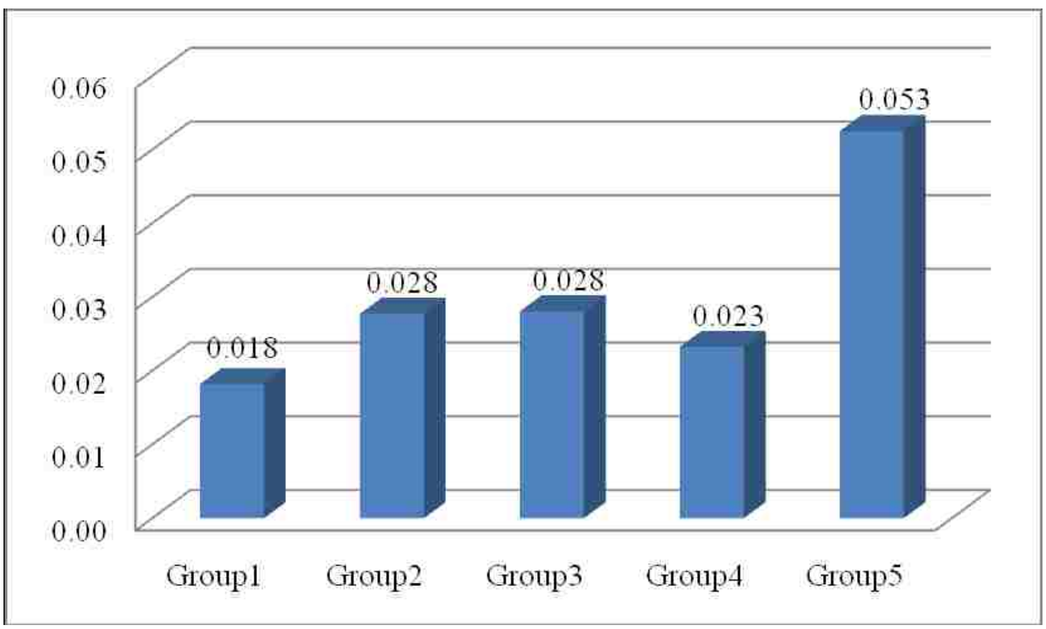


Figure 3.18: Theil Index of In-Degree of All Five Groups in the North American Square Ranked Network

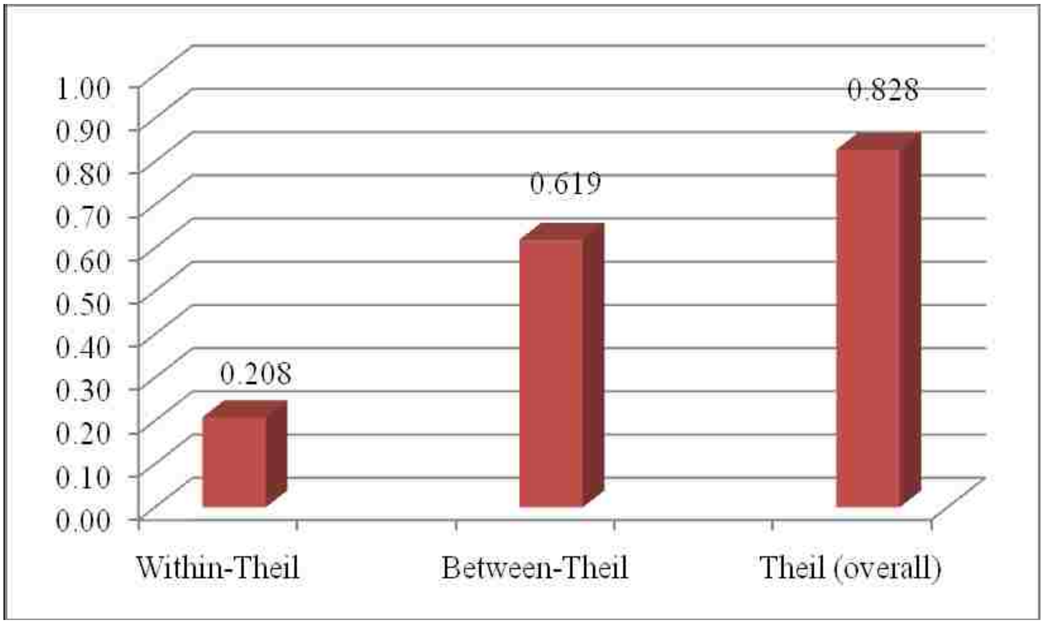


Figure 3.19: Theil Index of Out-Degree in the North American Square Ranked Network

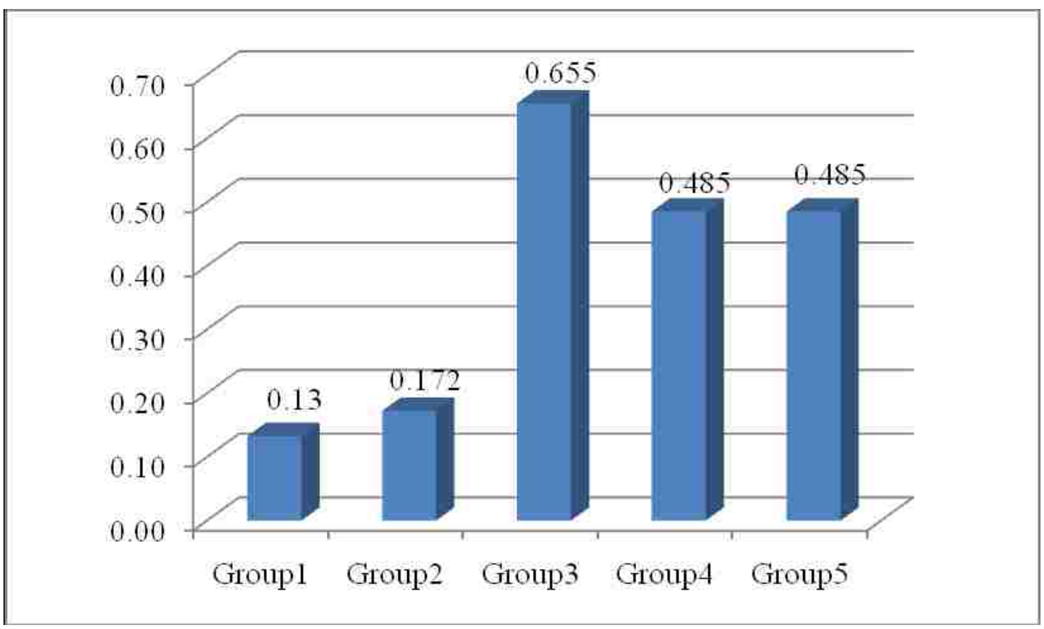


Figure 3.20: Theil Index of Out-Degree of All Five Groups in the North American Square Ranked Network

square ranked network are 0.47 and 0.12, respectively, the density of the entire network is 0.11. The order of density values is clarified in North American Square Ranked Network. They are ordered from the highest to the lowest value with the group quality ordering. The lowest density belongs to Group 5, which is the lowest quality, because most of universities in this group cannot place their Ph.D. within the same group at all. It is very rare ability in this group. Group 1 has the only clustering coefficient that is higher than the network's. These facts indicate closer ties within higher-level groups. Not only is the density of higher-level groups higher, the density of the local neighborhood is also higher. By contrast, the lowest-quality group has a near-zero clustering coefficient value because universities within this group have a very small chance of exchanging their graduates.

Although the relationship within the North American Square Ranked Network is denser than the square ranked network, and has shorter average shortest path than the square ranked network, the diameter is higher (Table 3.3 and 3.4). The diameter in North American Square Ranked Network is eight, while in Square Ranked Network the diameter is seven. There are more channels of links in the square ranked network than North American Square Ranked Network. One example of the maximum shortest path lengths in North America network is from UC Irvine to Yale. However, the route is reduced to three links in the square ranked network. In other words, UC Irvine does not directly grant connect to Yale, but can reach Yale with 8 edges (or links) in the North American Square Ranked Network but with 3 edges in the square ranked network. In the North American Square Ranked Network, any university can reach any other with at most 8 edges. When the non-North American universities are included in the network, they can play a role of a channel of connections in the network. Any university can indirectly connect to any other with shortest path through them. In the square ranked network, any one university can reach any other university with the maximum shortest path length of 7 edges, for example Laval can indirectly reach Harvard or UC Berkeley with 7 edges.

In North American Square Ranked Network, the diameters of the highest quality group are shorter than the diameter of the entire network. The diameters of the lower group are very small because most the universities in the lower group are not likely to employ within the same group. There are many vertices in the network of lower groups that are disconnected and many distances that cannot be measured. The characteristic path length in the lower groups seems short, however, quite meaningless.

Since there is a sense of hierarchy in the placement of Ph.D. graduates, it would pay to investigate the centrality of these networks in terms of node betweenness, which measures the number of shortest paths between other pairs of vertices that pass through any vertex. The higher the node betweenness score of any vertex implies its influential role as being central to the network. Tables 3.6 and 3.7 present node betweenness in the ranked university network and the North American ranked university network respectively. In the former, London School of Economics (LSE) has the highest node betweenness. In the latter, the university with highest node betweenness is Pittsburgh.

In both networks, most universities with a high value of node betweenness are not members of Group 1. Of the top five universities based on the node betweenness score in square ranked network, only Pittsburgh is in North America. As discussed above, the square ranked network which includes the non-North American universities has a shorter diameter than North America ranked university. This indicates the important role of many non-North American universities as being central to the network in terms of node betweenness. Within top ten

Table 3.6: Node Betweeness in the Square Ranked Network

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
1	Harvard	525.89	40
2	Chicago	1337.11	8
3	Penn	1393.70	6
4	Stanford	560.12	36
5	MIT	708.31	23
6	UCBerkeley	671.11	25
7	Northwestern	728.24	22
8	Yale	531.88	38
9	Michigan	361.28	55
10	Columbia	653.75	28
11	Princeton	914.67	17
12	UCLA	700.18	24
13	NYU	521.84	43
14	Cornell	271.78	73
15	LSE	2378.67	1
16	WiscMad	352.17	57
17	Duke	1031.05	12
18	OhioSt	438.58	47
19	Maryland	371.30	53
20	Rochester	665.02	27
21	UTAustin	81.76	115
22	Minnesota	1323.24	9
23	UIUC	895.72	18
24	UCDavis	163.02	92
25	Toronto	256.07	77
26	Oxford	668.62	26
27	UBC	279.50	70
28	UCSD	994.56	13
29	USC	136.17	99
30	BU	64.94	122
31	PennSt	732.45	21
32	CMU	581.62	33
33	Cambridge	228.39	84
34	Florida	93.68	111
35	MichSt	275.53	72

Table 3.6 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
36	Rutgers	72.95	117
37	UWash	307.29	65
38	UNC	252.85	78
39	TAMU	136.47	97
40	Indiana	493.15	45
41	Iowa	570.27	34
42	TelAviv	9.29	155
43	UVA	948.80	15
44	UCL	874.13	19
45	Hebrew	37.44	132
46	Brown	372.87	52
47	Tilburg	260.98	76
48	Pitt	1934.88	3
49	Warwick	597.65	32
50	Arizona	80.99	116
51	WOntario	1363.80	7
52	JHU	71.03	118
53	ANU	2365.52	2
54	Vanderbilt	124.19	100
55	Queen's	519.68	44
56	WUSTL	170.98	89
57	Montreal	224.64	85
58	GTown	374.70	51
59	COBoulder	171.33	88
60	UGA	10.05	153
61	VATech	141.98	94
62	Purdue	1082.08	10
63	UCIrvine	136.38	98
64	BC	166.94	90
65	IowaSt	523.32	42
66	Amsterdam	822.59	20
67	NCSt	567.17	35
68	Erasmus	323.58	63
69	Dartmouth	0.00	167
70	Louvain	86.80	114

Table 3.6 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
71	UYork	614.46	29
72	ASU	60.89	123
73	Toulouse	270.62	74
74	Essex	388.10	50
75	Stockholm	351.78	58
76	UCSB	204.21	86
77	LBS	12.81	152
78	FLSt	278.29	71
79	UNSW	524.24	41
80	Alberta	0.00	167
81	McMaster	161.80	93
82	Houston	29.37	136
83	Syracuse	137.74	96
84	UAB	404.78	48
85	Nottingham	1047.80	11
86	HKUST	0.00	167
87	Bonn	335.35	61
88	YorkU	39.61	130
89	CalTech	548.49	37
90	LSU	249.13	79
91	Southampton	60.28	124
92	UConn	20.32	142
93	GASt	0.00	167
94	UKY	180.56	87
95	GWU	117.14	104
96	INSEE	0.00	167
97	SMU	91.34	112
98	NotreDame	8.89	156
99	SSE	122.34	101
100	SFU	355.43	56
101	Oregon	242.62	81
102	GMU	606.74	31
103	Birkbeck	54.60	127
104	VUA	41.21	129
105	UMass	163.92	91

Table 3.6 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
106	SCarolina	0.00	167
107	ParisI	288.91	66
108	Bristol	4.91	160
109	Melbourne	1462.16	5
110	UIC	0.00	167
111	Copenhagen	280.51	69
112	McGill	268.04	75
113	Groningen	101.49	109
114	ChUHK	0.00	167
115	ULB	0.00	167
116	NewcastleuT	0.78	166
117	Tulane	27.55	137
118	American	66.65	120
119	Mannheim	23.56	140
120	Auburn	282.73	67
121	UPF	929.23	16
122	Buffalo	43.37	128
123	Manchester	1511.29	4
124	UCSC	38.02	131
125	Monash	56.64	125
126	Rice	242.26	82
127	Tennessee	20.25	143
128	Emory	0.00	167
129	NUSingapore	0.00	167
130	Laval	13.14	151
131	C3MU	138.98	95
132	Waterloo	0.00	167
133	WayneSt	2.06	164
134	WiscMil	363.79	54
135	Missouri	102.94	108
136	UCRiverside	531.55	39
137	Alabama	243.52	80
138	Quebec	18.06	144
139	Albany	25.68	138
140	Oslo	0.00	167

Table 3.6 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
141	MiamiFL	0.00	167
142	Maastricht	118.89	103
143	Delaware	0.00	167
144	Sydney	65.07	121
145	EHESS	350.57	59
146	Vienna	96.49	110
147	Munich	87.78	113
148	EAnglia	2.75	163
149	Geneva	14.99	148
150	INSEAD	0.00	167
151	Clemson	0.00	167
152	Birmingham	55.19	126
153	Guelph	15.72	147
154	Hitots	0.00	167
155	Tufts	16.45	146
156	BYU	0.00	167
157	Tokyo	35.43	134
158	CULon	8.09	158
159	Zurich	37.05	133
160	StonyBrook	122.01	102
161	Carleton	13.16	150
162	Reading	240.33	83
163	AcademiaS	0.00	167
164	KUL	114.75	105
165	Bar-Ilan	0.00	167
166	EUI	324.76	62
167	Bocconi	17.32	145
168	Utah	110.26	106
169	Brandeis	3.75	161
170	IUPUI	0.00	167
171	Exeter	24.89	139
172	Bologna	0.00	167
173	Wyoming	280.77	68
174	Nebraska	0.00	167
175	WVA	992.08	14

Table 3.6 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
176	Kansas	69.14	119
177	NHH	0.00	167
178	Temple	13.27	149
179	Glasgow	442.16	46
180	SIUC	338.29	60
181	KanSt	0.00	167
182	CUNY	609.15	30
183	Oklahoma	8.74	157
184	CWM	0.00	167
185	Strathclyde	21.46	141
186	Edinburgh	32.28	135
187	UHK	0.00	167
188	WashSt	398.31	49
189	Uppsala	1.63	165
190	Osaka	9.99	154
191	Tsukuba	5.59	159
192	UNM	0.00	167
193	UCDublin	3.34	162
194	CODenver	0.00	167
195	RomeLS	0.00	167
196	Concordia	317.64	64
197	SCU	0.00	167
198	QMUL	103.92	107
199	MontSt	0.00	167
200	URI	0.00	167

universities in the square ranked network, University of Pennsylvania has the highest node betweenness score and University of Michigan has the smallest node betweenness score.

However, in the square ranked network University of Pennsylvania is ranked as number 8 in terms of node betweenness and Michigan is ranked as number 50. There is no obvious pattern of centrality in terms of node betweenness. The standard correlation between Coupe's (2003) university ranking and the ranking based on node betweenness is about 0.60 in the square ranked network and 0.47 in North American Square Ranked Network shows there is no significant

Table 3.7: Node Betweeness in the North American Square Ranked Network

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
1	Harvard	215.95	30
2	Chicago	874.84	4
3	Penn	830.78	5
4	Stanford	82.09	58
5	MIT	165.72	38
6	UCBerkeley	199.64	33
7	Northwestern	188.53	35
8	Yale	155.31	41
9	Michigan	180.66	36
10	Columbia	58.92	64
11	Princeton	132.93	47
12	UCLA	444.44	11
13	NYU	151.49	44
14	Cornell	59.60	62
16	WiscMad	122.57	50
17	Duke	194.23	34
18	OhioSt	332.00	20
19	Maryland	211.88	32
20	Rochester	77.69	59
21	UTAustin	56.32	66
22	Minnesota	749.38	6
23	UIUC	120.02	51
24	UCDavis	33.93	79
25	Toronto	87.31	57
27	UBC	115.05	52
28	UCSD	56.73	65
29	USC	14.95	90
30	BU	23.79	82
31	PennSt	153.07	42
32	CMU	337.83	19
34	Florida	93.60	55
35	MichSt	163.03	39
36	Rutgers	43.14	71
37	UWash	162.54	40
38	UNC	256.65	25

Table 3.7 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
39	TAMU	60.79	61
40	Indiana	380.94	16
41	Iowa	358.93	17
43	UVA	682.12	7
46	Brown	54.80	67
48	Pitt	1731.92	1
50	Arizona	38.00	73
51	WOntario	427.05	13
52	JHU	26.61	81
54	Vanderbilt	63.64	60
55	Queen's	228.81	28
56	WUSTL	52.83	69
57	Montreal	301.77	22
58	GTown	22.87	84
59	COBoulder	114.43	53
60	UGA	7.30	98
61	VATech	59.51	63
62	Purdue	934.20	3
63	UCIrvine	35.02	78
64	BC	36.35	75
65	IowaSt	492.97	10
67	NCSt	587.14	8
69	Dartmouth	0.00	109
72	ASU	52.07	70
76	UCSB	124.46	49
78	FLSt	441.83	12
80	Alberta	0.00	106
81	McMaster	41.81	72
82	Houston	18.12	89
83	Syracuse	268.80	24
88	YorkU	27.12	80
89	CalTech	89.50	56
90	LSU	350.96	18
92	UConn	18.88	88
93	GASt	0.00	107

Table 3.7 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
94	UKY	147.31	46
95	GWU	0.00	117
97	SMU	103.90	54
98	NotreDame	7.81	95
100	SFU	9.38	94
101	Oregon	224.08	29
102	GMU	394.78	15
105	UMass	35.73	76
106	SCarolina	0.00	110
110	UIC	0.00	103
112	McGill	151.59	43
117	Tulane	23.75	83
118	American	35.51	77
120	Auburn	213.04	31
122	Buffalo	149.09	45
124	UCSC	36.37	74
126	Rice	53.05	68
127	Tennessee	21.05	86
128	Emory	0.00	120
130	Laval	5.08	100
132	Waterloo	0.00	113
133	WayneSt	1.32	101
134	WiscMil	9.84	93
135	Missouri	303.28	21
136	UCRiverside	0.00	105
137	Alabama	246.92	26
138	Quebec	13.00	91
139	Albany	20.08	87
141	MiamiFL	0.00	114
143	Delaware	0.00	115
151	Clemson	0.00	102
153	Guelph	6.86	99
155	Tufts	0.00	118
156	BYU	0.00	104
160	StonyBrook	22.71	85

Table 3.7 Continued

Coupe Rank	Employer	Nodes Betweenness	Node Betweenness Rank
161	Carleton	7.76	97
168	Utah	244.06	27
169	Brandeis	0.00	108
170	IUPUI	0.00	112
173	Wyoming	125.32	48
174	Nebraska	0.00	111
175	WVA	940.46	2
176	Kansas	178.36	37
178	Temple	10.93	92
180	SIUC	408.35	14
181	KanSt	0.00	116
182	CUNY	555.50	9
183	Oklahoma	7.76	96
184	CWM	0.00	119
188	WashSt	300.86	23
192	UNM	0.00	121
194	CODenver	0.00	122
196	Concordia	0.00	123
197	SCU	0.00	124
199	MontSt	0.00	125
200	URI	0.00	126

relationship between the centrality role and quality of university in the network. Hence, even though universities in Group 1 are privileged in terms of placement, they are not influential or central in terms of node betweenness. Figures 3.21 and Figure 3.22 present node betweenness score distribution plotted on a log-log scale for the square ranked network and the North American square ranked network. The node betweenness distribution in the latter follows the power law approximately, more so than in the former. The square ranked network seems to have more nodes of centrality than the North American square ranked network.

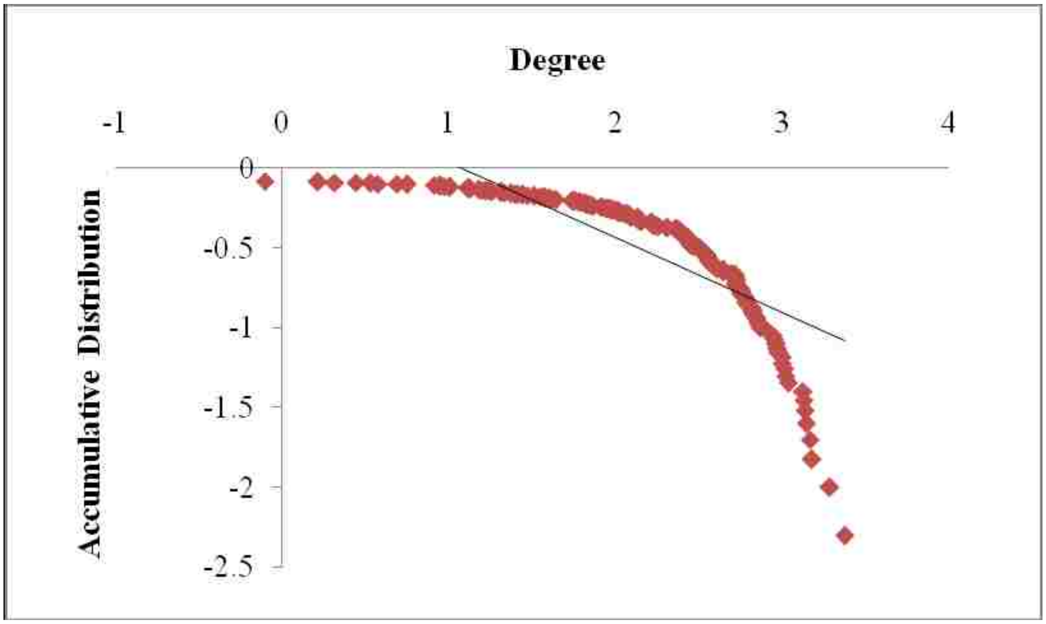


Figure 3.21: Log-Log Plot (P(X>a)) of Node Betweeness Distribution in the Square Ranked Network

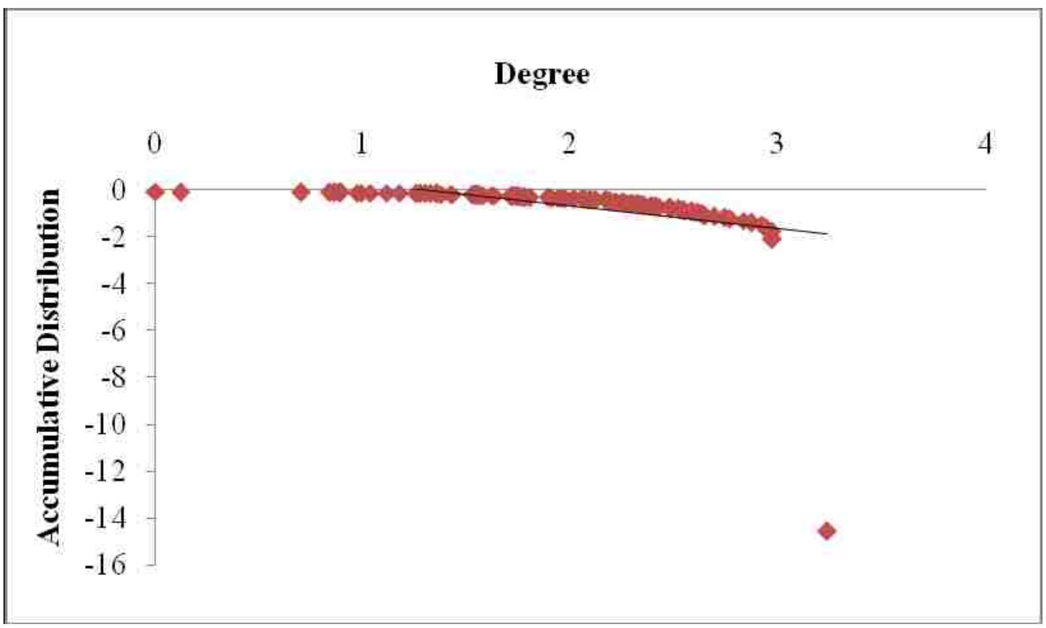


Figure 3.22: Log-Log Plot (P(X>a)) of Node Betweeness Distribution in the North American Square Ranked Network

3.5 Summary

Social relationships between universities in the network has formed the sustained pattern of Ph.D. granting and employment in the economics academic labor market. Every quality group of universities in the economics academic market can connect to every other. Interactions in the network exhibit its hierarchical structure or so-called "downstream pattern" described by Moore and Newman (1977). Higher ranked groups supply their graduates to lower ranked groups. In other words, every group in the labor market for academic economists tends to hire graduates from groups ranked above it. This chapter further investigates interactions between each university in the network instead of just groups as in chapter 2. This chapter is interested in the structural characteristics of the economic academic labor market particularly whether it can be characterized as a small-world network or not.

Empirical works on the small-world phenomenon are few. The small-world phenomenon properties was examined in the economics collaboration network by Goyal et al. (2004) and Goyal et al. (2005) finding evidence of an emerging of small-world phenomenon in the network of collaborating economists. Both the Lorenz curve construction and the degree distribution plot indicate the inequality of the degree of collaboration in the network and the network structure consists of a few stars as the centers of the network.

This chapter found that the North American Square Ranked Network and the square ranked network are single component networks. The North American Square Ranked Network resembles the overall square ranked network. The economics academic networks exhibit the small-world phenomenon, with the characteristic path lengths close to those of random networks. While the clustering coefficient of North American Square Ranked Network is more than twice as high as in a comparable random network, the clustering coefficient of square ranked network is more than four times as high. Any two universities can connect through only three links. Moreover, the maximum shortest path for any university to reach any other university is seven links.

Most of universities employ their faculty from higher ranked universities. The connections of ranked universities to others in terms of “employment” do not display any significant inequality. But, there is significant inequality in the degree of placement of economists in the academic labor market. It is due to the “downstream pattern” in academic labor market. Top ranked universities can place their graduates in lower ranked universities, but very few lower ranked universities can place their graduates in higher ranked universities. Lower ranked universities, then, connect to their same quality grouped universities in terms of “granting.” Moreover, in observed universities, many universities do not produce many graduates and some universities cannot place their graduates in ranked universities at all. This inequality of granting connections, then, significantly comes from the inequality between the different groups of quality, not the inequality within the group.

Even though the inequality of granting connections seems significant, there is no inequality of employment connections. The structure of economics academic network is such that in terms of node betweenness there are many universities, particularly non-North American universities, that play a role as the centrals of these single component networks, for example London School of Economics, Australian National University, Pittsburgh, Manchester, etc. These universities significantly reduce the paths between any two universities. It is hard to relate the influential universities in terms of node betweenness and the influential university in terms of placement in the academic network. Although most of these central universities are not influential in terms of placement their graduates in the network, were the removed from the network, any two ranked universities would connected through more than three links.

REFERENCES

Adamic, Lada A., and Eytan Adar. (2004): "How to Search a Social Network," 10 Sept. 2006 <http://arxiv.org/PS_cache/cond-mat/pdf/0310/0310120v2.pdf>.

Amaral, L. A. N., A. Scala, M. Barthelemy, and H. E. Stanley. (2000): "Classes of Small-World Networks," *PNAS*, 97(21), pp. 11149-11152. 4 Oct. 2006 <http://www.pnas.org>.

Barabasi, A.L., and Eric Bonabeau. (2003): "Scale-Free Networks," *Scientific American,* 86(25), pp. 50-59.

Barrat, A., and M. Weigt. (2000): "On the Properties of Small-World Network Models," *European Physical Journal B* 13, pp. 547-560.

Bonacich, Phillip. "The Strength of Weak Ties," 15 March 2006 <http://www.sscnet.ucla.edu>

Borgatti, S.P., M. G. Everett, and Freeman. L.C. (2002): Ucinet6 for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Collier, Irwin. (1999): "Lecture on Theil Index of Inequality," 20 Feb. 2005 <http://www.wiwiss.fu-berlin.de/w3/w3collie/>.

Conceicao, Pedro, James K Galbraith, and Peter Bradford. " The Theil Index Sequences of Nested and Hierarchic Grouping Structure: Implication for the Measurement of Inequality through Time with Data Aggregated at Different Levels if Industrial Classification," UTIP Working Paper, Number 15, U of Texas, Austin. 1-43. 20 Feb. 2005 <http://www.utip.gov.utexas.edu/web/workingpaper/utip15.pdf>.

Coupe, Tom. (2003): "Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1999-2000," *Journal of the European Economic Association,* 1(6), pp. 1309-1345.

Dusansky, Richard, and Clayton J. Vernon. (1998): "Rankings of U.S. Economics Departments," *Journal of Economics Perspectives,* 12(1), pp. 157-170. 10 Jan. 2007 <http://www.jstor.org>.

"Gini Coefficient," < http://www.techwranglers.net/dtree/gini.html>.

Goffman, Casper. (1969): "And What Is Your Erdos Number," *The American Mathematical Monthly,* 76(7), p. 791.

Goyal, Sanjeev, Macro van der Leij, and Jose Luis Moraga-Gonzalez. (2004): "Economics: An Emerging Small World," Tinbergen Institute 14 Sept.2006 <http://www.tinbergen.nl/~noraga/smallworld.pdf>.

Granovetter, Mark S. (1973): “The Strength of Weak Ties,” *American Journal of Sociology*, 78(6), pp. 1360-1380.

Grossman, Jerrold W. (1996): “Paul Erdos: The Master of Collaboration,” Oakland University, Rochester. 20 Nov. 2006 <http://citeseer.ist.psu.edu>.

Hanneman, Robert A., and Mark Riddle. (2005): “Introduction to Social Network Methods,” U of California, Riverside. 16 Jan. 2007 <http:// www.faculty.ucr.edu/~hanneman/>.

Heck, Jean L., Peter A. Zaleski, and Scott J. Dressler. (2006): “Economic Departments and Their Contributions to the Elite Economic Journals,” *Journal of the Academy of Business Education*, 7. 20 Mar. 2007 <www.abe.sju.edu/proc2006/heck.pdf>.

Jackson, Matthew O., and Brian W. Rogers. (2005): “The Economics of Small Worlds,” *Journal of the European Economic Association,* 3(2-3), pp. 617-627.

Kalaitzidakis, Pantelis, Thanasis Stengos, and Theofanis P. Mamuneas. (2003): “Rankings of Academic Journals and Institution in Economics,” *Journal of the European Economic Association,* 1(6), pp. 1346-1366.

Martin, Ryan. (2005): “Six Degree of Graph Theory: Kevin Bacon, Paul Erdos, William McKinley and Me,” Central College Mathematics Seminar. 15 Nov. 2006 <http://orion.math.iastate.edu/rymartin/talks/SixDegrees/6degMAA.pdf>.

Montgomery, J. D. (1991): “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis,” *American Economic Review,* 81(5), pp. 1408-1418.

Moore, William J., and Robert J. Newman. (1977): “Academic Placement of Econommists 1960-74,” *Job Openings for Economists,* American Economic Association.

Newman, M. E. J. (2000): “Models of the Small World: A Review,” *Santa Fe Institute Bulletin* 2 Oct. 2006 <http://aps.arxiv.org/PS_cache/cond-mat/pdf/0001/0001118v2.pdf>.

Newman, M. E. J. (2001): “The Structure of Scientific Collaboration Networks,” *PNAS,* 98(2), pp. 404-409. 5 Oct. 2006 <http://www.pnas.org>.

Newman, M. E. J. (2002): “Random Graphs as Models of Networks,” *Santa Fe Institute Bulletin* 2 Oct. 2006 <http://arxiv.org/PS_cache/cond-mat/pdf/0202/0202208v1.pdf>.

Newman, M. E. J. (2004): “Coauthorship Networks and Patterns of Scientific Collaboration,” *PNAS,* 101, pp. 5200-5205. 2 Oct. 2006 <http://www.pnas.org>.

Newman, M. E. J. (2005): “Power Laws, Pareto Distributions and Zipf’s Law,” *Contemporary Physics,* 46(5), pp. 323-351.

Roessler, Christian. (2004): “Rankings,” 20 Feb. 2007 <http://www.econphd.net/rankings.htm>.

Scott, John. (1991): *Social Network Analysis: A Handbook*, London: Sage.

Scott, Loren C., and Peter M. Mitias. (1996): "Trends in Rankings of Economics Departments in the U.S.: An Update," *Economic Inquiry*, 34, pp. 378-400.

Simao, Andrea Branco, Claudia Jullia Horta, and Simone Wajnman. (2001): "The Effect of Labor Force Aging and Female Participation Growth on the Income Inequality," 20 Feb. 2005 <http://www.iussp.org/Brazil2001/s30/S35_P05_Horta.pdf>.

Tervio, Marko. (2006): "Network Analysis of Three Academic Labor Markets," 11 Apr. 2007 <http://faculty.hass.berkeley.edu/marko/>.

Travers, Jeffrey, and Stanley Milgram. (1969): "An Experimental Study of the Small World Problem," *Sociometry* 32(4), pp. 425-443.

Van der Leij, Macro J., and Sanjeev Goyal. (2005): "Strong Ties in a Small World," Tinbergen Institute 17 Sept. 2006 <http://www.tinbergen.nl/discussionpaper/06008.pdf>.

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

Watts, Duncan J., and Steven H. Strogatz. (1998): "Collective Dynamics of Small-World Network," *Nature*, 393, pp. 440-442.

Watts, Duncan J. (1999a): *Small World: The Dynamics of Networks between Order and Randomness*, Princeton: Princeton University Press.

Watts, Duncan J. (1999b): "Networks, Dynamics, and the Small-World Phenomenon," *The American Journal of Sociology*, 105(2), pp. 493-527.

Weisstein, Eric W. *MathWorld* <http://mathworld.wolfram.com>

White, Douglas R., and Michael Houseman. (2002): "The Navigability of Strong Ties: Small Worlds, Tie Strength and Network Topology," Santa Fe Institute 2 Oct. 2006 <http://www.santafe.edu/research/publications/workingpapers/02-10-055.pdf>.

Zlatic, V., M. Bozicevic, H. Stefancic, and M. Domazet. (2006): "Wikipedias: Collaborative Web-Based Encyclopedias as Complex Networks," *Physical Review E*, 74(1), 10 Jan. 2007 <http://arxiv.org/PS_cache/physics/pdf/0602/0602149v3.pdf>.

APPENDIX A: UNIVERSITY NAMES AND DETAIL

Coupe Rank	Abbreviation	University Name	Country	Continent
1	Harvard	U Harvard	USA	North America
2	Chicago	U Chicago	USA	North America
3	Penn	U Penn (UPA)	USA	North America
4	Stanford	U Stanford	USA	North America
5	MIT	MIT	USA	North America
6	UCBerkeley	UC Berkeley (U CA Berkeley)	USA	North America
7	Northwestern	Northwestern U	USA	North America
8	Yale	U Yale	USA	North America
9	Michigan	U MI Ann Arbor (Michigan)	USA	North America
10	Columbia	Columbia U	USA	North America
11	Princeton	Princeton U	USA	North America
12	UCLA	UCLA	USA	North America
13	NYU	NYU	USA	North America
14	Cornell	Cornell U	USA	North America
15	LSE	London School of Econ	UK	Europe
16	WiscMad	Wisconsin-Madison (U WI Madison)	USA	North America
17	Duke	Duke U	USA	North America
18	OhioSt	Ohio-State (oh State U)	USA	North America
19	Maryland	U MD College Park	USA	North America
20	Rochester	U Rochester	USA	North America
21	UTAustin	U TX Austin	USA	North America
22	Minnesota	U MN Twin Cities	USA	North America
23	UIUC	U IL Urbana Champaign	USA	North America
24	UCDavis	U CA Davis	USA	North America
25	Toronto	U Toronto	Canada	North America
26	Oxford	U Oxford	UK	Europe
27	UBC	U British Columbia	Canada	North America
28	UCSD	U CA San Diego	USA	North America
29	USC	U Southern CA	USA	North America
30	BU	Boston U	USA	North America
31	PennSt	Penn State U (PA State U)	USA	North America
32	CMU	Carnegie Mellon U	USA	North America
33	Cambridge	U Cambridge	UK	Europe
34	Florida	U Florida (U FL)	USA	North America

Coupe Rank	Abbreviation	University Name	Country	Continent
35	MichSt	Michigan State (MI State U)	USA	North America
36	Rutgers	Rutgers U NJ	USA	North America
37	UWash	U Washington (UWA)	USA	North America
38	UNC	U NC Chapel Hill	USA	North America
39	TAMU	TX A&M U	USA	North America
40	Indiana	IN U, Bloomington (Indiana)	USA	North America
41	Iowa	U Iowa (U IA)	USA	North America
42	TelAviv	U Tel Aviv	Israel	Asia
43	UVA	U Virginia (UVA)	USA	North America
44	UCL	U College London (UCL)	UK	Europe
45	Hebrew	Hebrew U	Israel	Asia
46	Brown	Brown U	USA	North America
47	Tilburg	U Tilburg	Netherlands	Europe
48	Pitt	U Pittsburgh	USA	North America
49	Warwick	U Warwick	UK	Europe
50	Arizona	U Arizona (U AZ)	USA	North America
51	WOntario	U Western Ontario	Canada	North America
52	JHU	Johns Hopkins U	USA	North America
53	ANU	Australian National U (ANU)	Australia	Australia
54	Vanderbilt	Vanderbilt U	USA	North America
55	Queen's	Queens U, Canada	Canada	North America
56	WUSTL	Washington U, MO (WUSTL)	USA	North America
57	Montreal	U Montreal	Canada	North America
58	GTown	Georgetown U, DC	USA	North America
59	COBoulder	U Colorado Boulder (U CO Boulder)	USA	North America
60	UGA	U Georgia (UGA)	USA	North America
61	VATech	VA Polytechnic Institute& State U	USA	North America
62	Purdue	Purdue U in	USA	North America
63	UCIrvine	U CA Irvine	USA	North America
64	BC	Boston College	USA	North America
65	IowaSt	Iowa State (IA State U)	USA	North America
66	Amsterdam	U Amsterdam	Netherlands	Europe
67	NCSt	NC State U	USA	North America
68	Erasmus	Erasmus U Rotterdam	Netherlands	Europe
69	Dartmouth	Dartmouth College	USA	North America
70	Louvain	Catholic U Louvain	Belgium	Europe

Coupe Rank	Abbreviation	University Name	Country	Continent
71	UYork	U York, UK	UK	Europe
72	ASU	AZ State U (ASU)	USA	North America
73	Toulouse	U Toulouse I	France	Europe
74	Essex	U Essex	UK	Europe
75	Stockholm	U Stockholm	Sweden	Europe
76	UCSB	U CA Santa Barbara	USA	North America
77	LBS	London Business School	UK	Europe
78	FLSt	Florida State (FL State U)	USA	North America
79	UNSW	U New S Wales	Australia	Australia
80	Alberta	U Alberta	Canada	North America
81	McMaster	McMaster U	Canada	North America
82	Houston	U Houston	USA	North America
83	Syracuse	Syracuse U, NY	USA	North America
84	UAB	U Autonoma Barcelona	Spain	Europe
85	Nottingham	U Nottingham	UK	Europe
86	HKUST	Hongkong U of Science & Tech	China	Asia
87	Bonn	U Bonn	Germany	Europe
88	YorkU	York U Canada	Canada	North America
89	CalTech	CA Institute of Technology	USA	North America
90	LSU	LA State U	USA	North America
91	Southampton	U Southampton	UK	Europe
92	UConn	U Connecticut (U CT)	USA	North America
93	GASt	GA State U	USA	North America
94	UKY	U Kentucky (U KY)	USA	North America
95	GWU	George Washington U, DC	USA	North America
96	INSEE	INSEE	France	Europe
97	SMU	Southern Mathodist U	USA	North America
98	NotreDame	U Notre Dame IN	USA	North America
99	SSE	Stockholm School of Econ	Sweden	Europe
100	SFU	Simon Fraser U CN	Canada	North America
101	Oregon	U Oregon (U OR)	USA	North America
102	GMU	George Mason U, VA	USA	North America
103	Birkbeck	Birkbeck College, U London	UK	Europe
104	VUA	Free U Amsterdam (Vrije U Amsterdam)	Netherlands	Europe
105	UMass	U MA Amherst	USA	North America

Coupe Rank	Abbreviation	University Name	Country	Continent
106	SCarolina	U South Carolina (USC)	USA	North America
107	ParisI	U Paris I	France	Europe
108	Bristol	U Bristol	UK	Europe
109	Melbourne	U Melbourne	Australia	Australia
110	UIC	U IL Chicago	USA	North America
111	Copenhagen	U Copenhagen	Denmark	Europe
112	McGill	McGill U	Canada	North America
113	Groningen	U Groningen	Netherlands	Europe
114	ChUHK	Chinese U Hong Kong	China	Asia
115	ULB	Free U Brussels (ULB)	Belgium	Europe
116	NewcastleuT	U Newcastle upon Tyne	UK	Europe
117	Tulane	Tulane U	USA	North America
118	American	American U, Washington, DC	USA	North America
119	Mannheim	U Mannheim	Germany	Europe
120	Auburn	Auburn U	USA	North America
121	UPF	U Pompeu Fabra (UPF)	Spain	Europe
122	Buffalo	SUNY Buffalo	USA	North America
123	Manchester	U Manchester	UK	Europe
124	UCSC	U CA Santa Cruz	USA	North America
125	Monash	Monash U, Australia	Australia	Australia
126	Rice	Rice U, Houston, TX	USA	North America
127	Tennessee	U TX Knoxville	USA	North America
128	Emory	Emory U	USA	North America
129	NUSingapore	U National Singapore	Singapore	Asia
130	Laval	U Laval	Canada	North America
131	C3MU	U Carlos III Madrid	Spain	Europe
132	Waterloo	U Waterloo, Waterloo, Ontario	Canada	North America
133	WayneSt	Wayne State U, MI	USA	North America
134	WiscMil	U WI Milwaukee	USA	North America
135	Missouri	Missouri (U MO Columbia)	USA	North America
136	UCRiverside	U CA Riverside	USA	North America
137	Alabama	U Alabama (U AL)	USA	North America
138	Quebec	U Quebec Montreal	Canada	North America
139	Albany	SUNY Albany	USA	North America
140	Oslo	U Oslo	Norway	Europe

Coupe Rank	Abbreviation	University Name	Country	Continent
141	MiamiFL	U Miami, FL	USA	North America
142	Maastricht	U Maastricht	Netherlands	Europe
143	Delaware	Delaware (U DE)	USA	North America
144	Sydney	U Sydney	Australia	Australia
145	EHESS	EHESS	France	Europe
146	Vienna	U Vienna	Austria	Europe
147	Munich	U Munchen (Munich)	Germany	Europe
148	EAnglia	East Anglia	UK	Europe
149	Geneva	U Geneva	Switzerland	Europe
150	INSEAD	INSEAD	France	Europe
151	Clemson	Clemson U	USA	North America
152	Birmingham	U Birmingham	UK	Europe
153	Guelph	U Guelph	Canada	North America
154	Hitots	Hitotsubashi U	Japan	Asia
155	Tufts	Tufts U	USA	North America
156	BYU	Brigham Young U	USA	North America
157	Tokyo	U Tokyo	Japan	Asia
158	CULon	City U London	UK	Europe
159	Zurich	U Zurich	Switzerland	Europe
160	StonyBrook	SUNY Stony Brook	USA	North America
161	Carleton	Carleton U, Ottawa	Canada	North America
162	Reading	U Reading	UK	Europe
163	AcademiaS	Academia Sinica	China	Asia
164	KUL	Catholic U Leuven	Belgium	Europe
165	Bar-Ilan	Bar Ilan U	Israel	Asia
166	EUI	European U Institute, Firenze	Italy	Europe
167	Bocconi	U Bocconi	Italy	Europe
168	Utah	U Utah (U UT)	USA	North America
169	Brandeis	Brandeis U	USA	North America
170	IUPUI	IN U Purdue U, Indianapolis	USA	North America
171	Exeter	U Exeter	UK	Europe
172	Bologna	U Bologna	Italy	Europe
173	Wyoming	U Wyoming (U WY0	USA	North America
174	Nebraska	U Nebraska Lincoln (U NE Lincoln)	USA	North America
175	WVA	U West Virginia (WV U)	USA	North America

Coupe Rank	Abbreviation	University Name	Country	Continent
176	Kansas	U Kansas (U KS)	USA	North America
177	NHH	Norwegian School Econ & Business Admin. (NHH)	Norway	Europe
178	Temple	Temple U	USA	North America
179	Glasgow	U Glasgow	UK	Europe
180	SIUC	Southern IL U Carbondale (SIUC)	USA	North America
181	KanSt	Kansas State U (KS State U)	USA	North America
182	CUNY	CUNY Baruch College	USA	North America
183	Oklahoma	U Oklahoma (U OK)	USA	North America
184	CWM	College of William & Mary (CWM)	USA	North America
185	Strathclyde	U Strathclyde	UK	Europe
186	Edinburgh	U Edinburgh	UK	Europe
187	UHK	U Hong Kong (UHK)	China	Asia
188	WashSt	Washington State U	USA	North America
189	Uppsala	Uppsala U, Sweden	Sweden	Europe
190	Osaka	Osaka U	Japan	Asia
191	Tsukuba	U Tsukuba, Japan	Japan	Asia
192	UNM	U New Mexico (U NM)	USA	North America
193	UCDublin	U College Dublin	Ireland	Europe
194	CODenver	UCO Denver	USA	North America
195	RomeLS	U Rome La Sapienza	Italy	Europe
196	Concordia	Concordia U	Canada	North America
197	SCU	Santa Clara U, CA	USA	North America
198	QMUL	Queen Mary & Westfield College	UK	Europe
199	MontSt	Montana State U (MT State U)	USA	North America
200	URI	Rhode Island	USA	North America

APPENDIX B: DEGREE IN SQUARE RANKED NETWORK

Ranked by Coupe	University Name	In-Degree	Out-Degree
1	Harvard	15	90
2	Chicago	12	91
3	Penn	17	73
4	Stanford	12	89
5	MIT	12	97
6	UCBerkeley	13	94
7	Northwestern	18	75
8	Yale	14	90
9	Michigan	19	47
10	Columbia	17	45
11	Princeton	18	87
12	UCLA	20	44
13	NYU	20	32
14	Cornell	15	53
15	LSE	17	65
16	WiscMad	15	69
17	Duke	15	39
18	OhioSt	15	19
19	Maryland	17	22
20	Rochester	11	59
21	UTAustin	17	14
22	Minnesota	14	70
23	UIUC	21	35
24	UCDavis	15	15
25	Toronto	20	19
26	Oxford	12	49
27	UBC	16	27
28	UCSD	14	39
29	USC	17	12
30	BU	13	21
31	PennSt	17	19
32	CMU	22	26
33	Cambridge	6	40
34	Florida	13	6
35	MichSt	18	25

Ranked by Coupe	University Name	In-Degree	Out-Degree
36	Rutgers	21	8
37	UWash	15	28
38	UNC	17	18
39	TAMU	18	11
40	Indiana	19	19
41	Iowa	10	18
42	TelAviv	11	3
43	UVA	17	24
44	UCL	15	18
45	Hebrew	11	9
46	Brown	12	48
47	Tilburg	10	11
48	Pitt	21	21
49	Warwick	17	14
50	Arizona	14	6
51	WOntario	16	25
52	JHU	10	34
53	ANU	12	13
54	Vanderbilt	19	8
55	Queen's	16	25
56	WUSTL	14	21
57	Montreal	21	7
58	GTown	21	4
59	COBoulder	23	10
60	UGA	12	1
61	VATech	11	15
62	Purdue	15	35
63	UCIrvine	18	4
64	BC	15	13
65	IowaSt	19	7
66	Amsterdam	13	6
67	NCSt	18	8
68	Erasmus	9	7
69	Dartmouth	9	0
70	Louvain	3	14

Ranked by Coupe	University Name	In-Degree	Out-Degree
71	UYork	17	12
72	ASU	22	3
73	Toulouse	10	18
74	Essex	18	12
75	Stockholm	7	9
76	UCSB	17	11
77	LBS	3	4
78	FLSt	22	4
79	UNSW	17	2
80	Alberta	14	0
81	McMaster	17	9
82	Houston	17	2
83	Syracuse	14	7
84	UAB	16	9
85	Nottingham	21	5
86	HKUST	13	0
87	Bonn	5	10
88	YorkU	20	2
89	CalTech	13	19
90	LSU	11	2
91	Southampton	10	4
92	UConn	21	1
93	GASt	20	0
94	UKY	14	3
95	GWU	20	1
96	INSEE	3	0
97	SMU	15	3
98	NotreDame	18	1
99	SSE	7	7
100	SFU	20	6
101	Oregon	17	5
102	GMU	17	5
103	Birkbeck	8	6
104	VUA	4	6
105	UMass	13	5

Ranked by Coupe	University Name	In-Degree	Out-Degree
106	SCarolina	14	0
107	ParisI	1	10
108	Bristol	2	4
109	Melbourne	25	3
110	UIC	13	0
111	Copenhagen	9	5
112	McGill	24	4
113	Groningen	3	4
114	ChUHK	15	0
115	ULB	0	7
116	NewcastleuT	2	1
117	Tulane	10	4
118	American	13	3
119	Mannheim	8	2
120	Auburn	7	1
121	UPF	24	11
122	Buffalo	10	5
123	Manchester	18	13
124	UCSC	13	3
125	Monash	13	2
126	Rice	16	9
127	Tennessee	14	1
128	Emory	13	0
129	NUSingapore	27	0
130	Laval	16	2
131	C3MU	19	4
132	Waterloo	10	0
133	WayneSt	9	1
134	WiscMil	16	3
135	Missouri	13	3
136	UCRiverside	12	4
137	Alabama	11	1
138	Quebec	18	2
139	Albany	13	2
140	Oslo	3	0

Ranked by Coupe	University Name	In-Degree	Out-Degree
141	MiamiFL	12	0
142	Maastricht	7	5
143	Delaware	20	0
144	Sydney	20	3
145	EHESS	9	12
146	Vienna	4	5
147	Munich	9	4
148	EAnglia	4	2
149	Geneva	7	2
150	INSEAD	7	0
151	Clemson	14	0
152	Birmingham	13	2
153	Guelph	15	1
154	Hitots	6	0
155	Tufts	14	1
156	BYU	14	0
157	Tokyo	14	2
158	CULon	5	1
159	Zurich	5	2
160	StonyBrook	10	12
161	Carleton	14	2
162	Reading	10	1
163	AcademiaS	17	0
164	KUL	8	8
165	Bar-Ilan	7	0
166	EUI	9	18
167	Bocconi	16	2
168	Utah	9	2
169	Brandeis	14	1
170	IUPUI	16	0
171	Exeter	10	2
172	Bologna	20	0
173	Wyoming	7	1
174	Nebraska	12	0
175	WVA	15	3

Ranked by Coupe	University Name	In-Degree	Out-Degree
176	Kansas	15	3
177	NHH	2	0
178	Temple	17	1
179	Glasgow	11	2
180	SIUC	8	5
181	KanSt	12	0
182	CUNY	24	5
183	Oklahoma	9	1
184	CWM	15	0
185	Strathclyde	10	2
186	Edinburgh	9	3
187	UHK	16	0
188	WashSt	11	3
189	Uppsala	2	3
190	Osaka	7	1
191	Tsukuba	6	1
192	UNM	9	0
193	UCDublin	3	1
194	CODenver	8	0
195	RomeLS	3	0
196	Concordia	15	2
197	SCU	7	0
198	QMUL	15	2
199	MontSt	8	0
200	URI	7	0

Note: In-Degree and Out-Degree are measured in unweighted directed and un-looped networks.

APPENDIX C: DEGREE IN NORTH AMERICAN SQUARE RANKED NETWORK

Ranked by Coupe	University Name	In-Degree	Out-Degree
1	Harvard	13	69
2	Chicago	12	77
3	Penn	16	56
4	Stanford	10	70
5	MIT	10	73
6	UCBerkeley	10	74
7	Northwestern	16	60
8	Yale	13	70
9	Michigan	17	40
10	Columbia	12	34
11	Princeton	14	68
12	UCLA	18	38
13	NYU	16	20
14	Cornell	13	38
16	WiscMad	13	64
17	Duke	13	36
18	OhioSt	13	18
19	Maryland	17	19
20	Rochester	9	48
21	UTAustin	17	13
22	Minnesota	13	57
23	UIUC	18	31
24	UCDavis	12	11
25	Toronto	17	14
27	UBC	15	20
28	UCSD	12	26
29	USC	13	7
30	BU	12	14
31	PennSt	13	12
32	CMU	22	22
34	Florida	12	6
35	MichSt	17	23
36	Rutgers	20	7
37	UWash	14	25
38	UNC	16	18

Ranked by Coupe	University Name	In-Degree	Out-Degree
39	TAMU	17	10
40	Indiana	18	18
41	Iowa	10	15
43	UVA	17	23
46	Brown	10	45
48	Pitt	18	19
50	Arizona	12	5
51	WOntario	13	20
52	JHU	10	28
54	Vanderbilt	17	6
55	Queen's	14	20
56	WUSTL	12	18
57	Montreal	17	7
58	GTown	18	1
59	COBoulder	23	10
60	UGA	12	1
61	VATech	9	12
62	Purdue	15	32
63	UCIrvine	17	2
64	BC	12	8
65	IowaSt	17	7
67	NCSt	17	8
69	Dartmouth	9	0
72	ASU	21	3
76	UCSB	16	10
78	FLSt	22	4
80	Alberta	12	0
81	McMaster	14	7
82	Houston	15	2
83	Syracuse	14	6
88	YorkU	17	2
89	CalTech	12	18
90	LSU	11	2
92	UConn	21	1
93	GASt	20	0

Ranked by Coupe	University Name	In-Degree	Out-Degree
94	UKY	14	2
95	GWU	20	0
97	SMU	13	3
98	NotreDame	17	1
100	SFU	18	1
101	Oregon	14	5
102	GMU	16	4
105	UMass	12	5
106	SCarolina	13	0
110	UIC	13	0
112	McGill	19	3
117	Tulane	10	4
118	American	12	3
120	Auburn	7	1
122	Buffalo	10	5
124	UCSC	13	3
126	Rice	14	6
127	Tennessee	14	1
128	Emory	13	0
130	Laval	13	2
132	Waterloo	10	0
133	WayneSt	9	1
134	WiscMil	16	2
135	Missouri	13	3
136	UCRiverside	12	0
137	Alabama	11	1
138	Quebec	15	2
139	Albany	13	1
141	MiamiFL	12	0
143	Delaware	20	0
151	Clemson	14	0
153	Guelph	12	1
155	Tufts	13	0
156	BYU	14	0
160	StonyBrook	8	8

Ranked by Coupe	University Name	In-Degree	Out-Degree
161	Carleton	12	2
168	Utah	9	1
169	Brandeis	14	0
170	IUPUI	15	0
173	Wyoming	6	1
174	Nebraska	12	0
175	WVA	15	3
176	Kansas	15	3
178	Temple	17	1
180	SIUC	8	4
181	KanSt	12	0
182	CUNY	23	4
183	Oklahoma	9	1
184	CWM	15	0
188	WashSt	11	3
192	UNM	9	0
194	CODenver	8	0
196	Concordia	14	0
197	SCU	6	0
199	MontSt	8	0
200	URI	7	0

Note: In-Degree and Out-Degree are measured in unweighted directed and un-looped networks.

VITA

Nongnuch Soonthornchawakan was born in Phuket, Thailand. Her parents passed away when she was young. She and her sisters were raised by her aunts. She attended high school in Bangkok and graduated from the Wattana Wittaya Academy in 1983. She received a Bachelor of Arts degree in economics with first class honors from Thammasat University in 1987. In 1990 she received a Master of Arts degree in economics from the same university. She received a Japan Ministry of Education Scholarship (Monbusho) to study in Japan. She became a doctoral candidate in 1992 at the Faculty of Economics, Kobe University, Japan. After a very scary experience of the Kobe earthquake in 1995 she decided to return to Thailand. She became a lecturer and researcher at the Faculty of Economics, Thammasat University. She received fellowships from Japan Society for the Promotion of Science at Kyoto University in 1997 and 2000. In addition to teaching and research duties, she also held administrative positions. She served as Vice Dean for Student Affairs, member of the Board of Faculty, and Representative of the Faculty. She was promoted to an Assistant Professor in 2001. In 2002, she received a scholarship from the Faculty of Economics and Thammasat University to study for the Doctor of Philosophy degree in the United States. She received a Master of Economics degree from North Carolina State University in 2005 and a Master of Science degree in economics from Louisiana State University in 2006.