2020

Keystone Sector Identification: A Graph Theory-Social Network Analysis Approach

Maureen Kilkenny

Laura Nalbarte

The Web Book of Regional Science

Sponsored by

Keystone Sector Identification
A Graph Theory-Social Network Analysis Approach

By

Maureen Kilkenny
Laura Nalbarte

Published: 2000

Updated: October, 2020

Editors: Scott Loveridge, Professor, Extension Specialist, Michigan State University

Randall Jackson, Director, Regional Research Institute, West Virginia University

<This page blank>

When citing this book, please include the following:

Kilkenny, M. & Nalbarte, L. (2000). *Keystone Sector Identification: A Graph Theory-Social Network Analysis Approach.* Reprint. Edited by Scott Loveridge and Randall Jackson. WVU Research Repository, 2020.

<This page blank>

Contents

<This page blank>

ACKNOWLEDGEMENT

The research on "Keystone Sector Identification" was financed by the Tennessee Valley Authority (TVA) Rural Studies Program, 1997-8. The TVA Rural Studies Program's mission areas include finding new ways to describe and analyze rural economies, enhancing entrepreneurial/business activity, and supporting innovative community change. TVA funds supported author Nalbarte's work in partial fulfillment of the requirements for her Masters Degree in Economics at Iowa State University, granted December 1998.

1 INTRODUCTION

This is about a new a method for identifying **keystone sectors** in communities, where sectors are broadly defined to include churches, clubs, associations, and public institutions as well as firms and businesses. In an arch, the *keystone* is the one with the unique wedge shape at the top of the arch that is critical for the arch's structural stability. While all other stones in an arch substitute for one another and can be removed (in pairs), the arch will fall apart if the keystone is lacking. The term **keystone species** was first coined by ecologists in the late 1960s with respect to the species uniquely responsible for the structure and integrity of an ecosystem. We now coin the term for use in community development analysis. In a community, the keystone sector is one that plays a unique role, without which the community is fundamentally and detrimentally altered.

The method presented here allows a researcher to do three main things: (1) quantify a community's structure, (2) describe the roles of all types of entities in that community, and (3) identify the critical, or keystone, entities in the community. This new method can highlight the importance of entities that would not even be considered, much less identified as critical, using the prevailing 'critical sector' or traditional 'industrial targeting' approaches familiar to economists.

1.1 Prevailing Methods

1.1.1 *I-O Based Methods*

There is a long tradition among applied economists seeking useful answers to the question: "What should be given priority for economic development funds or special credit?" *Target Industry Analysis*, developed by TVA, is the prevailing method used to respond to this question. Target industry analysis (a.k.a. critical sector or **turnkey industry** analysis) attempts to match industries with locations for the mutual benefit of the firms and the communities. The method keys off the identification of relatively **dominant industries**, those sectors with relatively high **location quotients**, that also have a number of input-output linkages with other local industries with high location quotients. The inference is that industries that are already in the area are the best indicators of industries that should be recruited.

Two approaches are typically used to estimate the size of the potential market for the recruited industry. An industrial 'linkages' approach estimates local demand as the national input-output industry×industry intermediate use coefficients times local industry activity. A trade area approach estimates final demand, taking distance, population density, and local income into account using a gravity formula.

Consider the popular *linkages* approach. What economists have been calling critical **or key sectors** are those sectors whose structure of backward and forward linkages create above-average impacts on the rest of the economy (Sonis et al., 1998). Given a record of a community's industrial input-output transactions, the strength of the effects of a purchasing sector on the sectors it purchases from, or **backward linkages,** are given by the column coefficients for the sector in the **multiplier matrix** . The **forward linkages,** or the effects of a sector on the other sectors it sells to, are given by the row coefficients of the multiplier matrix. These would be the sectors whose column and/or row coefficients are higher than average.

> "Any unit change in final demand of a particular sector will affect demand and supply of intermediate inputs. The demand stimulates other domestic sectors to satisfy its intermediate requirements (backward linkage). The supply also stimulates domestic production because it may induce use of its output as an input in new activities (forward linkage)... Interdependencies among productive sectors indicate each sector's potential capacity to stimulate other sectors. Activities having the highest linkages are considered key sectors because by concentrating resources in them it should be possible to stimulate a more rapid growth of production, income and employment than with alternative allocations of resources." (Cella, 1984)

The strengths of input-output based methods for identifying important sectors are that (i) the biproportional sector-sector transactions table is a popular and useful way to document sectoral interdependencies; (ii) the definition of the system boundary is very operational (i.e., $Y = AY + X \geq Y = mX$ where $m = [I - A]^{-1}$); (iii) the I-O methods have proven useful in many other contexts.

The weaknesses of I-O based methods include (i) I-O data is prohibitively costly to collect; thus (ii) U.S. national tables are released only every five years, with a four-year lag (e.g., 1997 data will be available by 2001); and (iii) there is no comprehensive sub-national data; (iv) the I-O approach is limited in scope: only current account (money-for-things) transactions are considered, and (v) only between business entities. There is no consideration of other types of transactions or relations, such as capital account transactions (money-for-paper), charity, social relations, public goods and services, information flows, or institutional support. There is no consideration of the roles of any non-business entities in a community.

1.1.2 *Revealed Comparative Advantage*

Another way used to determine which industries to target for an area is to survey existing businesses in attempts to identify the *regional comparative advantage.* In general, a region's comparative advantage activities are those which employ their relatively abundant factors of production most intensively (Hechscher-Ohlin-Samuelson); or those industries in which local producers are more productive (fewer inputs per unit output) than other region producers (Ricardian). The regions with the resources or technology to produce certain items relatively more cheaply should be able to profit most by specializing in those items. Market frictions, however, such as labor or capital immobilities, or, high fixed or transport costs, can undermine these hypothetical gains from specialization and trade in comparative advantage goods.

To identify the comparative advantage sector requires measuring and comparing relative factor abundancies, or relative sectoral factor productivities, in the region to those of the trading partners. Alternatively, some analysts assume that a region's comparative advantage sectors are *revealed* by their existing production and trade patterns. The deduction is that if a region is specializing in and exporting something in a competitive market economy, it must be its comparative advantage sector.

This assumption means that rather than attempting to assay local factor endowments or productivity objectively, a direct survey approach is often used to identify a region's comparative advantage sector. Businesses are asked about the location's proximity to input suppliers and/or output consumers, labor force costs, energy costs, and the industrial climate established by the state/local government, among other things. Their answers are used to profile a region. The validity of the approach relies on the objectivity of the survey respondents and their abilities to compare their chosen locale's feature to the features of other locations. Their answers to the questions are then interpreted in comparison with other, previous, studies. Two types of conclusions can be drawn. One is that more of the existing comparative advantage activity will be better. The other is that that industries which have located in areas with comparable profiles may be interested in locating in the study area as well (Goode and Hastings, 1988).

1.2 Results of Targeting Critical or Comparative Advantage Sectors

The lists of businesses in industries that may be recruited to the region, identified by either of the methods described above, are then screened for desirability for the community. The criteria include wage rates or value-added per employee, demand for the unemployed portions of the local labor force, ability to serve regional excess demands, expected growth potential, and other location specificities. The highest ranked industry is then targeted for recruitment. The process is considered successful if the business comes. The degree of success is typically measured by the number of workers initially employed locally by the recruited firm.

The results of even the "successful" industrial targeting programs, however, have been criticized for merely crowding-out private activity, or, being an imbalanced use of funds (Jacobs, 1985). The 'crowding out' argument is that if no market failure has been identified, why should public support be needed by a key or comparative advantage sector that is valued and profitable in a location ? If there are no market failures, public support does only what the market would have done anyway (crowding out), and, at a higher cost. The 'imbalanced use of funds' argument is that the approaches used to identify the target industries probably lead to too much capital investment in that one type of industry. The particular over-investment outcome depends on the sector identification method employed.

The "critical sector" and "revealed" approaches are biased towards the outcome that all locations have the same mix of industries, especially those industries with high linkages according to nationwide I-O data. In that case, there may be too much investment in such sectors economy wide. The "comparative advantage" approach is biased in the opposite direction, away from industries common in other regions, and towards more unique industries already thriving in the region of interest. In this case the targeting of public support may lead to overinvestment in the region-specific industry, with an **immiserizing growth** outcome. Neither key sector ID approach attempts to ascertain whether or not the local industry has already achieved optimal scale in the location or in the relevant market. Neither approach attempts to ascertain why further expansion in an industry which started in response to private incentives is not possible without public subsidy.

In addition to unrealized income growth, the other outward signs of too much investment in one type of activity include (i) environmental degradation, (ii) stagnation or decline in other sectoral activity, (iii) leakages of capital-related income, and (iv) population migration out of the area (Jacobs, 1985).

Given that the inspiration for this new way to identify key sectors comes from ecological sciences, an appropriate analogy is that industries induced to locate in an area by subsidies and tax holidays are "predators." In addition to using up tax revenues, the industries may 'feed' on the earning capacity of privately initiated local firms in the same industry. Subsidized industrial expansion can eclipse the existing industries in markets for local labor, similar inputs, or similar consumers. Excessive expansion of comparative advantage sectors may destroy natural resources. The targeting of a limited type of industry reduces the diversity of the local economy. A less diverse economy is more sensitive to exogenous shocks originating in those sectors. Finally, especially if it is the major employer, an industry designated as "critical" can more effectively threaten to abandon the community if substrate supplies, subsidies, or other attractions diminish (Wolhgemuth and Kilkenny, 1998). If such industries succeed in obtaining further public subsidy, the "predator" may be become a "parasite."

1.3 Keystone Species Concept

The dominant industry and predator or parasite analogies introduced above are the natural point of departure for further elaboration of useful concepts from ecology. As noted in the introduction, this project develops the keystone species concept coined by natural scientists for application by social scientists in the analysis of economic systems. We adopt the ecologists' definition of the keystone as being unique and without which the system structure would be fundamentally altered. We develop our own approaches for describing the roles of entities in a system, and our own test of 'keystoneness,' using techniques of graph theory from mathematics and social network analysis from sociology. The new methods are demonstrated in an application to data about a single community.

The concept keystone species was first coined by ecologist Robert Paine in the late 1960s, in his identification of a predator as the critical species in an ecosystem. Paine described the effect of removing a species of starfish, the main predator on mussels, from a rocky intertidal zone. Without the starfish, the whole system changed. The mussels proliferated, leading to a significant change (increase) in biodiversity. Paine called the starfish a *keystone species*:

> "The species composition and physical appearance were greatly modified by the activities of a single native species high in the food web [starfish]. These individual populations are the keystone of the community's structure, and the integrity of the community and it's unaltered persistence through time...are determined by their activities and abundances." (Paine, 1969)

Over the decades, the keystone species concept has evolved beyond it's focus on predators, but it remains "poorly defined and non-specific in meaning," (Mills, Soule and Doak, 1993). Natural scientists have tended to proliferate studies arguing the 'keystoneness' of their favorite organism. The emphasis in that research has been on single species as the unit of observation, not on the ecosystem or the patterns of interdependence among all the fauna and flora in one system.

In addition, the implied assumption that there will always be one and only one keystone in a system is suspect. Plants, micro-organisms, prey, the physical environment, and other aspects in an ecosystem may all play critical roles. As Princeton University ecologist Simon Levin says, "focusing on particular species often misses a great deal of what's important in an ecosystem" (Stone, 1995). By analogy, in the analysis of community economic systems, the exclusive focus on only one type of species (private sector businesses) or one type of relation (current account money flows) may also 'miss a great deal of what is important.'

It is instructive to note that the original 'top-down' keystone species concept in Biology has been fundamentally challenged by a more holistic, 'bottom-up' concept of *functional groups*, which recognizes that small organisms on the bottom of a food chain can also determine an ecosystem. The *functional group* concept is exemplified by the work of Robert Stenek, who studied a "curious pattern of continuity in the face of drastic change" among algae clusters floating around the Caribbean island of St. Croix.

> 'Almost none of the dominant species I'd find in one season would I find as dominant in the next,' Stone quotes Stenek, but "what was going on in the mats varied little: They consumed nutrients at a fairly constant rate and resisted similar types of predators. Different algae, apparently, were performing the same biological role" (Stone, 1995).

It is now widely understood that whole suites of organisms have similar functions in an ecosystem, and they can be modeled as a single group rather than species by species.

The *functional group* notion in biology corresponds to the economic notion of **substitutes:** inputs in production, or outputs in supply. Ecologists are now trying to develop methods to identify influential functional groups; i.e., *keystone functional groups.* By analogy, we economists seek to identify keystone sectors, where a 'sector' or group of sectors corresponds to a 'functional group' whose members substitute for one another.

In what follows, we define a community to be the set of businesses and institutions in a single geographic location, such as a town or city. We define the keystone sectors, to be the type of businesses or institutions, private or public, which play a unique and critical role in achieving the objectives of a community. By critical we mean necessary for the existing structure: without that type of entity, the structure is destroyed. We use a subset of *graph theory* (Berge, 1962) called *social network analysis* (Wasserman and Faust, 1994) to describe the structure of the community system. We also use it analyze the interrelationships between varied entities and to identify "species" or types. We introduce new tests of the sensitivity of the system to the absence of an entity.

1.4 Preliminary Analyses

As a preliminary step, we analyzed a survey of small businesses (Kilkenny, Nalbarte and Besser, 1999) in 30 small communities in Iowa. We sought a way to document the interdependencies among businesses, institutions, and local citizens. In that work, we tested and found statistical support for a "**social embeddedness**" hypothesis. We found that businesses whose owners or managers make more donations to their community and/or who serve as a volunteer or as an elected public servant, feel twice as successful as those who do not. The service, however, must be reciprocated by community support of the business. There is no evidence of differences across sectors or across towns by size. Neither the activity of the business nor the size of the town was found to be relevant. The findings are the same regardless of community sizes or different business activities.

We concluded from our work on "Reciprocated Community Support" that the usual market interactions and economic characteristics of business activity, typically thought to be most relevant in explaining firm success, *were far less relevant than non-market interactions.* This was a surprising finding, but not an unprecedented one. Kranton (1996) has argued that non-market reciprocal exchanges are a substitute for search and/or transaction costs. Non-market transactions can also help reduce damaging opportunistic behavior among strangers.

Economists concerned with the political economy of growth have also focused recently on the relationship between social capital and macroeconomic performance (Putnam, et al., 1993; Knack and Keefer, 1997). Almost all of these studies attempt to measure a dependence of macroeconomic performance (growth in regional or national gross product) on region-wide indicators of associational activity, trust, and civic cooperation. The hypothesis is that high-trust societies waste fewer resources protecting themselves from malfeasance; have cheaper, more credible and stable governments institutions; have more access to credit; and risk more on innovation-all of which lead to higher rates of national investment and national growth (Fukuyama, 1995; Knack and Keefer, 1997).

In sum, our preliminary work verified the social capital arguments by sociologists and economists about the importance of non-market interactions and social institutions in the economic success of a community. This further compelled the search for a way to identify a community's key sector(s) that considers a broader range of entities (not just businesses) and a broader range of relations (not just sale/purchase money flows) than the prevailing methods. However, in the data for our preliminary empirical work, only one side of an interaction was interviewed. That analysis relied on one side's opinions of how the other side feels. The experience clarified that to test hypotheses concerning reciprocity, it is necessary to sample both sides of relations.

2 GRAPH THEORY and NETWORK ANALYSIS

This section introduces a method to identify keystone sector(s) from among *many* possible types of entities, taking into account *numerous* possible types of interdependencies in a community. The method must (1) describe interdependencies within and among agents, institutions, sectors in communities, (2) determine the degree of importance of an agent or groups of agents, and (3) show the sensitivity of the structure of the community to the absence of particular types of agents. Network analysis methods are appropriate for all three tasks.

Network analysis has been widely used in transportation system research (Hanson and Huff, 1986; Koppelman and Pas, 1985; Wright, 1979) and anthropology and sociology research (e.g. Granovetter, 1973, Freeman, 1977). Applications of graph theory or network analysis to identifying critical economic sectors, however, are scarce. The few attempts include the early work of Campbell (1975), and some recent initiatives by Kauffman (1988), Roy (1994, 1995), and Sonis and Hewings (1997).

2.1 Social Network Analysis

Statistical models based on graph theory have also been used by researchers to study **social networks** for almost 60 years. The goal of these models was (and remains) the quantitative examination of the stochastic properties of social relations between the actors of a particular network (Wasserman and Pattison, 1996). Applications range from studies of interactions between individuals: interpersonal relations, friendship, leadership, etc; to studies of interactions between groups: global studies of communities, studies of the élite and political behavior; and studies of power-sharing.

> "Social network analysis may be viewed as a broadening or generalization of standard data analytic techniques and applied statistics, which usually focus on observational units and their characteristics. A social network analysis must also consider data on ties among units" (Wasserman, and Faust, 1994).

The basic feature of network analysis, as distinct from the more usual data analytic framework common in the social sciences, is the use of relational information to study or test hypotheses. A **relation** is the collection of ties of a specific kind among a set of entities or actors. Relational data can include, for example, data on family ties, interactions between people, or individuals' attitudes about other individuals in a group. The relational link between a pair of actors is called **tie**. A tie is a property of the pair, therefore a tie can not be thought of pertaining simply to an individual actor.

Since ties exist only between pairs of actors, the relevant unit of analysis is the *dyad*. A **dyad** consists of a pair of actors and the possible ties between them (Wasserman, and Faust, 1994). For example, a father and son are a dyad, and their familial relation is implied by their labels. Two cities connected by a commuter's travel pattern between them, and a retail store and customer, are also dyads.

For social network analysis, relational data are collected by observing or interviewing individuals about their interactions with the others in the set or network. The unit of observation is an individual from whom we obtain information about their ties with other actors. For economic analyses, relational data may include data on the values of purchases or sales between firms, the existence (or lack) of contractual agreements or information flows between agents, or the value, volume, or frequency of international trade flows among countries.

Data on social interactions between entities presented in a matrix are called a **sociomatrix** . By convention, the rows of the sociomatrix represent the sending actors while the columns represent the receiving actors (note: this is the opposite of Social Accounting Matrix (SAM) conventions). A sociomatrix need not be square: the set of sending and receiving actors may be the same or different.

There are two main types of ties or relations: i) *dichotomous* or *valued*, and/or ii) *directional* or *non-directional*. A **dichotomous relation** is recorded as either the presence or absence of a tie between two entities in the set. Meanwhile, a **valued relation** records not only the existence of a relation but also the intensity or frequency of the relation (Wasserman and Faust, 1994). An example of a dichotomous relation is public safety agency A's provision of services enjoyed by business B. Since a public good is by definition non-rival, use or non-use is the relevant measure (rather than quantity used.) An example of a valued directional relation is the dollar value of exports recorded as being shipped from country A to country B.

A **directional relation** has an explicit origin and destination. A **non-directional relation** is non-specific about the origin or destination of the flow on the link (Wasserman and Faust, 1994). By convention in graph theory, a non-directional relation is represented by an **edge**. It is illustrated by a line between the interacting agents that has no arrowhead. A directional relation is represented by an **arc**. An arc is a line between entities with an arrowhead at the destination. For example, if country A exports manufactured goods to country B, the direction of trade is from A to B, reflecting B's imports from A. If the direction is not recorded, the trade flow could be misconstrued to be about exports of B. Graphically, the directed relation of A's international trade exports to B is shown as $A \rightarrow B$. The matrix form for recording this relation would have an entry in the A row and the B column (opposite to conventional SAM accounting).

2.2 Digraphs

As suggested above, relational data can be presented not only in matrix form but also by graphs. If the relational ties have direction (are arcs) the graphs are called *directed graphs* or **digraphs**. The entities in digraphs are called **nodes** and the relations are represented by arcs. Formally:

> A digraph is a finite, non-empty set $\boldsymbol{N}$, whose elements $n_i = \{n_1, n_2, \ldots n_g\}$ are called *nodes*, together with a set $\boldsymbol{A} = \{a_{12}, a_{13}, \ldots a_{1g}, \ldots a_{g-1,g}\}$ of ordered pairs a_{ij}, called *arcs*, where n_i and n_j are distinct members of $\boldsymbol{N}$ (Robinson and Foulds, 1980).

The graphic form shows how each agent in a system relates to all other agents in the system. If the number of agents (g) is not too large, a graph is an efficient way to show which agents are connected to which others, and which are isolated; which are senders or receivers. **Figure 1** is a digraph of 73 agents. It is a good example of the case where the number of agents is "too large." See Freeman (2000) for a comprehensive explanation of various ways to illustrate networks.

Figure 1. Digraph of the INFORMATION relation among Towertown entities

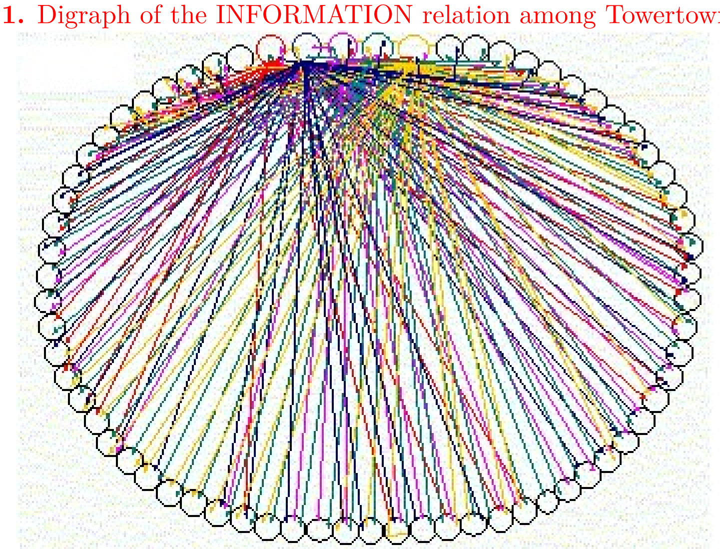


Figure 1 shows a multitude of dichotomous, directional ties between nodes. Clearly, the more fruitful form in which to analyze these ties is as a matrix of zeros and ones. Adjacency is the graph theoretic expression of the fact that two agents, represented by nodes, are directly related, tied, or connected with one another. Formally, given agents n_i and n_j in a set of agents $\boldsymbol{N}$, and the $\boldsymbol{A} = \{a_{ij}\}$ arcs denoting the existence of relations from agents i to agents j; agents i and j are *adjacent* if there exist either of the two arcs, a_{ij} or a_{ji}. Given the digraph $\boldsymbol{D} = (\boldsymbol{N}, \boldsymbol{A})$, its **adjacency matrix** $\boldsymbol{A}(\boldsymbol{D})$ is defined by $\boldsymbol{A}(\boldsymbol{D}) = \{a_{ij}\}$ where $a_{ij} = 1$ if either a_{ij} or a_{ji} exists, and 0 otherwise.

Note also that Figure 1 shows no *reflexive ties*. A **reflexive tie** is the one that an entity has with itself. In an adjacency or a sociomatrix, its presence would be recorded by $a_{ii} = 1$. In a digraph, reflexive ties are drawn as arrows that originate and end on the same node, that are curved back on themselves.

Looking at Figure 1 it seems like there are many interactions among the entities, but in fact it is far from *complete*. A **complete** graph is one in which all the actors have two-way ties to all other actors. In fact, in the digraph in Figure 1, only 24% of the entities have ties. The density of a complete graph is 100%. Formally, the **density** of a digraph ($\boldsymbol{D}$) is the actual number of non-reflexive arcs in proportion to the maximum possible number of non-reflexive arcs: $\boldsymbol{D} = \sum_i \sum_j a_{ij}/\boldsymbol{N}(\boldsymbol{N}-1)$. Note that the numerator is the count of non-reflexive actual ties: $\sum_i \sum_j a_{ij}$ for all $i \sum \neq j$. In the case of a complete network, all nodes are reciprocally adjacent to one another, and all elements of the sociomatrix are equal to one.

The nodes at the top of the oval in Figure 1 are the sources or sinks for the largest number of ties. The strength of a node as a source or a sink in a system is most easily measured using the sociomatrix or adjacency matrix data. The number of arcs beginning at a node is called the **outdegree** of the node. Given dichotomous sociomatrix data, outdegree is the row sum for the node:

of actor $i = \sum_j a_{ij}$. (1)

The number of arcs ending at a node is called the **indegree** of the node. The indegree is measured by the column sum for the node in a dichotomous sociomatrix:

indegree of actor j = $\sum_{\mathrm{i}} a_{\mathrm{ij}}$ (2)

According to social network analysis, the nodes at the top of the oval in Figure 1 are the most *prominent* because there are more links between them and the others. There are four measures of prominence: **local centrality**, **local prestige**, and *global centrality* in two forms: **closeness** and **betweenness**. Local centrality reflects the number of direct transmissions, and is thus measured simply by the outdegrees (or row sums) for each actor. Local prestige reflects the number of direct receipts, and is thus measured simply by the indegrees (or column sums) of each actor. Since these measures are based on the *degrees* of the nodes they are also known as *degree centrality and degree prestige* (Wasserman and Faust, 1994).

Actors may be *directly* related (with a one-step arc between them) or *indirectly* related through a third, fourth, or more remote entity. A **path** is a sequence of arcs from one node to another. When multi-step paths are taken into account, agents can be characterized more richly than just as sources or sinks. With respect to a path, an actor can be a **transmitter** (the arc is away from the node), a **carrier** (there are at least two arcs, one toward and one away), or a **receiver** (the arc is toward the node).

For example: consider the following digraph and adjacency matrix illustrating the relation between four entities A, B, C, and D:

$A \rightarrow B \rightarrow C$

D

In the above, A is a transmitter, B is a carrier, and C is a receiver; while D is **isolated**. An actor is **isolated** when there is no arc that relates the actor to any other in the network. The adjacency matrix for this digraph is:

	A	B	C	D
A	0	1	0	0
B	0	0	1	0
C	0	0	0	0
D	0	0	0	0

which shows that A relates *to* B, and B relates *to* C, while C does not relate OUT (only IN), and D does not relate at all.

Formally, a node is a proper source or *transmitter* if its indegree is zero and its outdegree is non-zero; that is, if the column sum is zero while the row sum is greater than zero. A node is a proper sink or *receiver* if its indegree is non-zero and its outdegree is zero, that is, if the column sum is greater than zero while the row sum is zero. A node is *isolated* if both indegree and outdegree are zero (Wasserman and Faust, 1994).

The distance between entities in a network is measured by the length of the path between them. Of course, there many be many paths, but the shortest path is the relevant one. The shortest path from n_i to n_j is called the **geodesic**. When nodes i and j are adjacent or tied in one step, the shortest path is the value of the sociomatrix element a_{ij}. When nodes i and j are related in a two-step path, the ijth element of $\boldsymbol{A}^2$ will be non-zero. If the shortest path is in three steps, the ijth elements of $\boldsymbol{A}$ and $\boldsymbol{A}^2$ will be zero, but the ijth element of $\boldsymbol{A}^3$ will be non-zero. In general, the pth power of the adjacency matrix shows the existance of all directed p-step paths. Consequently, the **geodesic**, $d(n_i, n_j)$, is measured as the first power p for which the ijth element of $\boldsymbol{A}^p$ is non-zero Wasserman and Faust (1994).

Now we can define the last two *global centrality* measures of prominence. The ubiquitousness of an actor's ties qualify them as *global* and, if the actor is in the shortest paths between others, the actor is *central*. The shortness of the paths from that actor to others is measured by **closeness**. The proportion of intermediary roles an actor plays measures **betweenness** . Both measures rely on the geodesic measure of distance.

Closeness, $\boldsymbol{C}(n_i)$, is the inverse of distance. The shorter the distances between node i and other nodes, the more central node i is. Formally, $\boldsymbol{C}(n_i) = [\sum_j^N d(n_i, n_j)]^{-1}$ where $d(n_i, n_j)$ is the geodesic (shortest path) distance between i and j entities in the network, and $\boldsymbol{N}$ is the network size.

Betweenness, $\boldsymbol{B}(n_i)$ measures the probability that a path from actor j to actor k takes a particular route through agent i. Assume each one-step tie has equal weight, and assume that interactions will occur through the shortest routes. Formally, $\boldsymbol{B}(n_i) = \sum_j^N \boldsymbol{g}_{jk}(n_i)/\boldsymbol{g}_{jk}$ where $\boldsymbol{g}_{jk}(n_i)$ is the total number of geodesics through i, and $1/\boldsymbol{g}_{jk}$ is the probability that a particular geodesic is chosen.

All these measures depend on the size of the network. Thus, the measures must be standardized before comparisons are made between networks of different numbers of entities. In degree centrality and closeness measures, the measures are standardized dividing by $\boldsymbol{N} - 1$. For betweenness measures, the standardization factor is $(\boldsymbol{N} - 1)(\boldsymbol{N} - 2)$, where $\boldsymbol{N}$ is the number of agents or network size.

Agents in a network that all relate to each other can be classified as a subset of the N actors as a *group*, a *sub-graph*, or as we shall say here, as a *component*. A **component** is the largest subset of related actors in a network. In the $ABCD$ example above, three of the actors are linked to one another through one or more step paths. These three, ABC, are thus a sub-graph, group, or component . The whole network has two components, the group ABC and the singleton D.

Components have two forms: *strong* and *weak*, depending on the directions of the ties, arcs, or relational links between the members of the component. A *strong component* is one in which the arcs that make up the paths are aligned in continuous chain without a change of direction. For example, ABC make up a strong component. A *weak component* is made of actors that are linked by non-directional edges (Scott, 1991).

3 NETWORK ANALYSIS of a COMMUNITY

In this section we develop these graph theory and social network techniques to answer the basic question: what entity(ies) are critical in a small community? We cannot conclude which sector is always key in communities because generalizations should not be drawn from observations about just one community. If we can answer this question for one community, however, we may be motivated to develop the techniques to answer the question about a sample of communities.

To conduct a network analysis toward keystone sector identification that considers more than input-output relations, one needs dyadic data on a variety of relations among a variety of entities in a community. We found this type of data in a collection complied by Lauman (1985). In the late 1970s, the sociologists Galaskiewicz and Marsden gathered information on the formal organizations in a small U.S. town. The community was nicknamed, and referred to since, as Towertown. Of the 73 entities they studied, less than a quarter are private businesses (**Appendix I**). They assayed three relations among these 73 entities: money, information, and support.

In this section we apply network analysis techniques to the Towertown data to do five things. First, we describe community-wide patterns: Is there a specific structure among agents, or do agents interact randomly with each other? Second, we describe the patterns of interactions among the individual agents in the community and compare their various roles. Third, we determine the groups of individual agents whose patterns of interactions are sufficiently similar that they can be treated as *types* of agents. This step is analogous to classifying numerous individual organisms into less numerous *species* or *functional groups.* If we cannot do this, the number of agents will remain "too large," as illustrated in Figure 1. Fourth, we assign ties to the sociomatrix data aggregated according to functional groups. Fifth, we test for the impact of the excision of each functional group from the community network. Are there any species that play such important roles that the community structure will be destroyed without them? This final step completes our identification of the "keystone" sector. Recall that we have defined the keystone as the entity that is relatively unique and without which the community structure will be fundamentally disrupted.

3.1 Towertown Survey

Towertown is the nickname given by to the midwestern community of 32,000 persons studied by Galaskiewicz and Marsden; Laumann (1985). A total of 109 organizations were identified in Towertown, and 73 were studied. The 73 included all business firms with more than 20 employees, banks, law firms, political organizations, associations, health institutions, educational institutions, service clubs, labor unions, city offices and departments, and churches.

Galaskiewicz and Marsden interviewed the executive officers of the 73 organizations. Each subject was presented with a list of the other 72 organizations in the community, and was asked the following six questions:

1. To which organizations on this list would your organization be likely to pass an important *information* concerning community affairs?
2. To which organizations on this list does your organization rely upon for information regarding community affairs?
3. To which organizations on this list does your organization give substantial *funds* as payments for services rendered or goods received, loans, or donations?
4. From which organizations on this list does your organization get substantial funds as payments for services rendered or goods received, loans, or donations?
5. Which organizations on this list does your organization feel a special duty to stand behind in time of trouble: that is, to which organization would give *support*?
6. Which organizations on this list would be likely to come to your organization's support in time of trouble?

From the responses to these questions, three dyadic relations were defined: INFORMATION (1,2), MONEY (3,4) and SUPPORT (5,6). An organization was determined to be "in relation to" another organization if the former organization answered yes to the first question in a pair, or the latter organization answered yes to the second question in the pair. Note that if either actor in the dyad reported the existence of a tie, a tie was recorded. In other words, for each relation R a 73×73 adjacency matrix ($\boldsymbol{A}^R$) was constructed with entries $a_{ij}^R = 1$ if the ith actor has a relation R tie with the jth actor and $a_{ij}^R = 0$ if not. (Also, $a_{ii}^R = 0$.)

3.2 Macro/Community wide structure

3.2.1 *Density*

The first step is to study the density and connectivity of the whole network. We describe community-wide patterns: Is there a specific structure among agents, or do agents interact randomly with each other? The **density** measure describes general level of linkage among the actors in the community. This measure compares the number of actual to possible relations, to show how far from completion the Towertown community network is.

Since all three sociomatrices about Towertown have the same symmetric number of actors ($\boldsymbol{N} = 73$), the maximum possible number of non-reflexive arcs is $73 \times 72 = 5256$. The arcs present in the INFORMATION sociomatrix are 1264, while the ones present in the MONEY and SUPPORT sociomatrices are 512 and 814 respectively. Consequently, the density measures are 24%, 9.7% and 15.5% for INFORMATION, MONEY and SUPPORT, respectively. INFORMATION linkages are more *dense*, and thus more *complete*. MONEY relations are the least dense and thus least complete in Towertown.

3.2.2 *Components*

If a sociomatrix is dense, we expect to find a single component. If not, there may be multiple components of various sizes. None of the three Towertown sociomatrices are dense (recall the INFORMATION sociomatrix is the most dense, having 24% of the maximum number of ties). So, we expected to find multiple components in the Towertown network. But we did not. A glance at the matrices indicates that every entity has at least one source or sink of INFORMATION, MONEY, and SUPPORT. There are no columns or rows that have all zeros. Furthermore, none of the connected entities are exclusive. The Towertown community appears to be one big, interconnected group. This analysis shows that in Towertown, there are no subgroups or components (no cliques), and no isolated actors.

3.3 Micro/individual roles

Now we know that there are no isolated actors in Towertown. But we also know that the links among all the actors are not all mutual. Next we apply the techniques to describe the roles of individual actors. Which actors are distinguished either because they receive a lot from other actors, or, because they give a lot to other actors? In this section we continue to build the network analysis lexicon by demonstrating the use of measures of indegree and outdegree to describe individual roles.

3.3.1 *Local Centrality and Prestige*

Prominent actors are those that are extensively involved in relationships with other actors. This involvement makes them more visible in the community. The prominence could be due to both receiving and transmitting. To determine which actors are prominent, we consider all the directed ties (arcs) originating from the actor (outdegree), plus all the received ties (indegree), plus all the indirect ties (multiple-step paths) as well.

Tables 1, 2 and 3 present the measures of **local centrality** and **prestige** for MONEY, INFORMATION, and SUPPORT. The measures are unstandardized with respect to the maximal number ties in each sociomatrix. In this case of equal dimension sociomatrices, and in which we are not comparing across networks in different size communities, standardization is irrelevant. But if the number of agents surveyed varied from relation to relation, or network to network, one would want to standardize each set of degree measures by normalizing with respect to the number of agents.

The first four rows in each of Tables 1-3 show the sample statistics for each degree measure for selected agents in Towertown. The mean of local centrality, for example, is the average outdegree among the 73 entities. With respect to MONEY ties (**Table 1**), on average, a Towertown entity *gives* money *to 7* other entities in Towertown. The mean of local prestige is the average indegree measure across all 73 entities. For example, with respect to INFORMATION ties (Table 2), on average, a Towertown entity *receives* INFORMATION *from 17* entities in Towertown.

The 5^{th} to 11^{th} rows in each table show the local centrality and prestige measures of the top six most central and prestigious individual entities in Towertown. For example, with respect to MONEY (**Table 1**), Towertown Newspaper is the most locally central, giving money to *33* other entities in Towertown. 1^{st} Towertown Bank is the most prestigious, receiving money (presumably deposits or interest payments) from *49* entities in Towertown.

Table 1. MONEY	LOCAL CENTRALITY	LOCAL PRESTIGE
sample statistics:		
Mean	7.01	7.01
Std. Dev.	6.35	8.97
Minimum	0	0
Maximum	33	49
top six entities:		
(11) **1^{st} TT Bank**	28	**49**
(12) TT Saving bank	17	38
(13) Bank of TT	23	26
(39) **TT Newspaper**	**33**	17
(69) Family Services	3	36
(71) YMCA	4	24

Table 1 shows that four of the top six agents are both locally central and prestigious . These agents are three of the four banks and Towertown Newspaper. Family Services and YMCA, on the other hand, are only locally prestigious (high indegree but low outdegree, and both outdegree measures are below the mean). These measures identify some intuitively logical patterns and relative positions. Banks both give (make loans) and receive (accept deposits) money; and the local centrality and prestige measures document that. Also, Family services and the YMCA both rely on volunteers and donations, so it is logical that they would show up as the major recipients of money. Note also that the newspaper is a source of money for many entities.

In the case of the INFORMATION relation (**Table 2**), again the top six agents appear to be both locally central and prestigious. All six agents receive and give information to a significantly (> 2 SD) higher number of agents in the community than the rest of the entities. The Radio Station appears to be the most locally central and prestigious in Towertown.

Table 3 shows the results with respect to the SUPPORT relation. The Community College both receives and gives more support than all other entities. "In times of trouble," the Community College is willing to give support to 59 entities, and 58 entities are willing to give support to the Community College.

Towertown Savings Bank appears to be locally central with respect to SUPPORT, but not locally prestigious (has high outdegree but the indegree value is close to the mean). This contrasts with its local centrality being lower than local prestige according to the MONEY relation. Note that Towertown Savings Bank has fewer out- than in-degrees, suggesting that it makes loans to fewer entities in the community than the number from whom it receives deposits. The bundling of many small deposits to make a few larger loans is the main business of financial intermediaries, known as "asset transformation." It is interesting to note, however, that this particular bank does not attract as much SUPPORT as it says it would give.

Table 2. INFORMATION	LOCAL CENTRALITY	LOCAL PRESTIGE
sample statistics:		
Mean	*17.29*	*17.29*
Std. Dev.	*11.21*	*11.22*
Minimum	*1*	*3*
Maximum	*63*	*62*
top six entities:		
(12) TT Savings bank	41	26
(25) City Council	32	43
(26) City Manager's office	43	44
(39) TT Newspaper	42	44
(40) WTWR Radio	**63**	**62**
(69) Family Service	45	36

Table 3. SUPPORT	LOCAL CENTRALITY	LOCAL PRESTIGE
sample statistics:		
Mean	*11.12*	*11.12*
Std. Dev.	*9.42*	*10.53*
Minimum	*0*	*0*
Maximum	*59*	*58*
top six entities:		
(12)TT Savings bank	37	14
(20) Small Bus.Assoc.	38	10
(56) **TT Comm.College**	**59**	**58**
(57) State University	20	42
(69) Family Service	12	52
(72) Mental Health	16	36

Note also that different actors are prominent across the MONEY, INFORMATION and/or SUPPORT relations. Towertown Savings and Loan and Towertown Family Service entities are, however, central and/or prestigious in all three of the relations.

In sum: we have shown how to classify entities with large outdegrees as important sources of INFORMATION, MONEY, OR SUPPORT. Actors with low in- or out-degrees are apparently less active, less important, or peripheral to a network.

Note, however, the limitations of dichotomous rather than value data. In dichotomous (0/1) data, what counts is the number of ties, not their level. While this may be quite appropriate for the analysis of non-rival types of interactions such as information and support ties, it does not fully describe rival transactions such as money ties. It is likely that agents with a few high-value MONEY ties play more important roles than can be perceived with dichotomous data. For example, a business that buys large amounts of goods and services from just a few other local businesses (who then buy from other local businesses, employ many residents, and provide many donations; and so on) may not have so many ties, but because of their magnitude, its few ties may be critical to the community.

3.3.2 *Global centrality*

The *global centrality* measures are based on the length and the number of *carrier* and multiple-step path roles. When an actor has a position of strategic significance in the overall network, that actor is considered *globally central.* Interactions between non-adjacent actors depend, by definition, on an intermediary. Thus, globally central actors can have widespread effects on a community because they are more closely tied with more of the other agents, and, act as intermediary for more agents; than any others. Globally central actors are detected using the closeness and betweenness measures.

The summary statistics for global centrality in Towertown, according to both measures, with respect to MONEY, INFORMATION and SUPPORT relations are shown in **Table 4**:

Table 4. GLOBAL CENTRALITY sample statistics	**CLOSENESS**			**BETWEENESS**		
	MONEY	INFO	SUPPORT	MONEY	INFO	SUPPORT
Mean	52.34	59.06	53.42	88.51	57.68	78.73
SD	6.74	7.13	8.38	244.87	131.43	238.37
Minimum	37.50	46.75	32.00	0	0.06	0
Maximum	78.26	88.89	82.76	1762.24	1013.23	1998.32

The higher the closeness of an entity, the more often entities are directly tied to it. Note that on average there are more direct INFORMATION ties (the mean of closeness is highest) than SUPPORT or MONEY ties. The mean of betweenness is the average number of times an actor is articulating (or, is the intermediary in the shortest path), in a relation between two other actors. With respect to MONEY flows, on average a Towertown entity serves as intermediary in *89* relations between other entities.

Table 5 shows the measures for the most *globally central* entities according to closeness and betweenness measures. The higher the value of closeness, the shorter are the paths to the entity relative to other entities in the community. For example, 1^{st} Towertown Bank reaches the most other agents with the fewest MONEY steps. 1^{st} Towertown Bank is significantly closer (more than 2 SD's higher than the mean) than other entities in Towertown. Comparing two of the banks, 1^{st} Towertown Bank (closeness value = 78) is closer to the rest of the agents in Towertown than the Bank of Towertown (closeness value = 65).

Table 5. GLOBAL CENTRALITY	**CLOSENESS**			**BETWEENNESS**		
	MONEY	INFO	SUPPORT	MONEY	INFO	SUPPORT
1^{ST} TT Bank (11)	78			1762		
TT Savings (12)	71			686		
Bank of TT (13)	65			409		
Ch. Commerce					222	
City Council (25)		73				
City Mng's office		76			308	
TT News (39)	71	75		900	344	
RADIO (40)	67	89			1013	
CmtyCollege(56)			83		253	1998
State Univ. (57)			70			222
Family Serv (69)		76	77			191

The higher the values of betweenness, the more potential an actor has to control third-party relationships. Again, 1st Towertown Bank shows the highest betweenness measure, it appears to be in the strongest control position. In both MONEY and SUPPORT flows, the entities that appear to be the most 'close' also appear to be the most 'between' actors. This analysis of betweenness in the MONEY relation shows that banks and the newspaper are the 'main roads' to other entities. They act as intermediaries for more other agents than any other entities in Towertown. This helps explain why the newspaper is the most **locally central** (has the highest outdegree in the MONEY relation) according to the measures shown earlier in **Table 1**.

If we consider the flow of SUPPORT, the Community College is a hub in Towertown. The *global centrality* of the Community College is shown both by its closeness to the rest of the agents and because it has a major articulation role as evidenced by its maximum betweenness score.

The findings for flow of the INFORMATION relation are also consistent with our intuition or priors. Our analysis of the INFORMATION network in Towertown shows that the media organizations: the newspaper and radio; and the community's coordinating office: the city manager's office are *globally central*.

3.3.3 *Peripheral Entities*

This digression concerns how to identify (and exclude from further consideration?) the entities that are the least likely to be keystones. The actors with the lowest values of local and global centrality are the actors that are *peripheral* to the network. According to the Towertown data, the entities that do not have many interactions with other entities include the unions, the churches, and some of the government offices (measures not shown). Their counts of interactions are significantly (more than 2 SD) below the means.

Two actors are notoriously peripheral in all the relations: the bankers' association and the county bar association. Why that might be the case? The network analysis shows that BANKS relate to many entities. Banker-members of the banking association clearly have many ties as professionals. But as an association, the association's relations with other entities are few. This may be because there is no need for the association. It would be redundant to replicate the plentiful ties the member-bankers already have.

3.3.4 *Entities Critical to Network Structure*

Now that we know about the *centrality* and *prestige* of agents, we have filled in some gaps left from our analysis of components . In the analysis of components we showed that all Towertown entities are interconnected and there are no cliques or subgroups. This does not provide any hints that any particular set of agents display unique patterns of relations. The analyses of *centrality* and *prestige*, however, showed that there are differences among entities. Some, notably banks, are better connected than others. But we still do not know whether banks' connections are critical to maintaining the structure of the Towertown community. This next section presents tests for this aspect of "keystoneness". If the single-component structure of the Towertown network is destroyed by the excision of an entity, that entity is considered critical for the maintenance of Towertown's structure.

3.3.5 *Cut-points*

Entities that are vital for the connectivity of a network are those without which other entities will become isolated. Now we shall test if any individual actors that have been shown to be *central* are also critical to maintenance of the network structure. The central entities are, one at a time, removed from the data (delete their rows and columns from the three sociomatrices). Then we conduct another analysis of components. If the new network has more subgroups or components, the excised entity is a **cut point** . More formally, a cut-point is the actor whose removal from the system would increase the number of components by dividing the graph into 2 or more separate components, between which there are no connections (Scott, 1991).

The results are that some of the entities we identified above as *central* are also cut-points. In the MONEY relation, isolated actors were produced by removing three entities: 1st Towertown bank, Towertown Savings Bank, and/or Family Services. Removing any one of the two banks and/or Family Services breaks the single Towertown network into 5 components: two isolated actors : 1) the County Medical Society, and 2) the Association of Churches; two isolated dyads comprised of 3) the Municipal Employees Union I and the

Central Labor Union; and 4) University Methodist Church plus the Association of Churches in another dyadic component, with 5) all 64 others in the remaining component.

Similar results were found with respect to the SUPPORT relation. Removing central actors Towertown Savings and Loan, Towertown Community College, and/or State University breaks the single component structure in three components. The three components are: 1) The four Unions in one component, 2) The democratic Committee and the County Housing Authority in another component and, 3) all the 64 others in the remaining component.

With respect to INFORMATION, however, no entity appears to be critical. There is no change to the structure due to the removal of any actor. The single component structure in INFORMATION relation is robust. There is no single agent that does not have a substitute as a vector of information in Towertown.

3.4 BLOCK MODELING

To this point we have treated each entity as unique. There are at least two reasons to group similar entities. One, as the number of entities in a community increases, the patterns of interaction are more complex and detailed. The main patterns (forest) can become obscured by the details (trees). Two, it is also useful to know which entities play similar roles. By 'similar role' is meant that patterns of interaction with all other entities are similar or, that the entities can **substitute** for each other. In this section we show ways to group or block entities into species, types or sectors in a network. Then we analyze the network structure with respect to types rather than individual actors. This last step is called **block modeling** .

Block modeling is another form of network structural analysis (Holland and Leinhardt, 1979). It was introduced by White, Boorman and Breiger (1976) for the descriptive algebraic analysis of social roles. A *blockmodel* is created from an elemental sociomatrix data base by doing two things. First, the entities are partitioned into discrete subsets. This step is called "blocking." Second, 0/1 ties are assigned between each pair of subsets. The following sections explain how this is done by example using the sociomatrix data on the Towertown community.

3.4.1 *Blocking*

The blocking procedure ideally consists of partitioning of the population of entities into a set of 'equivalent' subgroups or blocks. The network analysis software UCINET has a procedure called CONCOR (CONvergence of iterated CORrelations). This procedure delineates blocks of stochastically **structurally equivalent** entities. CONCOR is used to identify sets of entities with distinct behavioral patterns. By analogy, we want to group the 73 individual entities in Towertown into a smaller number of species, types, or sectors.

For example, actors 1 and 2 are structurally equivalent if actor 1 relates only with 3 and 4, then actor 2 relates only with 3 and 4. When there is strict structural equivalence between two of $\boldsymbol{n}$ entities in a network, the sociomatrix will be of rank $\boldsymbol{n} - 1$; one of the rows (columns) contains exactly the same information as another row (column). One of those two can be deleted, or, the two can be collapsed into one, without altering the connectivity of the network.

The first step in the CONCOR procedure is to calculate the correlations between all ties between dyads in the sociomatrices. The result of this first step is a square case-by-case correlation matrix. The second step is a clustering procedure to group the cases into (stochastically) structural equivalent sets (Scott, 1991). CONCOR splits the whole set of cases into subsets that have all positive correlation within each subset, and negative correlation between. Using this assignment to groups, the rows (and columns) of the original sociomatrix are permuted (their order is switched) to put the members of the subsets next to each other. The resulting matrix puts the stochastically structurally equivalent individuals blocked together.

The structural equivalence criterion is widely used by sociologists to block individuals into groups. We learned from this experience with Towertown data, however, that for economic analysis, one must take particular care interpreting the nature of the composition of the groups that result from its application. The problem is that structural equivalence may group entities that are **complements** together, while entities that are substitutes may be separated. We propose the following more formal definition of substitutability in a network:

from the perspective of entity K, entities I and J are perfect substitutes if a slight decrease in the desirability of interacting with I leads to interacting with J instead

For example, consider consumers (K) brown eggs (I) or white eggs (J): the different eggs are perfect substitutes if consumers buy whichever is cheapest, however slight the price difference.

In perfect substitutes cases, according to our definition, one would observe either a tie between K and I, or a tie between K and J, but not both. The dyadic relations of perfect substitutes are mutually exclusive in a world without risk, which generally induces diversification across substitutes. Another violation of the perfect substitutes case is when potentially substitutable relationships are subjected to quantitative constraints. In that case, the most preferred would be used up to the limit imposed by the constraint, then the next best substitute would be used. In these two cases, where there is risk and when there are constraints, dyadic relations of perfect substitutes would NOT be mutually exclusive.

Even in the riskless, unconstrained cases, however, the CONCOR procedure would treat perfect substitutes that had mutually exclusive relations as structurally **non-**equivalent. Entities that substitute perfectly would be in separate groups. This is not desirable.

Alternatively, *entities I and J are perfect complements to K when K's interaction with I is always accompanied by K's interaction with J*. An example of complementarity is tea leaves (I), hot water (J), and tea cups (K): if a tea cup is used, it typically always has both tea and water in it. While the tea and the water go together, they are not substitutes for each other.

Note that the definition of structural equivalence is very close to this definition of complementarity. The CONCOR algorithm is expected to group perfect complements together.

To verify the (in)applicability of the algorithm, we first applied CONCOR to each relational sociomatrix separately. The results of the CONCOR blockmodeling procedure applied to single relation Towertown data was not satisfactory. The resulting groups were comprised of heterogeneous entities: mixes of businesses, voluntary organizations and government offices. As suspected, CONCOR grouped rather than separated complementary entities, and it separated rather than grouped substitutable entities. Few of the resulting groups displayed a clear role division of species. Also, there was a different block structure for each relation.

But the CONCOR group structure based on all three relations simultaneously, shown in **Appendix II**, is quite reasonable. Compare the dozen types of entities as listed in Appendix I with the fifteen groups determined by CONCOR listed in Appendix II. Most of the groups in the second list are comprised of what we would consider to be similar or substitutable types of entities (exceptions are **bolded**). The only agents that do *not* group in expected ways are the three law firms and the Savings and Loan. These entities appear grouped with otherwise dissimilar entities.

According to the full range of relations, all seven private sector businesses, except banks, (entities 2-8) are in Block 9. Almost all of the health services, boards, centers, and associations in grouped together in Block 2. There are five groups (blocks) that are private voluntary associations, clubs, or churches (Blocks 1,4,5,6,10). All three learning institutions are grouped together in Block 3; with TT Savings and Loan (not intuitive.) All three banks are grouped together (Block 11). Note that the distinct behavior of Towertown Savings and Loan, identified when centrality and prestige were studied, also sets the S&L apart from the commercial banks. Government offices are in two types of groups: public service providers are in Block 8, and the largely administrative offices are grouped with the major information providers in Block 7. The two political party committees are grouped together in Block 13; and there are two groups of unions, Blocks 14 and 15.

Since some of the blocks formed by CONCOR are clearly comprised of 'substitute' entities (the two blocks of unions, the five blocks of private voluntary associations), we decided that further aggregation of the fifteen blocks would be appropriate. The minimum number of groups is three: private businesses, voluntary organizations, and public sector institutions. From all previous analyses, since the banks displayed significantly more critical network behavior compared to all other private businesses, we decided to distinguish banks from other businesses. The final aggregation consists of four groups 1) banks, 2) other business, 3) government, and 4) voluntary organizations (**Appendix III**).

We proceed to develop the methodology with this very small number of groups for ease of demonstration. Ideally, given the CONCOR results and our preference to collapse substitutes into blocks, we would analyze a nine block structure. Those blocks would be: (1) businesses, (2) banks, (3) health care, (4) schools, (5) associations (including churches and clubs), (6) public service providers, (7) media and public administration, (8) political parties, and (9) the unions.

Furthermore, we expanded the set of relations (and sociomatrices) from three to four by calculating a sociomatrix we called "SERVICE." SERVICE is the transpose of the MONEY relation. The transpose shows the relation "gives something to" from the row actor to the column actor. The assumption is that all flows of money are reciprocated by something: the provision of a service, delivery of a good, or even a "good feeling" from a charitable act. This is the same convention that records foreign aid given out as an *import* of the donor country in the current account of the donor's Balance of Payments statistics. The SERVICE relation sociomatrix is qualitatively different from the MONEY relation because arc patterns are not symmetric.

3.4.2 *Tie Assignment*

After assigning individuals to blocks, ties must be assigned between blocks. We consider three common approaches to defining ties between subsets (blocks). These are 1) *Zeroblock,* 2) *Oneblock*, and 3) α *density* (Wasserman and Faust, 1994). The first assigns a tie if even a single pair of entities between the two blocks have tie, the second assigns a tie only if all pairs across the two blocks have ties, and the third is a convex combination of the first two. We then propose and use a fourth approach based on our own criteria developed to reflect substitutability between sinks (or sources): the "*sinks substitute*" and "*sources substitute*" criteria.

First, some notation: individual agents in Towertown are indexed by subscripts $i, j = 1, \dots 73$, while groups in the blockmodel are indexed using subscripted $A, B = \{$banks, BIZ, Vol, and Gov$\}$. Let $\boldsymbol{T}$ be any one of the four elemental Towertown sociomatrices of dimension 73^2, so that $\boldsymbol{T} = \{$MONEY, INFORMATION, SUPPORT, AND SERVICE$\}$; and let $\boldsymbol{B}$ denote the corresponding set of four blockmodel (dimension 4^2) adjacency matrices. The blockmodel adjacency matrices are also called **image matrices** (Wasserman and Faust, 1994). Arcs (and ties) in $\boldsymbol{T}$ are denoted $\boldsymbol{t}_{ij}$, while arcs (and ties) between blocks in $\boldsymbol{B}$ are denoted $\boldsymbol{b}_{AB}$.

Zeroblock Criterion: the arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 0 only if there are no arcs ($\boldsymbol{t}_{ij} = 0$) from any actor (i) in the row block to any actor (j) in the column block; otherwise the block arc is 1:

$$\boldsymbol{b}_{AB} = 0 \text{ if } \boldsymbol{t}_{ij} = 0 \text{ for all } i \in A \text{ and all } j \in B \text{ else } \boldsymbol{b}_{AB} = 1.$$

Oneblock Criterion: the arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 1 only if all possible arcs ($\boldsymbol{t}_{ij}$) from all actors in the row block to actors in the column block exist, otherwise the block arc is 0:

$$\boldsymbol{b}_{AB} = 1 \text{ if } \boldsymbol{t}_{ij} = 1 \text{ for all } i \in A \text{ and all } j \in B;\ \text{ else } \boldsymbol{b}_{AB} = 0.$$

$\boldsymbol{\alpha}$ density Criterion: the arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 1 if the observed density of arcs between the two blocks ($\boldsymbol{d}_{AB}$) is at least as large as α, and zero otherwise. Recall that density is the actual number of non-reflexive arcs in proportion to the maximum possible number of non-reflexive arcs. Applied to calculate the density of ties between blocks ($\boldsymbol{d}_{AB}$), the formula is

$$\boldsymbol{d}_{AB} = \left[\sum\sum \tau_i \in A, j \in B\right] / \boldsymbol{N}_A \boldsymbol{N}_B$$

The analyst can choose the reference value of α. One reasonable choice is to set $\alpha =$ the density ($\boldsymbol{D}$) of the original sociomatrix ($\boldsymbol{D}$). If α is chosen to equal $\boldsymbol{D}$: a formal statement of the criterion is:

$$\boldsymbol{b}_{AB} = 1 \text{ if } \boldsymbol{d}_{AB} \geq \boldsymbol{D} \text{ else } \boldsymbol{b}_{AB} = 0$$

Sinks Substitute Criterion: the arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 1 if there is an arc ($\boldsymbol{t}_{ij}$) from every actor in the row block to at least one actor in the column block, otherwise the block tie is 0:

$$\boldsymbol{b}_{AB} = 1 \text{ if for all } i \in A \exists j \in \boldsymbol{B} \text{ s.t. } \boldsymbol{t}_{ij} = 1; \text{ else } \boldsymbol{b}_{AB} = 0$$

Sources Substitute Criterion: the arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 1 if there is an arc ($\boldsymbol{t}_{ij}$) to every actor in the column block from at least one actor in the column block, otherwise the block tie is 0:

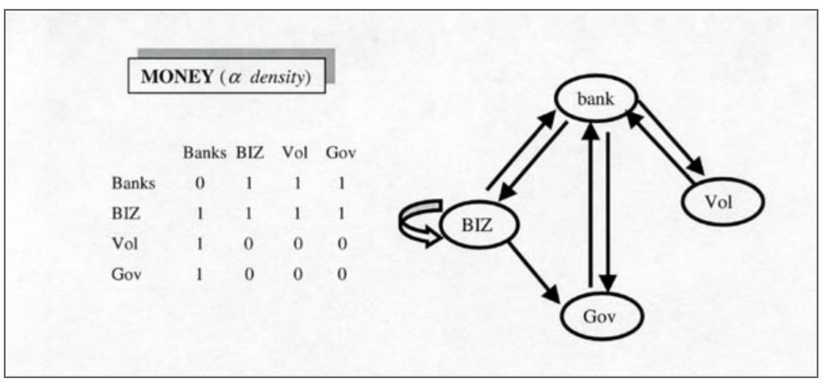
$$\boldsymbol{b}_{AB} = 1 \text{ if for all } j \in B \exists \in A \text{ s.t. } \boldsymbol{t}_{ij} = 1; \text{ else } \boldsymbol{b}_{AB} = 0$$

Which blocking procedure to use should correspond to the specific problem and objective. Our objective is to represent the relations between types of entities in communities. We have grouped together entities in Towertown which can substitute for one another. We know that they are not perfect substitutes. But the zeroblock procedure is appropriate only for the perfect substitute case. If all entities within blocks are perfect substitutes, a tie by any individual in the block could be made by any other individual in the block, so even a single tie would be recorded as a block tie. But our entities are not perfect substitutes. So the zeroblock procedure would be meaningless.

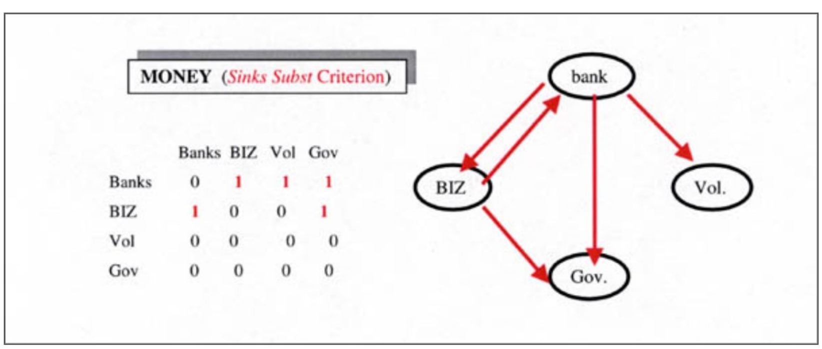
The oneblock procedure may be appropriate when *all* agents within each block are complements for one another. Since each entity is tied in tandem, the whole block should be tied in tandem, or not at all. But we also know that the Towertown entities in the groups are not perfect complements. So the oneblock procedure would not be useful for Towertown data either.

The first two procedures also lead to image matrices that are either all 1's or all 0's. The α density criterion is somewhere between those unlikely extremes, as are the sink/source substitute criteria. In sum, we developed the new criteria for the express use in analyzing blocks comprised of substitutable entities.

We tested all the criteria to assign ties between the four blocks for the MONEY relation. According to the zeroblock criterion, all blocks are reciprocally related (the image matrix, as predicted above, is all 1s.) According to the oneblock criterion, none of the blocks are related (matrix of 0s). According to the α density criterion (setting $\alpha^R = \boldsymbol{D}^R$) the following blockmodel and **reduced graph** is obtained for R=MONEY:

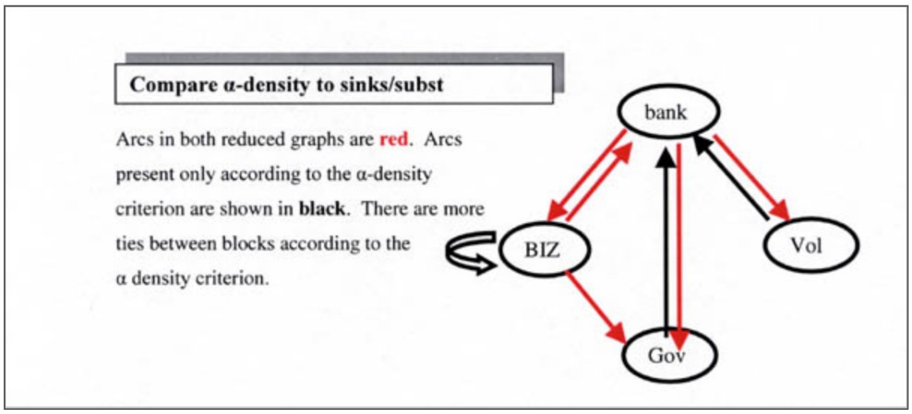


In contrast, applying the **sinks substitute** criterion resulted in the following blockmodel and reduced graph:

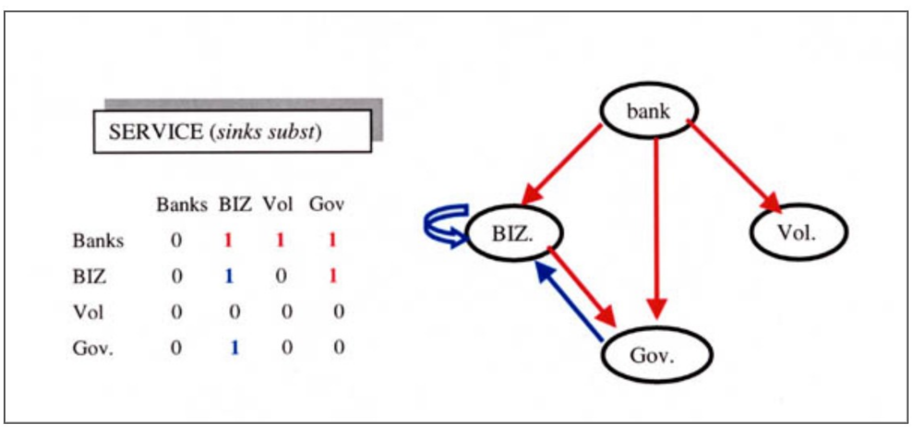


Both of the reduced graphs of the MONEY relation show that banks and business have reciprocated money transactions as lenders, borrowers, depositors, and creditors. Furthermore, businesses and banks give money to government agencies, i.e. pay taxes. Banks give money to voluntary organizations, i.e. donations and/or loans. Note that not enough business supply funds to voluntary associations directly to warrant the recording of a block tie using this criterion. Businesses, however, are linked indirectly to voluntary associations through banks. Banks are clearly articulating agents.

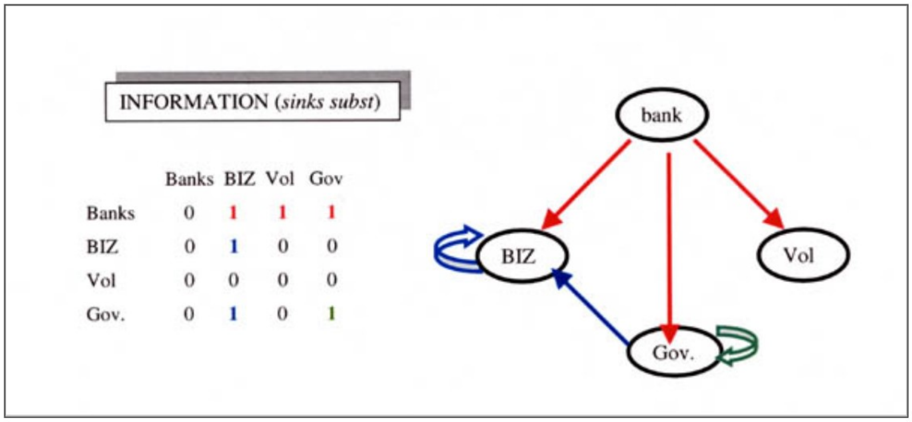
We prefer the sinks substitute criterion because it makes it clear that not all businesses have money transactions with other local businesses in the Towertown data, and that not all government entities nor all voluntary associations give money to local banks. We think the tie assignment using the α-density criterion is too generous when the groupings are supposed to reflect substitutability of entities within the blocks.



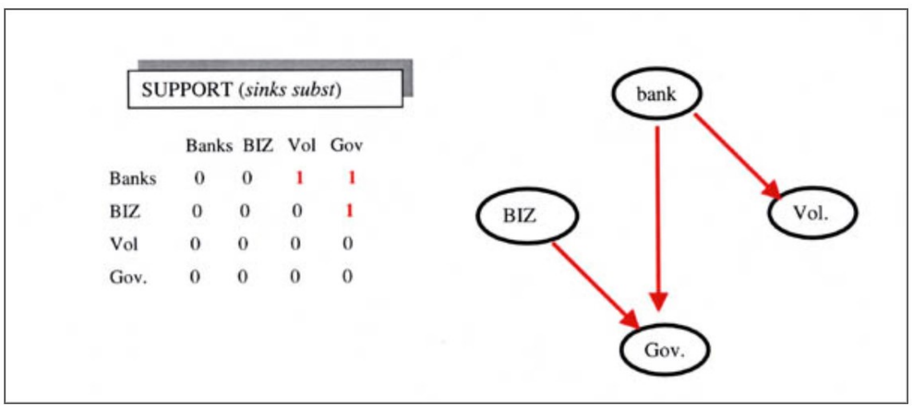
Applying the sinks-substitute criterion to the other three relations resulted in the following blockmodels and reduced graphs. The arcs in **red** are the same in the relation shown as in the MONEY relation. New ties relative to the MONEY relation are shown in **blue**.



The reduced digraph of SERVICE, the transpose of the MONEY relation, clearly shows that SERVICE is not simply a reverse or mirror image of the MONEY relation. In contrast with MONEY, we observe no SERVICE reciprocity between business and banks. This digraph makes it clear that local banks provide services (manage savings, provide credit) to local businesses. The MONEY relation showed that Businesses pay for those services, and that if local Banks buy from businesses, not all are local businesses. This digraph also shows a reciprocal relationship between local businesses and government agencies. The assumptions behind the SERVICE relation include that businesses produce the services paid for by the public sector, and these services also go to businesses. Interindustry transactions are also documented in the SERVICE image matrix. There is a **reflexive tie** (arrow from/to businesses) showing that, according to our sinks substitute criterion, all local businesses in Towertown serve at least one other business in Towertown. Note that banks are shown again to be a critical proper source or transmitter in the system.



The reduced graph for INFORMATION shows in **red** the ties that are the same as in both the MONEY and SERVICE relations, and in **blue** the ties that are the same as in just the SERVICE relation. A unique tie is in **green**. This graph shows that information flows parallel money or service flows. Banks are again the key suppliers of information. All local businesses supply information to at least some other local businesses. And all government agencies provide information to at least some local business, as well as to at least one other public agency.



Finally, the reduced graph of the SUPPORT relation shows that all Banks "would come to the aid" of at least one of the government entities, and, at least one of the local voluntary associations in Towertown. It is not true that all the banks would support at least one local business. Also, all the local businesses would support at least one of the government agencies; but not all local businesses support at least one of the voluntary associations (or banks). Again, banks are significant sources in the Towertown community.

3.5 THE KEYSTONE TEST

In the introduction we posed the following definition: in a community, the **keystone sector** is one that plays a unique role, without which the community is fundamentally and detrimentally altered. To this point we have identified the roles of individual entities, assigned individuals that play similar roles to groups (sectors), and identified the roles of sectors. From the image matrices and the reduced graphs it is clear that all of the sectors play unique roles. Now it is time for the test of sufficiency: which sectors are critical to the network? We present two tests.

For a sector to be *necessary* to maintain the structure of a community, we must show that for every community, community structure will be maintained only if that particular sector is present. If that sector is not present in every community, each community structure would be detrimentally altered. We cannot conduct this test because we have data only on one community. For a sector to be *sufficient* for a community structure, we must show that the sector is responsible for its structure. Sufficiency does not mean that the sector is the only type that can maintain the structure, just that it is one possible sector responsible for it. We can show that a sector maintains a community's structure by contradiction: without that sector, its structure is not maintained. In sum, we can test for the keystoneness of sectors in Towertown by excising the nodes that represent them from the blockmodels and checking if this disrupts the paths between the remaining sectors. The first test considers the effect of excision on the number of components . The second test considers the effect of excision on the geodesics .

3.5.1 *Fracture Test*

Our first test is non-directional: if the excision of a sector fractures the connectivity of the network or breaks up the **component** structure, that sector is a **keystone**. There can be more than one keystone sector according to this criterion. The results of this application of the test for cut-points in the image matrix or the reduced graph, which we call the **fracture test** are summarized in **Table 6**.

Table 6. FRACTURE TEST	*Relation*			
	MONEY	SERVICE	INFO	SUPPORT
excised sector	*number of components*			
none(intact)	1	1	1	1
BIZ	1	1	1	1
Vol	1	1	1	1
Govt	1	1	1	**2**
BANKS	**2**	**2**	**2**	**2**

When the Towertown network is intact, the four sectors form a single component. The polar opposite case would be four components: each sector being a singleton. A completely dissociated community would display the maximum number of components, that is, $\boldsymbol{N}$ which is 4 for the blockmodel of Towertown. In what follows, we excise each sector in turn and count the components. If removing the sector increases the number of components, that sector is a keystone.

If businesses are excised, Towertown remains associated into one component, across all relations. Thus, businesses are not keystones in this data according to this criterion. The scenario corresponds, for an extreme illustration, to how a retirement community is affected by the loss of all private businesses: money, services, information, and support can still flow between the banks, voluntary associations, and public institutions even if there were no local businesses. Businesses are not critical in a community if other sectors can keep the network intact without them.

There is also no fracture of the remaining network if voluntary associations are excised. For example, a town of individualistic or self-sufficient atheists could nonetheless exchange money, service, information, and support without any voluntary associations. The central intermediation role of banks with respect to money, information, *and* support means that community structure will be intact without voluntary associations.

If there are no government entities, not all businesses would remain tied in the SUPPORT network. As indicated in **Table 6**, the excision of GOV would break the single component into 2 with respect to SUPPORT. Thus, public entities, especially the community college (as shown in the section on Micro/individual roles are keystones. The GOVT sector is a keystone with respect to one relation: SUPPORT.

If there were no local banks, however, across ALL relations, the single Towertown network would be fractured into two. With respect to MONEY, SERVICE, INFORMATION, and SUPPORT, without local banks, the voluntary associations would be cut off, breaking Towertown into two components. One component would be the business - government dyad, the other would be the voluntary organization sector singleton.

Recall the digraphs for the blockmodels , above. With respect to the first three of the four relations, the Towertown network is comprised of two sub-components, with banks being the common element in both sub-components. This illustrated that banks provide connectivity to the community. The removal of banks leads to two disconnected components.

In the case of the SUPPORT relation, the blockmodel in the previous section showed that there are three sub-components. Both banks and the government sector play articulating roles in this relation, which is why the excision of either banks or gov't institutions leads to fracturing the single Towertown component into parts.

3.5.2 *Efficient Path Test*

The fracture test is easy to apply graphically. But it focuses on one-step ties rather than complete paths and directional relations (arcs). There is only one 2-step path in the Towertown network. In more complex networks, an analytical alternative to the simple fracture test, which we call the **efficient path**, test may be more appropriate. The efficient path test measures the effect of the excision of a sector on the lengths of the shortest paths between the remaining sectors. If the excision of a sector makes those geodesics longer or (equivalently) reduces the closeness of the remaining entities, the excised sector is a keystone.

Recall that a path is the set of arcs between any two nodes in a relational network. If establishing and maintaining dyadic relations is at all costly, the fewer the arcs (and the shorter the paths) needed for a given number of components, the more efficient the network. For example, $\boldsymbol{N}$ entities could each connect with each other unilaterally, requiring $\boldsymbol{N}(\boldsymbol{N}-1)$ arcs (or one-step paths). If any one entity is excised, the network remains fully-connected. But it is an *inefficient* network. Alternatively, each of the $\boldsymbol{N}$ entities could all belong to an $\boldsymbol{N}$+1st entity, such as an association based on shared values and principles. In that case, the single component (of $\boldsymbol{N}$+1 entities) would require only $\boldsymbol{N}$ arcs (or $\boldsymbol{N}$ one-step paths). Again, the network remains fully connected despite the excision of any one (or even any number of) the original $\boldsymbol{N}$ entities. But if the $\boldsymbol{N}$+1st sector, which is the articulating association, is excised; the network would fracture into $\boldsymbol{N}$ singletons. Comparing the length of paths required to maintain the component structure of a network with/without each sector will indicate the key role the articulating entities play.

In the example above, the $\boldsymbol{N}+1$ association is by far the most central actor. It is also the only articulating actor. Articulating sectors not only provide coherence, they also enhance efficiency. Our **efficient path test** is designed to highlight a sector's contribution to efficiency. It documents any reduction of efficiency in a network following the excision of a sector. The first step in the test is to document the shortest paths (geodesics) between sectors in the intact network. These are the **closeness** measures, but take care to normalize so that comparisons can be made later, because excision will change the size of the network. Then, excise a sector and recompute normalized **closeness** for the remaining sectors and the average for the network. If the excision increases any of the other shortest paths, the average closeness will fall. This suggests that the excised sector had an efficiency-enhancing role, and is a candidate for keystone sector. Again, there may be more than one keystone sector according to this criteria.

For example, with respect to the two networks discussed in the preceding paragraph, for the network with $\boldsymbol{N}(\boldsymbol{N}-1)$ arcs and $\boldsymbol{N}$ sectors, normalized closeness is 1 for each entity (and thus on average). For the network with $\boldsymbol{N}$ arcs and the $\boldsymbol{N}$+1 articulating sectors, closeness is $\boldsymbol{N}/(2\boldsymbol{N}-1)$, which is less than 1. The first network is more close, the second is more efficient. If any single entity is excised from the first network, the normalized closeness for each other entity is $[(\boldsymbol{N}-2)/(\boldsymbol{N}-2)]^{-1}$. It is still equal to 1, which indicates that no entity is key. If the articulating agent is excised from the second network, closeness becomes undefined $(1/\infty)$, since the remaining entities do not relate anymore. The method is a numerical way which will show if indeed an articulating agent is a keystone. If there is more than one articulating sectors, excising one will not affect the geodesics among the remaining sectors.

The procedure was applied to the asymmetric Towertown blockmodel for the MONEY relation. The results are shown in **Table 7**. The closeness of BANKS is highest ($\boldsymbol{C}_{bank}=1$) in the intact blockmodel , and the average closeness in the network is 0.875. Note that since neither GOVT and VOLuntary organizations are MONEY sources, there are no arcs originating from them, so their closeness measures are undefined. Removing BANKS isolates the VOLuntary group from businesses. The distance becomes infinite and the network closeness becomes undefined ($\rightarrow 0$). Removing GOVT also increases the distances, and reduces average network closeness from 0.876 to 0.83. When BIZ or VOL are excised, closeness for the remaining entities increases from 0.875 to 1, because those two entities have the longest path between them (among all relating dyads).

In sum, the efficient path test shows that the Towertown network is sensitive to the excision of businesses and government, but not as sensitive as it is to excision of BANKS. BANKS are the most key to MONEY efficiency.

Table 7. EFFICIENT PATH TEST	*excised sector:*				
closeness measure, MONEY relation	none (intact)	**BANKS**	BIZ	VOL	GOV'T
BANKS	1		1	1	1
BIZ	.75	1/∞		1	.66
normalized average	.875	**→ 0**	1	1	.83

The efficient path test is not applied to the other relations, because INFORMATION, SUPPORT, and SERVICE relations are not necessarily **transitive**. The MONEY relation is transitive because money is a generic unit of account and medium of exchange. It is reasonable to assume that when an entity receives money from one agent and then gives money to another, it is effectively the same money (but not necessarily the same amount). This is not true with respect to information. For example, Sally may talk to Henry, and Henry may talk to Steve, but not necessarily about the same things. Information is not generic. Neither is support or service. The Towertown data on dyadic relations do not include enough information to claim that along a 2-or more step path of SUPPORT, or SERVICE that the same thing is flowing along the entire path.

4 SUMMARY

This report has presented a new a method for identifying keystone sectors in communities, where *sectors* are broadly defined to include churches, clubs, associations, and public institutions as well as different types of businesses and industries. In an arch, the keystone is the one with the unique shape at the top of the arch that is critical for the arch's structural stability. The term keystone species was first coined by ecologists in the late 1960s with respect to the species responsibility for the structure and integrity of an ecosystem. In this paper, we developed a new method for identifying a keystone entity in any kind of system, but explicitly for community development analysis. We proposed that a keystone sector in a community plays a unique role, without which the community is fundamentally and detrimentally altered.

4.1 A Replicable Keystone Sector ID Method

This report has described and applied new ways to (1) quantify the structure, (2) describe the variety of roles, and (3) identify the keystone sectors in a community. We adapted existing methods of graph theory and social network analysis to the analysis of a local economic system. We do not know of any precedents for community-level economic analysis.

The *Keystone Sector ID* method can be easily replicated. First, we recommend reading the book by Wasserman and Faust (1994) for a general foundation in social network analysis. Next, given a community to study, and the relations to be investigated (money, information, ...), assay/list all the entities in the community and decide which ones to interview as subjects. Randomly sample, or include all entities larger than a threshold minimum size, which is what the Towertown surveyors did. Present each subject with a list of the other subjects. Ask them to indicate the others *to whom* the subject is a source, for each relation. Then, ask *from whom* the subject is a sink. Subjects could even be asked to record their ties in vector form ("In the first column, place a "1" next to each entity on this list *to whom* you **give** information about community affairs. In the second column, place a "1" next to each entity on this list *from whom* you **get** information about community affairs").

Summarize the subjects' responses into a single matrix for each relation. The dichotomous sociomatrices should indicate all the ties reported by all the subjects. Note that in the Towertown data, an arc is recorded as long as either side of the dyad reports it.

Analyze the sociomatrix data using a spreadsheet program or network analysis software. The basic measures of in- and out-degrees are simple to calculate. Density and most of the other measures can also be calculated using LOTUS 1-2-3. We recommend 1-2-3 over Microsoft EXCEL because LOTUS 1-2-3 offers a very easy to use matrix multiplication command. Or, use a social network analysis program like UCINET. Determine the numbers of components, cliques, or subgroups. Identify major sources and major sinks. Identify transmitters, brokers, articulators. Aggregate the elemental data into sectors using UCINET's CONCOR algorithm. Assign ties using the sinks substitute or sources substitute criteria (introduced in this paper); generate the blockmodels and reduced digraphs. Apply the fracture and/or efficient path tests introduced in this paper to identify the keystone sector *s*. (Obviously, those tests are not available on UCINET, since they are new to this project. In this work, we used LOTUS 1-2-3 to assign blockmodel ties and to conduct the fracture and efficient path tests.)

4.2 Strengths and Weaknesses

We have also raised many caveats. Dichotomous data treats all interactions as if they are of equal significance. This assumption is acceptable for non-rival interactions, but it is less appropriate for rivalrous transactions, such as those concerning money. It is possible that a sector with a few large links may be much more critical than the sectors with many (smaller) links. A methodology based on dichotomous data may fail to identify large sectors as key.

Also, the articulatory roles of some entities are exaggerated when analyses are based on measures for which symmetry is imposed. When we ignore the direction of a relation, we incorrectly assume that an exchange is reciprocated, when it may not be. Many network analysis software packages impose symmetry before calculating measures, *caveat emptor* ("buyer beware").

Other potential pitfalls or misleading results may arise from the blocking and tie assignment procedures. The patterns of ties to- and from- a block depend on the entities in the block, so the blocking procedure matters. We also showed that different criteria for tie assignment gives a different blockmodel. In the Towertown application, it turned out that the results of our tests for 'keystoneness' are the same regardless of the procedure used for tie assignment. The α-density network is just as sensitive to the excision of banks as our sinks substitute networks are. But Towertown is just one observation, it need not generally be the case. With other data, the results of the fracture and efficient path tests for keystone-ness may be different across blockmodels developed using different criteria.

By replicating these methods and analyzing the network structure of other communities we may learn how sensitive the *Keystone Sector ID* method is to these potential pitfalls. Meanwhile, it is clear that the approach can offer significantly new types of insight. The finding that financial intermediators play a critical role in a community is an example of one of those insights. Previous studies in the critical sector tradition, based on material input-output relationships, have ignored the banking sector (along with government entities and voluntary associations). This is because input-output data concerns only current account, private business transactions. The *Keystone ID* method can consider not only business-to-business current account transactions, but also capital account transactions, information flows, political support, and service transactions not only among businesses but with and among non-business entities as well. In some communities, public service providers, private voluntary associations, or other non-business entities may be key.

The lesson is that many types of interactions and many types of entities may be important to a community, not just businesses and money interactions. Here we have demonstrated a method that can be used to quantify and analyze almost any kind of interaction information between any kinds of entities.

4.3 Results and Implications

The outcome of all the analyses of the Towertown community data is the robust finding that *banks are a keystone sector*. The banking sector appears to be key according to all the various network analytic measures of centrality, prestige, and connectivity. At the individual level, banks were identified as the cut-points, and their excision isolated various voluntary organizations from the network, without disrupting the other ties. Furthermore, at the blockmodel level, banks were shown to play a critical role as a group. The Towertown network is very sensitive to the excision of banks. Their excision from the blockmodel, as shown by the fracture and efficient path tests, left voluntary associations isolated.

Now that we have a way to identify the Keystone Sector in one community, how might we use this information to enhance entrepreneurial/business activity or to support innovative community change? In our opinion, *we must first apply the Keystone Sector ID procedures to many small communities before concluding that any one sector (such as banks) is 'the keystone' in small communities* in general. The testable (alternative) hypotheses include "communities that do not have a bank or S&L are not fully-connected."

Furthermore, communities that are not fully-connected may have (hypothetically) fewer viable voluntary associations. Can we say that disconnected network communities have lower levels of "social capital?" We also expect to find that property values in communities that are not attractive to people or businesses do not rise. In contrast, we expect property values in fully-connected communities with more than one individual actor in their keystone sector (a competitive rather than monopolized keystone sector) to rise. We need to replicate the *Keystone Sector ID* study many times on different rural communities before we can draw inferences about what is the keystone sector in general. Then we should proceed to test for associations between community development outcomes and network patterns.

In closing, we hope that this work inspires other researchers to apply a *Keystone Sector ID* method to analyze other communities. We look forward to collaborating on ways to (1) compile single community network studies into a sample, so that we can draw inferences from a sample of communities, and (2) test for the dependence of rural development outcomes on network structure.

APPENDIX I: Entities in "Towertown"

1 - Farm Bureau
2 - Farm Equip. Co.
3 - Clothing Mfg. Co.
4 - Farm Supply Co.
5 - Mechanical Co.
6 - Electric Equip. Co.
7 - Metal Products Co.
8 - Music Equip. Co.
9 - Chamber of Commerce
10 - Banker's Association
11 - 1st Towertown Bank
12 - Towertown Savings and Loan
13 - Bank of Towertown
14 - 2nd Towertown Bank
15 - Brinkman Law Firm
16 - Cater Law Firm
17 - Lenhart Law Firm
18 - County Bar Association
19 - Towertown Board of Realtors
20 - Towertown Small Bs. Association
21 - Municipal Employees Union 1
22 - Municipal Employees Union 2
23 - Teacher's Union
24 - Central Labor Union
25 - City Council
27 - County Board
28 - Fire Department
29 - Human Relations Commission
30 - Mayor's Office
31 - Police Department
32 - Sanitary District
33 - Streets and Sanitation
34 - Park District
35 - Zoning Board
36 - Democratic Committee
37 - Republican Committee
38 - League of Women Voters
39 - The Towertown Newspaper
40 - WTWR Radio Station
41 - Towertown Public Hospital Board
42 - Towertown Public Hospital
43 - County Medical Society
44 - County Board of Mental Health
45 - County Board of Health
46 - County Health Service Center
47 - State Highway Authority
48 - Kiwanis Club 1
49 - Kiwanis Club 2
50 - Rotary Club
51 - Lions Club
52 - United Fund
53 - School Board
54 - Towertown High School
55 - Towertown Parent-Teacher Association
56 - Towertown Community College
57 - State University
58 - Association of Churches 1
59 - Association of Churches 2
60 - St. Hilary's Catholic Church
61 - 1st Baptist Church
62 - 1st Church of the Light
63 - 1st Congregational Church
64 - 1st Methodist Church
65 - Unity Lutheran
66 - University Methodist Church
67 - State Department of Public Aid
68 - County Housing Authority
69 - Family Services
70 - State Employment Services
71 - YMCA
72 - Mental Health Center
73 - Towertown Youth Services Bureau

APPENDIX II: CONCOR results

Block1: Farm Bureau, St. Hilary's Catholic Church, University Methodist Church, Association of Churches 1, 1st Congregational Church and 1st Methodist Church

Block 2: County Medical Society, County Board of Mental Health, TT Public Hospital Board, County Housing Authority, TT Public Hospital, State Employment Services, State Department of Public Aid, Mental Health Center, and County Board of health.

Block 3: TT Community College, State University, TT High School, and **TT Savings and loan**

Block 4: YMCA, United Fund, County Health Service Center, Family Services, and **Brinkman Law Firm**

Block 5: 1st Church of the Light, League of Women Voters, 1st Baptist Church, School Board, TT Small Business Association, Bankers Association, and Chamber of Commerce

Block 6: Kiwanis Club 1, Kiwanis Club 2, Rotary Club, Lions Club, TT Parent Teacher Association, and **Park District**

Block 7: WTWR Radio Station, Association of Churches 2, City Manager's Office, TT Newspaper, City Council, County Bar Association and **Cater Law Firm**

Block 8: Human Relations Commission, Fire Department, Police Department, Sanitary District, Mayor's Office, County Board, and Zoning Board

Block 9: Farm Equip Co., Clothing Mfg. Co., Farm Supply Co, Mechanical Co, Electric Equip Co, Metal Products Co and Music Equip Co

Block 10: Unity Lutheran Church

Block 11: 1st TT Bank, 2nd TT Bank , and Bank of TT

Block 12: Streets and Sanitation, State Highway Authority, TT Board of Realtors, and **Lenhart Law Firm**

Block 13: Democratic and Republican Committees

Block 14: Municipal Employees Union 1 and Teachers' Union

Block 15: Municipal Employees Union 2 and Central Labor Union

APPENDIX III: BLOCKMODEL SECTORS

Sector 1: **BANKS** =

Block 11: 1ST TT Bank, 2ND TT Bank , and Bank of TT and TT Savings and loan

Sector 2: **BUSINESSES** =

Block 9: Farm Equip Co., Clothing Mfg. Co, Farm Supply Co, Mechanical Co, Electric Equip Co, Metal Products Co and Music Equip Co

+ not from intact CONCOR blocks: Lenhart Law Firm, Brinkman Law Firm, Cater Law Firm, WTWR Radio Station, TT Newspaper **

Sector 3: **GOVERNMENT AGENCIES** =

Block 2: County Medical Society, County Board of Mental Health, TT Public Hospital Board, County Housing Authority, TT Public Hospital, State Employment Services, State Department of Public Aid, Mental Health Center, and County Board of health +

Block 3: TT Community College, State University, TT High School +

Block 8: Human Relations Commission, Fire Department, Police Department, Sanitary District, Mayor's Office, County Board, and Zoning Board

+ not from intact CONCOR blocks: County Health Service Center, Family Services, Streets and Sanitation, State Highway Authority, City Manager's Office, City Council, and Park District**

Sector 4: **VOLUNTARY ASSOCIATIONS** =

Block1: Farm Bureau, St. Hilary's Catholic Church, University Methodist Church, Association of Churches 1, 1st Congregational Church and 1st Methodist Church +

Block 5: 1ST Church of the Light, League of Women Voters, 1st Baptist Church, School Board, TT Small Business Association, Bankers Association, and Chamber of Commerce +

Block 6: Kiwanis Club 1, Kiwanis Club 2, Rotary Club, Lions Club, TT Parent Teacher Association +

Block10: Unity Lutheran Church +

Block 13: Democratic and Republican Committees +

Block 14: Municipal Employees Union 1 and Teachers' Union +

Block 15: Municipal Employees Union 2 and Central Labor Union

+ Not in an intact CONCOR block: YMCA, United Fund, Association of Churches 2, County Bar Association, TT Board of Realtors**

*** CONCOR groups that were split: Blocks 4, 7, and 12*

GLOSSARY

Adjacency is the graph theoretic expression of the fact that two entities, represented by nodes, are directly related, tied, or connected with one another. Formally, given entities n_i and n_j in a set of agents $\boldsymbol{N}$, and the $\boldsymbol{A} = \{a_{ij}\}$ arcs denoting the existence of relations from i to j; i and j are *adjacent* if there exist either of the two arcs, a_{ij} or a_{ji}. Given the digraph $\boldsymbol{D} = (\boldsymbol{N}, \boldsymbol{A})$, its **adjacency matrix** $\boldsymbol{A(D)}$ is defined by $\boldsymbol{A(D)} = \{a_{ij}\}$ where $a_{ij} = 1$ if either a_{ij} or a_{ji} exists, and 0 otherwise.

$\boldsymbol{\alpha}$ density Criterion: a procedure for tie assignment in the construction of a blockmodel. According to this criterion, an arc ($\boldsymbol{b}_{AB}$) between two blocks (A, B) for a given relation is 1 if the observed density of arcs between the two blocks ($\boldsymbol{d}_{AB}$) is at least as large as $\boldsymbol{\alpha}$, and zero otherwise. Often, α is chosen to equal ($\boldsymbol{D}$), the density of the original sociomatrix. Formally, for $\boldsymbol{\alpha} = \boldsymbol{D}$; $\boldsymbol{b}_{AB} = 1$ if $\boldsymbol{d}_{AB} \geq \boldsymbol{D}$ else $\boldsymbol{b}_{AB} = 0$.

arc: in graph theory, an arc is a directional relation that has an explicit origin and destination, represented by a line between entities with an arrowhead at the destination.

backward linkage = transaction link(s) *to* the buying sector of interest *from* industries that supply intermediate goods or services. The supplying industries are also known as 'upstream' sectors in the flow of goods from extraction through raw, processed, and final retailed form to consumers. The magnitude of a sector's backward linkages are given by its column coefficients in the multiplier matrix.

betweenness, a measure of the *global centrality* of the actor i, is the probability that the shortest path from actor j to actor k takes a route through agent i. Formally, $\boldsymbol{B}(n_i) = \sum_j^N \boldsymbol{g}_{jk}(n_i)/\boldsymbol{g}_{jk}$ where $\boldsymbol{g}_{jk}(n_i)$ is the total number of geodesics through i, and $1/\boldsymbol{g}_{jk}$ is the probability that the particular geodesic is chosen.

blockmodel is an aggregated sociomatrix representation of a network. A blockmodel is created from an elemental sociomatrix data base by (i) partitioning the entities into discrete subsets, called "blocking," and (ii) assigning 0/1 ties between each pair of blocked subsets (White, Boorman and Breiger, 1976; Holland and Leinhardt, 1979).

closeness is a second measure of the *global centrality* of the actor i. $\boldsymbol{C}(n_i) = [\sum_j^N d(n_i, n_j)]^{-1}$ where $d(n_i, n_j)$ is the geodesic (shortest path) distance between i and j entities in the network, and $\boldsymbol{N}$ is the network size. To normalize, divide the sum of geodesics by $\boldsymbol{N} - 1$ before inverting.

complement: generally, an input or consumer good whose use decreases as the cost of another rises is considered a *complement in demand* to that other good. An output or product whose supply decreases as the price of another falls is considered a *complement in supply* to the other product. Complements generally occur in fixed proportions in an economy. For **Keystone Sector Analysis:** *entities I and J are perfect complements when K's interaction with I is always accompanied by K's interaction with J* (Kilkenny and Nalbarte-this paper).

A **complete** graph is one in which all the actors have two-way ties to all other actors.

A **component** is the largest subset of actors in a network that all relate to each other, also known as *group* or a *sub-graph.* A **strong component** is one in which the arcs that make up the paths are aligned in continuous chain without a change of direction. A **weak component** is made of actors that are linked by non-directional edges (Scott, 1991).

cut-point is the node whose removal from the system would increase the number of components by dividing the graph into 2 or more separate components, between which there are no ties (Scott, 1991).

density is the measure of how many entities are related to others in a set. Density ($\boldsymbol{D}$) is measured by the ratio of the actual number of non-reflexive arcs in proportion to the maximum possible number of non-reflexive arcs: $\boldsymbol{D} = \sum_i \sum_j a_{\boldsymbol{ij}}/\boldsymbol{N}(\boldsymbol{N}-1)$.

dichotomous relation is a tie between two agents (dyad) that either does or does not exist. It is recorded as a binary variable with (1) indicating the presence and (0) indicating the absence of the relation between the two set entities in the dyad.

directed graphs or **digraphs** are the graphic representations of directional relational data among entities in a set. Entities are illustrated as nodes, and the directional relations, if they exist, are illustrated as arcs, with the arrowhead pointing from the source or sending node to the destination or receiving node. Formally, a **digraph** is a finite, non-empty set $\boldsymbol{N}$, whose elements $n_i = \{n_1, n_2, \ldots n_g\}$ are called *nodes*, together with a set $\boldsymbol{A} = \{a_{12}, a_{13}, \ldots a_{1g}, \ldots a_{g-1,g}\}$ of ordered pairs a_{ij}, called arcs, where n_i and n_j are distinct members of $\boldsymbol{N}$ (Robinson and Foulds, 1980).

directional relation is an interaction between entities that is specific about which entity is the source versus the sink, which is the origin or destination. It is represented graphically by an arc, a line between entities with an arrowhead at the destination. The sociomatrix for a directional relation will not generally be symmetric, unless all ties are reciprocated.

dominant industry: the sector(s) with relatively high location quotients, that also have a number of input-output linkages with other local industries with high location quotients Cella 1984

A **dyad** consists of a pair of actors and the possible ties between them (Wasserman and Faust, 1994).

edge: in graph theory, this represents a non-directional relation or tie which is non-specific about the origin or destination of the flow on the link. It is illustrated by a line between the interacting agents that has no arrowhead.

efficient path test is applied to a blockmodel of a community to determine if a sector is a keystone. If the excision of the sector from the image matrix (or reduced graph reduces the **closeness** measures for the remaining sectors, that sector is a **keystone**. (Kilkenny and Nalbarte, this paper)

forward linkage the transaction link(s) *from* the supplying sector of interest *to* industries that demand intermediate goods or services. The using industries are also known as 'downstream' sectors in the flow of goods from extraction through raw, processed, and final retailed form to consumers. The magnitude of a sector's forward linkages are given by its row coefficients in the multiplier matrix.

fracture test is a cut-point test based on ties applied to a blockmodel of a community to determine if a sector is a keystone. If the excision of the sector from the image matrix (or reduced graph) destroys the connectivity of the network by increasing the number of **components**, that sector is a **keystone**. (Kilkenny and Nalbarte, this paper)

geodesic $d(n_i, n_j)$ is the shortest path between two nodes i and j. It is measured as the number of arcs required to get from i to j, which is the first power p for which the ijth element of $\boldsymbol{A}^p$ is non-zero: $d(n_i, n_j) = \min\{p\}|\boldsymbol{A}_{ij}^P > 0$ (Wasserman and Faust, 1994). Equivalently, for any digraph with adjacency matrix $\boldsymbol{A}$, each cell a_{ij}^p of $\boldsymbol{A}^p$ equals the number of paths of length p from node i to j in A, for any positive integer p (Robinson and Foulds, 1980).

image matrix is the aggregated form of an elemental sociomatrix, representing the blockmodel in a matrix comprised of [0,1] cell entries.

immiserizing growth is expansion in a region's real economic output that earns less rather than more real income. This outcome can arise only under two conditions (1) the region must be large in the market for the product, where 'large' means that changes in the region's supply of the product will affect market prices, and (2) demand for the product must be highly inelastic, such that revenues fall even as more is sold.

indegree measures the strength of a node as a sink in a system. It is the number of arcs ending at the node, measured by the column sum for the node in a dichotomous sociomatrix: formally the indegree of actor $j = \sum_i a_{ij}$.

keystone sector is the type of entity (business, institution, organization, etc) in a community that plays a unique role and without which the community is fundamentally and detrimentally altered (Kilkenny, 1997). Based on the terms Keystone Species and turnkey or key sector.

keystone species is the species responsible for the structure and integrity of an ecosystem. The term was first coined by ecologists in the late 1960s (Paine, 1969).

local centrality is a measure of prominence which reflects the number of direct transmissions from the entity, measured by the outdegree (or row sum) for the entity. Also known as *degree centrality.*

local prestige is a measure of prominence which reflects the number of the entity's direct receipts, measured by the indegree (or column sum) of the entity. Also known as *degree prestige.* (Wasserman and Faust, 1994).

location quotient: a measure of relative concentration in an area. Given a local proportion, such as the percentage of total region r employment in industry $i(e_{ir})$, and a reference proportion, such as the share of nationwide employment in industry $i(e_{ir})$, the location quotient (LQ_{ir}) is e_{ir}/e_i, more formally: $LQ_{ir} = (e_{ir}/\sum_i e_{ir})/(\sum_r e_{ir}/(\sum_i \sum_r e_{ir})$.

multiplier matrix is an array of numbers which show the amount change in the row sector due to a unit change in the column sector. It is calculated from an industry-by-industry transactions matrix in three steps as follows. First, divide the transaction cells by their column totals. This gives a matrix of sectoral expenditure shares $[A]$. Second, subtract this matrix from the Identity matrix $[I - A]$. Third, invert this subtrahend. The result is the multiplier matrix "m". Formally, $m = [I - A]^{-1}$.

nodes represent the individual entities or actors in networks.

non-directional relation is an interaction between entities that is non-specific about which entity is the source versus the sink, or which is the origin or destination. It is represented graphically by an edge or tie, a simple line between the entities, and by a symmetric sociomatrix.

Oneblock Criterion: a procedure for tie assignment in the construction of a blockmodel. According to this criterion, an arc $(\boldsymbol{b}_{AB})$ between two blocks (A, B) for a given relation is 1 only if all possible arcs $(\boldsymbol{t}_{ij})$ from all actors in the row block to actors in the column block exist, otherwise the block arc is 0. Formally, $\boldsymbol{b}_{AB} = 1$ if $\boldsymbol{t}_{ij} = 1$ for all $i \in A$ and all $j \in B$; else $\boldsymbol{b}_{AB} = 0$.

Outdegree measures the strength of a node as a source in a system. It is the number of arcs beginning at a node, measured in dichotomous sociomatrix data as the row sum for the node: outdegree of actor $i = \sum_j a_{ij}$

reduced digraph the illustration of arcs between the blocks of nodes for a blockmodel. See also image matrix.

reflexive tie is the relation that a particular entity has with itself. In a sociomatrix, a reflexive tie by the ith entity is recorded by $a_{ii} = 1$. In a digraph, reflexive ties are drawn as arrows that originate and end on the same node, that are curved back on themselves.

A **relation** is the collection of ties of a specific kind among a set of entities. Alternatively, consider the mathematical definition of **binary relation** (Robinson and Foulds,1980):

> Given two sets $\boldsymbol{S}$ and $\boldsymbol{T}$, each member of set $\boldsymbol{S}$ may be related to a number (perhaps zero) of members of set $\boldsymbol{T}$. The mathematical description of this situation is called a binary relation. If $s \in \boldsymbol{S}$ and $t \in \boldsymbol{T}$, then (s, t) is a member of this set when s **is related to** t.

Sinks Substitute Criterion: a procedure for tie assignment in the construction of a blockmodel. According to this criterion, an arc $(\boldsymbol{b}_{AB})$ between two blocks (A, B) for a given relation is 1 if there is an arc $(\boldsymbol{t}_{ij})$ from every actor in the row block to at least one actor in the column block, otherwise the block tie is 0. Formally, $\boldsymbol{b}_{AB} = 1$ if for all $i \in A \exists j \in B$ s.t. $t_{ij} = 1$; else $\boldsymbol{b}_{AB} = 0$ (Kilkenny and Nalbarte, this paper)

social embeddedness in traditional societies, economic life is submerged ('embedded') in social relations. For example, businesses employ locals and buy locally, no matter what, because it is expected of them by society. In contrast in modern life, social relations are often an epiphenomenon of the market (Granovetter, 1985). For example, when businesses employ locals or buy locally simply because it is profitable, this is the reverse of *social embeddedness.*

A **social network** consists of a finite set of *actors* and the *relation* or relations defined on them. **Actors** are social entities, discrete individuals, corporate or collective social units. (Wasserman and Faust, 1994).

sociomatrix is an entity-by-entity array of data on the relational ties between them. Rows of the sociomatrix represent the sending actors while the columns represent the receiving actors. See also digraph.

Sources Substitute Criterion: a procedure for tie assignment in the construction of a blockmodel. According to this criterion, an arc $(\boldsymbol{b}_{AB})$ between two blocks (A, B) for a given relation is 1 if there is an arc

($\boldsymbol{t}_{ij}$) to every actor in the column block from at least one actor in the column block, otherwise the tie is 0. Formally: $\boldsymbol{b}_{AB} = 1$ if for all $j \in B \exists i \in A$ s.t. $\boldsymbol{t}_{ij} = 1$; else $\boldsymbol{b}_{AB} = 0$. (Kilkenny and Nalbarte, this paper)

structurally equivalent entities have exactly the same directional ties (arcs) to and from all other entities with whom either has ties. Formally, entities i and j are *structurally equivalent* if, for all other actors $k = 1, 2, \dots \boldsymbol{N}; k \neq i, j$, and all the relations $r = 1, 2, \dots, \boldsymbol{R}$, $a^r_{ik} \iff a^r_{jk}$ and $a^r_{ki} \iff a^r_{kj}$.

substitute: generally, an input or consumer good whose use increases as the cost of an alternative rises is considered a *substitute in demand* for that alternative. An output or product whose supply increases as the price of an alternative falls is considered a *substitute in supply* to the alternative product. Substitutes generally satisfy similar demands (play similar roles) in an economy. For **Keystone Sector Analysis:** *entities I and J are perfect substitutes if a slight decrease in the desirability of K interacting with I leads to K interacting with J instead (*Kilkenny and Nalbarte, this paper*).*

tie is a relation between two entities in a network. If it has direction it is an arc, if non-directional it is an edge.

transitive a relation R is *transitive* over the set $\{\boldsymbol{X}\}$ if for all $x_{i,j \text{ or } k}$ in $\{X\}$, if $x_i R x_j$ AND $x_j R x_k$, then $x_i R x_k$. For example, let R be "like" and $\{X\}$ be people. Like is a transitive relation if every time a person A likes B, and person B likes C, then person A likes C. It is easy to find a counterexample to prove that in fact, "like" is **not** a transitive relation for people.

turnkey or key sector sectors whose structure of backward and forward linkages create above-average impacts on the rest of the economy. Activities having the highest linkages are considered key sectors because it is thought that concentrating resources in them will stimulate more production, income and employment than alternative allocations of resources. (Cella, 1984). See also dominant industry.

valued relation is a tie that has a magnitude that reflects the level, intensity, or frequency of a relation.

Zeroblock Criterion: a procedure for tie assignment in the construction of a blockmodel. According to this criterion, a tie ($\boldsymbol{b}_{AB}$) is assigned between two blocks (A, B) for a given relation is 0 only if there are no arcs ($\boldsymbol{t}_{ij} = 0$) from any actor (i) in the row block to any actor (j) in the column block; otherwise the block arc is 1. Formally, $\boldsymbol{b}_{AB} = 0$ if $\boldsymbol{t}_{ij} = 0$ for all $i \in A$ and all $j \in B$ else $\boldsymbol{b}_{AB} = 1$.

BIBLIOGRAPHY

Berge, C. (1962). *The Theory of Graphs.* New York: Wiley.

Campbell, J. (1975). Application of graph theoretic analysis to interindustry relationships: The example of Washington state. *Regional Science and Urban Economics*, 5:91-106.

Cella, G. (1984). The input-output measurement of interindustry linkages. *Oxford Bulletin of Economics and Statistics*, 46(1):73-84.

Fienberg, S.E., Meyer, M.M and Wasserman, S. (1981). Analysing data from multivariate directed graphs: an application to social networks. in *Interpreting Multivariate Data*, ed. V. Barnett, London: John Wiley. 289-306.

Fienberg, S.E. Meyer, M.M. and Wasserman, S. (1985). Statistical Analysis of multiple sociometric relations. *Journal of the American Statistical Association*, 80(389):51-67.

Freeman, L. (1977). A set of Measures of Centrality Based on Betweenness. *Sociometry* (1):35-41.

_______________(2000). Visualizing Social Networks. *Journal of Social Structure*, (1):1.

Fukuyama, F. (1995). *Trust :The Social Virtues and the Creation of Prosperity.* New York: The Free Press.

Galaskiewicz, J. and Mardsen, P.V. (1978). Interorganizational resource networks: formal patterns of overlap. *Social Science Research*, 7:89-107.

Goode, Frank, and Steven Hastings (1988). *Northeast Industrial Targeting (NIT) and Economic Development Data Base (EDD) System User's Manual* version 1, Dept. of Agricultural Economics and Rural Sociology, The Pennsylvania State University, University Park, PA 16802; March.

Granovetter, M. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, 78(6):1360-1380.

____________________(1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology*, 91(3):481-510.

Hanson, S. and Huff, J. (1986). Classification Issues in the Analysis of Complex Travel Behavior. *Transportation*, 13(3):271-293.

Holland, P.W., and Leinhardt, S. (1979). *Perspectives on Social Network Analysis.* Volume of Quantitative Studies in Social Relations. New York: Academic Press.

Jacobs, Jane (1984). *Cities and the Wealth of Nations: Principles of Economic Life.* New York: Random House.

Kauffman, S.A. (1988). The evolution of economics webs. in P.W.Anderson, K.J. Arrow and D.Pines (eds). *The Economy as a Complex Evolving System.* New York: Addison-Wesley: 125-146.

Kilkenny, Maureen (1997). Identifying Rural Keystones *research project proposal to the Tennessee Valley Authority (TVA) Rural Studies Program.*

Kilkenny, M., Nalbarte, L. and Besser, T. (1999). Reciprocated Community Support and Small-Town, Small-Business Success. *Entrepreneurship and Regional Development*, 11:231-246 (August).

Knack, S. and Keefer, P. (1997). Does Social Capital have economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 112(4):1251-1288.

Koppelman, F.S. and Pas, E. (1985). Travel Activity Behavior in Time Space: Methods for Representation and Analysis pp.587-627 in *Measuring the Unmeasurable* P.Nijkamp, H. Leitner, and N. Wriley, editors; Netherlands, Doedrecht: Martinus Nijhoff Publishers.

Kranton, R. (1996). Reciprocal Exchange: A self-sustaining System. *The American Economic Review*, 86(4):830-51.

Lauman, E.O. (1985). Interorganizational Resource Links in Towertown, U.S.A., from *A collection of problem from many Fields for the Student and Research Worker* by Andrews and Herzber, New York: Springer-Verlag.

Paine, Robert (1969). A Note on Trophic Complexity and Community Stability. *American Naturalist* 103:91-93; cited in Mills, Scott, M. Soule, and Daniel Doak (1993). The Keystone Species Concept in Ecology and Conservation. *BioScience*, 43(4):219-224.

Putnam, R. (with R. Leonardi and R.Nanetti) (1993). *Making Democracy Work: civic traditions in modern Italy.* Princeton New Jersey: Princeton University Press.

Robinson, D. F. and Foulds, L.R. (1980). *Digraphs: Theory and Techniques.* New York: Gordon and Breach.

Roy, J. R. (1994). Trade with and without intermediaries: some alternative model formulations. *Annals of Regional Science*, 28:329-344.

Roy, J. R. (1995). Dispersed spatial input demand functions. *Annals of Regional Science*, 29:329-334.

Scott, J., (1991). *Social Network Analysis.* London: SAGE Publications Ltd.

Sonis, M., Hewings, G. (1997). Economic complexity as network complication: multiregional input-output structural path analysis. *Regional Economics Applications Lab*, 97-T-1.

Sonis, M., Hewings, G. and Guo, J. (1998). A new image of classical key sector analysis: minimum information decomposition of the Leontief inverse. *Regional Economics Applications Laboratory.*

Stone, Richard, (1995). Taking a New Look at Life through a Functional Lens. *Science*, 269(21):316-317 (July).

Sudgen, R. (1984). Reciprocity: the supply of Public Goods through Voluntary Contributions. *The Economic Journal*, 94:772-787.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis.* Cambridge U.K., Cambridge University Press.

Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks. *Psychometrika*, 61(3):401-425.

Wolhgemuth, Darin, and Maureen Kilkenny (1998). Firm Relocation Threats and Copy Cat Costs. *International Regional Science Review*, 21(2):139-162.

Wright, C.C. (1979). Arcs and Cars: An Approach to Road Traffic Management based on Graph Theory. in *Graph Theory and Combinatorics* edited by R.J. Wilson; London: Pitman.