Mat-2.108 Independent research projects in applied mathematics

Centrality Measures and Information Flows in Venture Capital Syndication Networks

Mikko Jääskeläinen 26.10.2001
45728s
Department of Engineering Physics and Mathematics
mikko.jaaskelainen@hut.fi
Supervisor: Prof. Ahti Salo
# Table of Contents

1 INTRODUCTION ............................................................................................................ 2

2 THEORY AND RESEARCH SETTING ......................................................................... 4
   2.1 VENTURE CAPITAL ............................................................................................................................ 4
   2.2 SYNDICATION ..................................................................................................................................... 5
   2.3 SOCIAL NETWORKS ............................................................................................................................ 7
   2.4 CENTRALITY ........................................................................................................................................ 8
   2.5 RESEARCH SETTING ......................................................................................................................... 12

3 DATA AND METHODS ................................................................................................. 15
   3.1 DATA .................................................................................................................................................. 15
   3.2 SIMULATION ...................................................................................................................................... 16
   3.3 CALCULATIONS ................................................................................................................................. 17

4 RESULTS ...................................................................................................................... 19
   4.1 BEHAVIOUR OF THE SIMULATION ................................................................................................ 19
   4.2 DIFFERENCES IN RANKINGS ........................................................................................................ 21
   4.3 DIFFERENCES IN SCORES ............................................................................................................... 23

5 DISCUSSION AND CONCLUSIONS ........................................................................... 25

6 REFERENCES ................................................................................................................ 27
Centrality Measures and Information Flows in Venture Capital Syndication Networks

Abstract: This paper examines the performance of four centrality measures in describing flowing and accumulation of information in venture capitalists syndication networks. Although centrality measures should differentiate between network positions indicating those units that are more central than others, the basis of the differentiation does not lie on any theoretical foundation. Thus, applicability of these measures to describe flows and accumulation of information on certain actors is not evident. Utilising an actual syndication network of 161 US venture capitalists, we build a simulation emulating the information flows in this network in order to examine whether the measures correlate with the accumulation of information. The results demonstrate that there are unexpected differences between measures with respect to their explanation of information accumulation. The measure the simulation was built on failed worst to meet the information index, whereas simpler measure performed without flaws. The results contribute to the understanding of centrality measures and their performance in the context of information flows.
1 Introduction

A distinctive characteristic of venture capital firms is the large amount of co-operations involved in an investment process (e.g. Bygrave 1987, 1988, Lerner 1994, Brander, Amit and Antweiler 2001). This co-operation takes an easily observable form in syndicated investments, in which two or more venture capital firms invest at the same time to a start-up company. These syndicated investments form ties between venture capitalists constituting a syndication network (Bygrave 1988, Sorenson and Stuart 2001). The more connected a venture capitalist is, the better it is able to receive information from its syndication partners (Sorenson and Stuart 2001). This information contributes to the performance of the venture capitalist, as it is able to find better investment targets than it would find alone. In addition, syndication helps venture capitalist to do better investment decisions as it receives a second opinion and additional information from its partners (Brander et al. 2001). Furthermore, the complementary skills of venture capitalists in a syndicate may increase the value of the target company, contributing to the overall performance of the venture capitalists (Brander et al. 2001).

The research of syndication networks has remained so far on a descriptive level. In his two pioneering studies, Bygrave (1987, 1988) described networks that venture capitalists form through syndication. Bygrave concluded that information sharing constitutes a major rationale for syndication. In their recent study, Seppä and Jääskeläinen (2001a) have established an empirical connection between the amount of syndication of an individual venture capitalist and its performance. Furthermore, preliminary results from another study of authors (Seppä and Jääskeläinen 2001b) show that in addition to the individual level, also the venture capitalist’s position in the syndication networks affect the performance of the venture capitalist.

The methods used to describe the position of an organisation in an interorganisational network are based on the social network analysis (e.g. Podolny 1997, Ahuja 2000, Sorenson and Stuart 2001). The social network analysis itself, as noted by multiple authors (e.g. Freeman 1979, Friedkin 1991), is not based on any specific theory, but is more of a tool to approach empirical cases in which multiple actors are connected to each other. There is a wide range of indices describing the centrality of a unit in a network. The one considered here, developed by Bonacich (1987), is thought to describe the centrality based on the information flows in a network. However, as not derived from any theory, the basis of the measure is not fully established.

Hence, the applicability of this centrality measure for presenting the amount of information that each unit in a network receives is unclear.

This paper sets out to examine the centrality measure presented by Bonacich (1987) in the light of information flows. The objective is to compare the Bonacich’s centrality measure with the results of simulated information accumulation, and hence to seek to validate the model behind the Bonacich’s measure. The simulation is constructed using an actual syndication network among the 161 largest US venture capitalists. The results contribute to the understanding of the performance of centrality measures, especially of the Bonacich’s measure.

The paper is organised as follows. Chapter 2 introduces the concepts involved in venture capital syndication and social networks. Chapter 3 presents the data with methods and construction of simulation model. Chapter 4 presents the results, which are discussed in Chapter 5, presenting the implication and conclusions.
2 Theory and research setting

2.1 Venture Capital

Wright and Robbie (1998) defined venture capital as the investment by professional investors of long-term, unquoted, risk equity finance in new firms where the primary reward is an eventual capital gain, supplemented by dividend yield. In addition, venture capitalist are usually actively involved in their investment steering their development towards desirable outcomes (Sahlman 1990). This definition corresponds to the perspective taken in this study.

The equity stake in companies separates the venture capitalist from banks and other financiers that lend capital for collaterals. In addition, the active role in investments distinguishes venture capitalists from other private capital investors, such as holding companies. Venture capital includes also those investors who invest in manner characterised by Wright and Robbie above, but who target developed companies. A typical example is e.g. a leveraged buy-out.

Venture capitalists function as a financial mediator between entrepreneurs seeking for the financing and investors investing capital to venture capital funds. Venture capitalists raise funds by taking investments from investors, who thus become limited partners in venture capital partnership. Investors are typically institutions such as pension funds, insurance funds or endowments (Sahlman 1990). Managers of venture capital firms are general partners of firms, and they take care of the daily operations of investing in new companies and steering and monitoring existing investments.

When a venture capitalist invests in a new company, the process takes usually certain steps that are relatively identical in each investment (Bygrave and Timmons 1992, Tyebjee and Bruno 1984). Venture capitalists receive constantly new proposals for investments. This stream of proposals is called deal flow. From this deal flow, the venture capitalist picks those that appear to have potential for an investment. After screening and evaluating potential proposals, the most promising ones are taken step further for more detailed screening and valuation of the proposal. If the venture capitalist decides to invest, the new company becomes a part of the venture capitalist’s portfolio. To earn profits for the limited partners and for themselves, the general partners start to steer and nurture the new portfolio company to help it grow and increase its...
value. Finally, as the company has grown enough, the venture capitalist exists from the investment by selling its stake in the company either on a public marketplace or to another company willing to acquire the portfolio company. However, only two out of ten portfolio companies generate high returns, whereas six out of ten return little more than the invested capital, and the remaining two are complete failures (Bygrave and Timmons 1992).

When a venture capitalist decides to finance a venture, it rarely provides all the capital at once; on the contrary, the investments to ventures are staged. The venture capitalist provides the company with enough capital for it to proceed to the next development stage. Once reached, the progress and future of the company are reassessed. If they meet the VC’s criteria, a new investment is done. Otherwise, the project is terminated and venture capitalist recedes from the venture. By staging the investments, a venture capitalist can preserve an option to abandon the venture if its outlook turns weak. (Gompers 1995)

2.2 Syndication

The syndication is a distinctive feature in the venture capital deals. A syndicated investment is defined as one in which two or more venture capitalists invest in the same company within the same round (e.g. Bygrave 1987, Lerner 1994, Brander et al. 1999). As investments to companies are usually staged, syndication may occur on each financing round. The requirement of simultaneity is however slightly strict for the use of researchers. Once a VC invests in a company, it remains an investor although it would not invest on the next round. Hence, all VC that have invested in a single company have syndicated their investment in a sense. On the other hand, syndication, as Wilson (1968) defines it, is a jointly formed decision to invest in a company, which requires concurrent activity from each part. As a compromise, syndication is often defined as those investments that constitute a single round of financing, although the actual timing of investments would differ.

Syndication networks

The first to handle the topic was Bygrave in his two successive papers (1987, 1988). His focus was on the patterns that US venture capitalists formed through syndication relationships. Bygrave found out that venture capitalists are tightly connected to each other. The largest venture
capitalists were also most connected, forming links to the whole industry. In addition, the networks were centred around a few firms and geographic locations.

The primary interest of Bygrave was in the reasons of co-operation between venture capitalists. Bygrave (1987) claimed that two important formal linkages between VCs are the jointly done investments and seats in the boards of these companies. Hence, syndicated investment forms a node between two venture capitalists that ties them together and allows both formal and informal information flows between the two. He used syndication to represent this linkage to examine the structure of the network of venture capitalists. Sorenson and Stuart (1999) study took similar approach using syndication to represent contact between VCs. They concentrated also on the informational aspect of syndication. They found out that syndication relationship lessens the impediments caused by geographical distance. Venture capitalists that had ties with other firms were likely to invest on larger geographic region, which suggests that they were able to receive information about distant investment proposals through their contacts.

**Rationales for syndication**

The information perspective is present in the majority of suggested rationales of syndication. Additional information first increases the amount of proposals a venture capitalist receives (Bygrave 1987, Sorenson and Stuart 2001). Second, it enhances the decision making by bringing in a second opinion about the quality of proposal (Sah and Stiglitz 1986, Lerner 1994) as well as increasing the capacity and perspectives used to assess the potential of the proposal (Lerner 1994). Finally, as venture capitalists in syndicate have different skills and contact networks, it enhances the value adding of the venture (Brander et al. 2001).

Another perspective to syndication is spreading of financial risk. As venture capitalists syndicate, they are able to increase the size of their portfolio and they can invest to a wider variety of ventures than they could do alone. Although the rationale sounds promising, there has been only little and weak evidence for this rationale. Bygrave (1987), comparing the relative importance of information sharing and financial risk spreading, concluded that risk played only little role in syndication.

Two of the suggested rationales stem from structural sources. Lerner (1994) suggested that venture capitalists do window dressing, that is, they try to demonstrate their quality by entering successful investments on later stage to earn a merit from their publicity. To be able join these
investments, they have to syndicate. Other rationale stemming more from the structure than actions is information asymmetry between entrepreneur and original investor. Admati and Pfleiderer (1994) suggested that information asymmetry forces venture capitalist to hold a constant share of a venture. As the venture grows, additional venture capitalists are needed to invest, if the original investor does not increase its share. Thus, venture capitalists syndicate.

The research has been able to propose more rationales than it has been able to reject. Lerner (1994) tested window dressing, decision-making and information asymmetry hypothesis, finding some support for each of them. Bygrave (1987) noted that information sharing explains syndication, whereas risk spreading does not. Brander, Amit and Antweiler (1999) in turn showed value adding to be more meaningful than selection hypothesis. Although these results does not reject any of the hypotheses, it seems that information and the sharing of it plays a meaningful role in syndication, and thus provides a basis to investigate the subject further.

2.3 Social networks

Social networks are networks formed between actors on the basis of social relations. The connections are based on social interactions, such as friendships, transactions or hierarchies. Usually, when presenting these social relations analytically, the number of actors is low and a social network can be presented as a graph. Due to this graph presentation, the field of study is also referred to as ‘graph theory’ (e.g. Freeman 1979). Although named as ‘theory’, the theoretical basis of the field is almost inexistent. Social graphs are more of a way to approach social relations and they serve as a tool to analyses these relations. The hypotheses presented in the studies of social networks are without exception ad hoc formalisations of plausible ideas, and there is no underlying theory concerning the field (Friedkin 1991).

Figure 1 presents an example of social graph. The connections between units are undirected, that is, the units are in equitable positions with respect to each other. A directed graph would be marked with directions of the influence. These are usual when a graph describes a power relations or hierarchy, where one actor gives orders to another. The terminology on social graphs is similar to all networks. The actors are presented with ‘points’ and the connections between points are called ‘edges’.

When the number of actors is low, a graph is a convenient way to present the network. However, for calculation and especially when the number of actors is large, a matrix presentation becomes more useful. The relational matrix can be either binary, indicating only the presence of relationship, or weighted, giving different importance to each relationship. If relations are undirected, the matrix is symmetric; directed relations yield asymmetric matrix. Figure 2 illustrates the matrix presentation of the graph in Figure 1.

```
1 0 1 1 1 0
2 1 0 1 1 1
3 1 1 0 0 0
4 1 1 0 0 0
5 0 1 0 0 0
```

Figure 2 Matrix presentation of the social network

2.4 Centrality

Although the theory of social networks is weakly established, there are some useful tools developed to analyse them. One of the primary uses of social networks is the identification of those actors that are in more important positions than others are. It is intuitively clear that those with multiple connections are more central when compared to those who have only few contacts. Centrality measures are used to bring up these differences among actors based on their connections to others, and thus to describe the positions in which actors are in a social network.

Centrality can refer both to the position of a single unit or to the overall characteristics of a network. When referring to a network, term ‘graph centrality’ is used. Graph centrality is based on the compactness of the graph. If all units are close to each other, the graph has high centrality. This does not make comparison between units, but describes the graph on more aggregated level.
Centralisation of a graph describes how much some units have tendency to be more central than other units do. In this study, we concentrate solely on point centralities.

The incoherent theoretical base has resulted to a multitude of different kinds of centrality measure. The research on the topic stems from late 1940’s and is still quite active (Freeman 1979). However, there are some more established measures that are used in research utilising social networks. Freeman (1979) reviewed in his acknowledged paper the earlier research on centrality measures and came to suggest three basic measures for point centrality: degree, closeness, and betweenness. These measures are based on the shortest paths connecting units. Bonacich (1987) based his measure on continuous flows in a network. Unlike Freeman’s measures, his measure puts emphasis on the quality of the units one is connected to. Both Freeman and Bonacich represent the most developed forms of measures in their respective traditions.

There are also other recently suggested measures. Stephenson and Zelen (1989) based their measure on the variance of the information received from other units. The more there are connections, the more observations a unit has on specific information, and the smaller is the variance. Friedkin (1991) was first to take an attempt to derive the measures from theoretical foundations. He developed three measure of centrality for process of social influence.

As we shall later see, the measure of Bonacich corresponds best to our research setting of information flows. We will also use the Freeman’s measures in order to compare and to check the validity of our results. The four measures are described in following sections.

Degree, Closeness and Betweenness

First of the Freeman’s measures is based on degree, i.e., the number of units directly connected to the unit under scrutiny. All the measures are divided by the largest possible size the measure to make measure comparable and more intuitive. In the case of degree, the unit can be connected to all units in the network but itself. Hence, the measure is divided by n-1. The centrality measure of degree is

\[ C_D(p_k) = \frac{\sum_{i=1}^{n} a(p_i, p_k)}{n-1}, \quad (1) \]

where \( p_i \) is unit i, and \( a(p_i, p_k) = 1 \) if and only if \( p_i \) and \( p_k \) are connected. Otherwise \( a(p_i, p_k) = 0 \).
The second measure is based on closeness of other units. It reflects the mean length of the shortest paths between the unit i and the other units. The longer the paths, the further is the unit from others, and the less central it is. Thus, the measure is defined as inverse centrality,

\[ C'_{c(i)} = \frac{n-1}{\sum_{j=1}^{n} d(p_i, p_k)}, \]

where \( d(p_i, p_k) \) is the number of edges between in the shortest path connecting \( p_i \) and \( p_k \).

The third measure is based on the unit role on the connecting paths of other units, and thus it is called betweenness. This is defined as

\[ C'_{b(i)} = \frac{2 \sum_{i < j} \sum_{k} b_j(p_k)}{n^2 - 3n + 2}, \]

where

\[ b_j(p_k) = \frac{g_{ij}(p_k)}{g_{ij}}. \]

\( g_{ij} \) is the number of all geodesics (i.e. shortest paths) between i and j. \( g_{ij}(p_k) \) is the number of those geodesics between i and j that pass through unit \( p_k \). Thus, the measure is based on the probability that \( p_k \) is part of the communication of i and j. The denominator of equation (3) is the largest possible value of the measure, as above.

**Bonacich’s centrality**

Bonacich (1987) defined a centrality measure that is based on the centralities of units that a unit is connected to. The measure \( c_i(\alpha, \beta) \) is defined as

\[ c_i(\alpha, \beta) = \sum_j (\alpha + \beta c_j) R_{ij}, \]

The elements of \( R \) are the number of companies in which the firms i and j have invested together. \( \beta \) is the degree to which the centrality of i is function of centralities of others. It can be though also of as the radius of the influence of i. If we do not expect the relationships of j to directly benefit i, \( \beta \) should be small, and vice versa.
\( \alpha \) is used as normalisation factor. Because centrality measures vary, when the size of the network varies, we need to make the measures from two networks comparable. To standardise the centrality measure across networks, \( \alpha \) is chosen so that the sum of squared centralities equals the number of units in the network (Bonacich 1987), that is,

\[
\sum_i c_i(\alpha, \beta)^2 = n \tag{6}
\]

Doing this, the unit that has \( c_i(\alpha, \beta) = 1 \) does not have an unusually large or small centrality, irrespective of the number of positions in the networks (Bonacich 1987).

**Differences between measures**

The differing characteristics imply different results for point centralities. Table I presents the results of each centrality measure for the graph presented in Figure 1. Each measure contributes the same unit, number 2, with highest centrality score. In addition, the second largest score also coincides across measures. Except the betweenness, also the other ranks are the same. However, the differences between units are not proportional. E.g. closeness reports smaller difference between unit 5 and units 3 and 4 than the degree measure. Betweenness does not make separation between these units.

<table>
<thead>
<tr>
<th>unit</th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweenness</th>
<th>Bonacich</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.8</td>
<td>0.08</td>
<td>1.16</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0.58</td>
<td>1.39</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>0.67</td>
<td>0</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>0.67</td>
<td>0</td>
<td>0.87</td>
</tr>
<tr>
<td>5</td>
<td>0.25</td>
<td>0.57</td>
<td>0</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Bolland (1988) examined the performance of these measures in real and simulated networks. He concluded that the betweenness yields the different centralities than the others, which were more similar to each other. This can be seen also in the Table I. Betweenness tends to give higher score to units that connect to parts of the network, while Bonacich’s measure deflate the centrality from outskirts more to the centre of the network. However, Bolland used an earlier version of Bonacich’s measure lacking the parameters, which gives more weight for the longer, indirect paths. In all, we can expect differing results from each of the measures.

---

2.5 Research setting

Information as a resource

The purpose of the study that originated the problem of measuring centrality from the point of view of information flows was to see, whether centrality amounts to increased performance of venture capitalists. If causality exists, it implies that the venture capitalist is able to utilise the information received to enhance its performance. From the point of view of the resource-based view of firm, the information hence serves as a valuable resource creating potential for competitive advantage (e.g. Barney 1991, Peteraf 1993).

To hold such a potential, the information has to be either something enabling the venture capitalist to add value to its portfolio companies or then information about potential investment targets. The network is based on realised investments, which means that each tie is based on existing co-operation relationship. In addition, syndication with other venture capitalists implies that a venture capitalist is willing to share information on potential investments, which makes it seem likely that information relevant to performance of venture capitalist may very well pass through the connections of the network.

If we assume that venture capitalists receive information through their linkages to other venture capitalists, the amount of linkages should affect the amount of information received. In addition, if information is a potential source of competitive advantage, those firms with most information should perform better than those with less information, ceteris paribus. Hence, the more a firm has contacts, the better it should perform.

To test this hypothesis, we need to have a measure describing the contacts of a venture capitalist, that would differentiate between more and less connected units. Centrality measures offer a promising opportunity.

Centrality measures on information

The applicability of the centrality measures to describe the relative amounts of information that venture capitalists receive in different points of a network is not self-evident. As noted earlier assumptions behind the measures do not correspond to the information sharing between venture capitalists.

The measures of Freeman can be thought to base on communication. The degree is the number of most potential partners to communicate with, as it is the number nearest units. This measure is quite intuitive, but has a significant shortcoming. Degree takes into account only the immediate neighbours neglecting the connections behind the closest actors. Closeness, on the other hand, captures the network better. Since it is the mean distance to all other units, depicting the length of the path the communication has to take to reach other units, it describes better the centrality of the unit. The nearer the unit is to other units, the more it is likely to receive information. Betweenness, on the other hand, is based on communication from the beginning. It bases on the unit’s role in communication of other units. If it is on the shortest path of two units, the communication between these units will likely pass it on its way. Thus, it can tap into this communication and in addition, it can control this communication to an extent, hence giving it power over the other units.

The basic shortcoming of Freeman’s measures is that they assume that communication takes only geodesic paths (Stephenson and Zelen 1989). Thus, they neglect the possibility that communication could take a longer route. In addition, each measure seems slightly inappropriate from the point of view of communication between venture capitalists. Most likely the information flows originate further than the degree measure assumes. Closeness, weight similarly all paths, although it is likely that the longer the path, the less likely it will reach the unit. Finally, the betweenness is more a measure of power than communication.

Bonacich’s measure is however more promising. It utilises all paths, both direct and indirect, between actors. The paths are also weighted inversely to their lengths. Hence, the measure takes into account both the whole network and the probability that communication is transmitted through these paths. If we take parameter $\beta$ as a probability, the Bonacich’s centrality measure $c(\alpha, \beta)$ can be interpreted as the expected number of paths in a network activated directly or indirectly by each individual. When $\beta$ is less in absolute value than the reciprocal of the largest eigenvalue of $R$, $c(\alpha, \beta)$ can be presented as infinite sum,

$$c(\alpha, \beta) = \alpha \sum_{k=0}^{\infty} \beta^k R^k 1 = \alpha (R1 + \beta R^2 1 + \beta^2 R^3 1 + ... ) \quad (7)$$

Thus, $c(\alpha, \beta)$ is the total number of paths from position $i$, when each path is weighted inversely to its length. This is useful when we consider a communication network. In this case $c(\alpha, \beta)$ can
be though as expected value, while $\beta$ is the probability that once information is received, it will be transmitted to any receiving contact. Hence, $c(\alpha, \beta)$ is the expected amount of communication the unit $i$ receives. However, $c(\alpha, \beta)$ contains also a sort of redundancy when considered from the information point of view. The paths presented in equation (7) may pass the same point more than once. While communication on a certain piece of information involves multiple contacts, the information is new when it is received for the first time. Hence, the following communication involving the information is redundant. Although referred to as information centrality (Wasserman and Faust 1994), the measure of Bonacich potentially contains a bias when used to present the accumulation of information.

**Simulation of information accumulation**

Since the performance of the centrality measures is in doubt, we construct a simulation to examine the behaviour of the measures. We simulate the assumed information transmission in a network of venture capitalists. Our purpose is to use the simulation to identify those venture capital firms that receive most information through their connections. Once we have conducted the simulation, we compare its results with the corresponding results of the centrality measures.
3 Data and methods

3.1 Data

The network underlying the simulation is constructed from actual venture capital investments. Some studies (e.g. Bolland 1988, Donninger 1986, Snijders 1981) have utilised simulated, or random, networks, but their aim has been on the characteristics of the centrality measures, while we aim to examine the correlation between the measures and the information accumulation.

The venture capital investment data, from which the network is constructed, is obtained from Securities Data Corporation’s Venture Economics database. This extensive source has been used also in previous venture capital research (e.g. Bygrave 1987, Lerner 1994, Gompers 1995, Gompers and Lerner 1998). Venture Economics has gathered venture capital investment data since the 1970s using annual reports of venture capital funds, personal contacts to funds’ personnel, initial public offering prospectuses, and acquisitions announced in the media. The database contains information on over 150,000 private equity investments (one whole financing round consists of several single investments) and it is widely recognised as a leading source of U.S. venture capital investment data.

We define our sample as the 161 largest private U.S. venture capital organisations and identify them on the basis of the number of portfolio companies the firms had invested in by the end of the year 2000. These numbers were extracted from the Venture Economics investments database. All investments are made by U.S. venture capitalists in U.S. companies. Altogether, the data includes 54,700 investments into 10,057 portfolio companies in the years 1980-2000.

We constructed the syndication networks by first screening the data for investor-investment pairs and then connecting the investors using their investment targets. As a result, we had a list of syndication pairs with indication of the number of co-investments. This list was then translated to a relational matrix.

It is notable, that the availability of data has affected to the selection of the measure connectedness. The investments of venture capitalists are closely tracked and recorded and thus there is sufficiently data available for researchers. However, one should remember that a

syndicated investment is only a proxy for all the connections between two venture capitalists. There is a multitude of informal relations between venture capitalists. It very likely that most of these relationships never lead to a syndicated investment, and still contribute to the operations of these VCs. There are no readily available data in existing databases, and mapping of these informal relationships would require an extensive survey. However, the purpose of this study is to validate the measure of Bonacich, not to examine the actual relationships. Thus, the nature of data does not affect the results.

3.2 Simulation

Model of information flow

We construct a simulation to presents the flow of information among venture capitalists. We build this simulation on a model of emergence and transmission of information. In reality, each venture capitalist deploys its resources and contacts to conduct research in order to find new investment targets. We suppose that this process produces information that is novel to each venture capitalist. Once a venture capitalist comes up with new information, it may pass the information to its partners. We identify these partners using syndicated investments.

Once a venture capitalists has acquired information, it does not transmit information automatically, but with some probability of $\lambda$. Hence, new information emerges to the network of venture capitalist randomly through the members of the network, who subsequently pass it to their partners with some probability. These partners pass it on in a similar manner until either the information has spread through the network or transmissions have ceased.

We randomise the source of information, which can hence emerge anywhere in the network. Each unit has equal probability to originate information. Then, the transmissions of this information are followed as it proceeds in the network, and each unit that receives the information, is recorded. Once the transmissions of this particular information come to end, the records of units that received the information are stored to keep record of the overall amount of information each unit receives during the simulation. After this, new piece of information is launched from random source and the process takes place again.
Hence, we are able to find out how much information each unit receives. Comparing these amounts and ranking them, we can compare them against the centrality scores presented above. To make the simulation comparable with the Bonacich’s model, we use set the probability of transmission $\lambda=\beta$. We follow the practice of earlier studies and set $\beta$ as three quarters of the reciprocal of the largest eigenvalue of $R$ (Sorenson and Stuart, 1999, Podolny, 1993).

**Information indices**

We call the measure created as a result of the simulation as information index to separate in from centrality measures. The information index is based on the amount of information a unit receives through its contacts. The measure is scaled with highest score so that measure receives values between zero and one.

We calculate two variations of the index. In the first case, syndicated investments reflect only an existing relationship and the strength of these ties are taken as equal. The second variation utilises also the number of syndicated investments as an indication of the strength of the tie. The probability of transmitting information is made directly proportional with the number of investments. This assumes that venture capitalists prefer those syndication partners with whom they have invested more to those that are involved in fewer investments.

**Calculations**

Each simulation run for 80 000 rounds, meaning that 80 000 separate pieces of information were tracked through the network. Simulations were repeated for 240 times. Thus, we have 240 observations for each 161 units of the network.

In addition, we run the simulation using multiple smaller networks to examine the behaviour of the simulation. In these examinations, we used networks of size 10, 20, 40 and 60. These networks were generated randomly with arbitrarily chosen probability of 0.5 for a tie to exist between any two units. The generated networks are denser than the actual network, i.e., there are more connections between units than in actual network. However, these generated networks are only used to examine the characteristics of the simulation and to determine optimal number of rounds to run the simulation. Thus, this discrepancy does not affect the final results, but helps to optimise calculations.
The main reason for using smaller, randomly generated networks is that the simulation is computationally heavy. The 240 repetitions with 80,000 rounds took nearly 70 hours of CPU time on relatively powerful computers. The total time of CPU used more than doubles when we add the time used for testing the model and examining its behaviour. There is clearly room for improvement of the simulation. It is also questionable whether the Matlab offers suitable platform to run the simulation on, or should one create a separate computer program from running simulations.
4 Results

4.1 Behaviour of the simulation

In order to assess the performance of the simulation, we carried out a series of tests to examine the behaviour of the simulation. We generated random networks of varying size to relate the number of units in network and the required amount of rounds in simulation. The actual network of 161 units is too large for comfortably to examine characteristics of the simulation.

In addition, by using randomly generated smaller networks for testing, we are able to acquire information on the effects of the value of $\beta$ and the number of rounds on the correlations. In the actual network, we follow the convention of using the reciprocal of the largest eigenvalue. Thus, simulated networks allow us some room to test and explore the model and the simulation.

Convergence

To determine the required amount of rounds in simulation, we examined the convergence of the correlation coefficient between the measure of Bonacich and the results of simulation. We altered the value of the parameter $\beta$, as well as the size of the network and the number rounds we run the simulation.

The smaller the value of $\beta$, the more rounds are needed to produce reliable results. When $\beta$ is very small, a small number of rounds may be too little to produce observable results. If $\beta=0.001$, it takes on average 1000 repetition to transfer information only once, whereas with $\beta=0.01$, the number is 100. Figure 3 illustrates this effect of $\beta$ on the convergence of the simulation. The figure presents the correlation coefficient between the result of simulation and the Bonacich's measure. The higher the $\beta$, the faster the simulation converges and the smaller the number of rounds needed. The simulation used randomised networks of ten units.
Corresponding tests conducted with different sizes of networks and number of rounds yielded similar results. The larger the network, the later the simulation converges. With 40 units and \( \beta = 0.001 \), the model started to converge after 60000 rounds, while 80000 rounds were needed when the number of units were 60. However, the randomised networks were denser than the actual network. That is, there were more connections between units than in the actual case. With fewer connections there are less interactions that the simulation has to demonstrate and hence the actual simulation should converge relatively faster than the denser, randomised networks.

We tested this by running the simulation with the actual network and value of \( \beta \). The simulation seems to converge with relatively low number of rounds, as there is not much difference with the results whether the number of rounds is 40000 or 100000 (see Figure 4). Hence the use of 80000 rounds is a compromise between robustness and time used to calculations.
Saturation

The probability of transmission affects how wide the information will spread from the origin. The higher the probability that the information will be passed on once received, the further the information is spread. As noted in equation (7), the infinite sum will converge if $\beta$ is small enough; otherwise, the sum diverges to infinity. Much the same applies to the simulation. If $\beta$ is large enough, every piece of information will receive every unit in network, and the network is saturated. Thus, $\beta$ has to be small enough to prevent the saturation.

However, $\beta$ is not limited to the theoretical maximum of reciprocal of largest eigenvalue, but the simulation tolerates also higher values. As the model is discrete and finite, the flow of information is likely to terminate rather fast even if $\beta$ is higher.

Although the simulation is not sensitive to the exceeding the theoretical maximum of $\beta$, it is not able to handle all possible values. The larger the $\beta$, the further the information is spread. Finally, with sufficiently large $\beta$, the information reaches nearly each unit regardless of the starting point. In our tests, this happened at latest when $\beta=0.5$. The simulation is unable to generate differences between units and the correlation with Bonacich’s measure vanishes. Although we can calculate the centralities with Bonacich’s measure with any $\beta \in [0,1]$, the simulation is applicable only with low values of $\beta$.

4.2 Differences in rankings

Each centrality measure yields a score for each unit. Given the four measures added with two versions of simulated indices, we have six alternatives to use in describing the centrality of a unit. Based on different views on the role of the unit in the relationship network and hence using different methods to measure the centrality, each measure attaches different score to the unit.

We begin our examination of the differences between measures by reviewing the ordering of the units according to each measure. Each measure attaches score to units, which are then used to order the units from the most central to the most peripheral. This way we are able to assess whether measure identify the same unit as most central, and if not, how much the ranks differ.
Table 2 presents the ten most and least central units according to each measure. Some measures, like degree and closeness, seem to rank units very similarly, while there are larger differences among others. Especially the Bonacich’s measure seems to yield considerable different results than others. While all the other measures identify the unit 102 as most central and the unit 37 as the least central, the Bonacich’s measure does not rank 102 in top ten, and identifies two less central units than 37.

### Table 2 Highest and lowest ranking units

<table>
<thead>
<tr>
<th>Rank</th>
<th>Information dichotomised</th>
<th>Information weighted</th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweenness</th>
<th>Bonacich</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>72</td>
</tr>
<tr>
<td>2</td>
<td>66</td>
<td>85</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>142</td>
</tr>
<tr>
<td>3</td>
<td>124</td>
<td>92</td>
<td>124</td>
<td>124</td>
<td>136</td>
<td>94</td>
</tr>
<tr>
<td>4</td>
<td>85</td>
<td>126</td>
<td>85</td>
<td>85</td>
<td>109</td>
<td>79</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>109</td>
<td>25</td>
<td>25</td>
<td>35</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>29</td>
<td>76</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>71</td>
</tr>
<tr>
<td>7</td>
<td>126</td>
<td>124</td>
<td>126</td>
<td>126</td>
<td>124</td>
<td>69</td>
</tr>
<tr>
<td>8</td>
<td>107</td>
<td>94</td>
<td>107</td>
<td>107</td>
<td>93</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>109</td>
<td>150</td>
<td>109</td>
<td>109</td>
<td>25</td>
<td>124</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>77</td>
<td>35</td>
<td>35</td>
<td>85</td>
<td>91</td>
</tr>
</tbody>
</table>

To quantify the differences between rankings, we calculate the Spearman correlation between measures. Table 3 presents the results. All measures correlate highly with each other. The Bonacich’s measure presents the lowest correlations with other measures, as suggested by the listing of rankings. Degree and closeness measures have identical rankings. The dichotomised information index yields also very similar results. Furthermore, difference to weighted information index remains small, as well.

Table 3 Rank correlations between centrality indices

<table>
<thead>
<tr>
<th></th>
<th>Information</th>
<th>Weighted</th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweenness</th>
<th>Bonacich</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>-</td>
<td>0.9932</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9941</td>
<td>0.9547</td>
</tr>
<tr>
<td>Weighted</td>
<td>0.9932</td>
<td>-</td>
<td>0.9932</td>
<td>0.9841</td>
<td>0.9583</td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>0.9999</td>
<td>0.9932</td>
<td>-</td>
<td>0.9942</td>
<td>0.9546</td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>0.9999</td>
<td>0.9942</td>
<td>0.9942</td>
<td>-</td>
<td>0.9546</td>
<td></td>
</tr>
<tr>
<td>Betweenness</td>
<td>0.9941</td>
<td>0.9583</td>
<td>0.9546</td>
<td>0.9546</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Bonacich</td>
<td>0.9547</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In conclusion, all but the Bonacich’s measure yields very similar results and places the same units in same rankings.

4.3 Differences in scores

When used as independent variable in a regression model, the mere identification of most central unit is not sufficient, but also the value of a centrality measure counts. In this case, the measure has to answer the question of how much more central a unit is from another.

The similarities between measures can be easily assessed with ordinary correlations. Table 4 provides the correlation coefficients between the measures. Results show more differences between measures than the rank correlations. While the Bonacich’s measure ranked units almost identically as others, the correlations provide evidence for bigger differences to other measures. Most interestingly, the Bonacich’s measure correlates less with the simulated indices than with any of the other measures. Instead, the degree seems to correlate with the dichotomised information index almost perfectly. Closeness is also very similar to the dichotomised index.

Table 4 Correlations between centrality indices

<table>
<thead>
<tr>
<th></th>
<th>Information</th>
<th>Weighted</th>
<th>Degree</th>
<th>Closeness</th>
<th>Betweenness</th>
<th>Bonacich</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>-</td>
<td>0.841</td>
<td>0.999</td>
<td>0.988</td>
<td>0.889</td>
<td>0.699</td>
</tr>
<tr>
<td>Weighted</td>
<td>0.845</td>
<td>-</td>
<td>0.888</td>
<td>0.865</td>
<td>0.593</td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>0.990</td>
<td>0.893</td>
<td>-</td>
<td>0.936</td>
<td>0.663</td>
<td></td>
</tr>
<tr>
<td>Closeness</td>
<td>0.936</td>
<td>0.663</td>
<td>0.936</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Betweenness</td>
<td>-</td>
<td>0.467</td>
<td>0.663</td>
<td>0.663</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Bonacich</td>
<td>0.699</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are more differences also between the two simulated indices than the rankings suggested. If rankings are nearly identical, but values correlate less, it implies that the differences between units are differently distributed. In fact, as we can see from Figure 5, some of the measures create
Results

higher differences in the ends of the scale. These figures present the centrality scores of units in descending order. This illustrates how point centralities are distributed among units.

The weighted information index emphasises the more central score more than the dichotomised information index, but on the other hand makes creates smaller differences in the lower end. This explains why these to indices correlate less than what rankings had suggested. It is notable that most measures have almost linear relation in the middle of interval, while the other or both ends are then either emphasised or downplayed.

Figure 5 Distributions of centralities

5 Discussion and conclusions

This paper set out to examine the centrality measures in the light of information flows. The setting was based on two assumptions derived from earlier literature. First, those venture capitalists with most contacts to other venture capitalists are more likely to receive more information than others do. Second, centrality measures should differentiate between network positions indicating those units that are more central than others are. The purpose of the study was to validate whether the measures correspond with these assumptions.

Results of this paper indicate that there are differences between centrality measures with respect to how well they correspond to the information accumulation model. The centrality measure of Bonacich that served as the basis of the model for the simulation and hence was the primarily under scrutiny, corresponded least with the model. The degree measure, on the other hand, correlated almost perfectly with the index describing the amount of received information.

There are a few potential explanations for this surprising outcome. First, the measure of Bonacich is based on the weighted sum of all direct and indirect paths to other units. Therefore, there is overlapping among these paths, which results to redundant connections. Information is new only on the first time and does not serve as a source for an advantage if received multiple times. Thus, there is difference between simulation, which counts only the first time information is received, and the actual measure with redundant paths. Second, the probability of transmission was based on the parameter of the Bonacich’s measure. This was set as three quarters of the reciprocal of the largest eigenvalue of the relations matrix. This cumbersome definition stems from the interpretation of the Bonacich’s measure as information transmission model, where parameter serves as probability of transmission. However, as this interpretation is based on infinite sum, the parameter has to be small enough for sum to converge. Now, as we use this low probability in our simulation, it means that the information is very unlike to flow any further than to closest neighbours. Thus, the multiple indirect paths that are counted in actual measure distort the correlation. This is also the reason why the information index had higher correlation with the degree measure than with the measure of Bonacich. Third, there is a possibility that the simulated model does not correspond to the Bonacich’s measure. The social processes are many dimensional and involve multiple actors and incentives. It may be that the nature of information that was taken as the basis of the simulation does not correspond to the view Bonacich had when

developing his measure. We are concerned of information about investment possibilities. To this kind of information there is tied both social and financial commitments. However, if the information is nothing more than a joke spreading as word of mouth, it results to very different kind of social process. However, we can place this process under the title of ‘information flow’ as easily as the investment information process. Thus, the social context may not be the one intended originally.

Although the results may well be based on misinterpreted social model, they have nevertheless a significant implication. Clearly, we are not able to use the measure of Bonacich to describe the information accumulation among venture capitalist. For this purpose, the degree measure seems to be both more accurate and easier to calculate. However, the results do not imply that the Bonacich measure would be inapplicable in venture capital context. The results only show that information flow is insufficient as only interpretation of the measure. It may well capture other aspects, such as power, that are essential to the centrality.

This study has answered to one question, but simultaneously risen several others. Now that the information flow model used in this study did not correspond to the measure of Bonacich, it raises question that what model would. This opens the door for more profound sociological pondering on the nature of the information sharing process and its dimensions. On the other hand, the simulation was limited to one version should have best corresponded to the Bonacich’s model. With higher transmission probability, the simulation may yield more interesting results, which in turn raises the question whether it would be in this case applicable to some other social process.

6 References


Podolny, M. and A. Feldman. 1997. Is it better to have status or to know what you are doing?: An examination of position capability in venture capital markets. Stanford University research paper no. 1455.


