Statistisk sentralbyrå
Statistics Norway

Relational capital

Do board networks exist and are they valuable?

TALL SOM FORTELLER

Solveig Bjørkholt, Erik Fjærli and Marina Rybalka

RAPPORTER / REPORTS

2019/ 36

Reports 2019/36

Solveig Bjørkholt, Erik Fjærli and Marina Rybalka

Relational capital

Do board networks exist and are they valuable?

Statistisk sentralbyrå • Statistics Norway
Oslo–Kongsvinger

In the series Reports, analyses and annotated statistical results are published from various surveys. Surveys include sample surveys, censuses and register-based surveys.

Published 29 November 2019

Print: Statistics Norway

ISBN 978-82-587-1024-7 (printed)
ISBN 978-82-587-1025-4 (electronic)
ISSN 0806-2056

Symbols in tables	**Symbol**
Category not applicable	.
Data not available	..
Data not yet available	…
Not for publication	:
Nil	-
Less than 0.5 of unit employed	0
Less than 0.05 of unit employed	0.0
Provisional or preliminary figure	*
Break in the homogeneity of a vertical series	—
Break in the homogeneity of a horizontal series	|
Decimal punctuation mark	.

Preface

This report is a part of a project entitled “Capturing the value of intangible assets in micro data to promote the EU's growth and competitiveness” — GLOBALINTO, with financial support from the European Commission/Research Executive Agency under the Horizon 2020 program. The work presented here is related to the deliveries under work package 3, i.e. delivery number 3.2 “Develop, validate and refine micro-level measures of intangibles related to owners’ business network participation”.

Statistics Norway, 18.11.2019

Jan Henrik Wang

Abstract

This report attempts to seize the term relational capital, defined as an enterprise's network participation.

We use network analysis tools on Norwegian register data and demonstrate that there indeed exists a network between enterprises, through boards forming interconnected nodes. The network also contains many sub-networks (clusters).

Opposed to most of the previous studies our data contain information on small and medium-sized enterprises (SMEs). We find that the SMEs are highly network-forming and that this sub-group of enterprises forms a network with a very lumpy structure, i.e., having the tendency to form clusters.

We also examine how the network evolves from our starting point in 2004 to 2017 and find that the network expands as the enterprise population increases. Both the number of boards and board members has increased where the former has increased faster, implying more boards per board member over time.

Board networks can be analysed at the individual level, where the nodes in the network consist of board members, or at the enterprise level where the nodes consist of enterprises. At the individual level, we find that gender affects the density of networks, women having higher network density. However, this could possibly be an effect of women's network being smaller than men's network.

Looking at nodes defined by enterprises, we present in chapter 5 an analysis of how R&D spending is related to an enterprise's connectedness to the network. On the proviso that we do not know much about the mechanisms or the nature of causal relationships, we find that a high network connectedness has a positive impact on R&D spending. Most interesting, we find that this correspondence is particularly strong among SMEs.

Our findings suggest that network participation may add value to enterprises through transmission of information or other incentives to increased R&D effort, and that board networks may represent such a channel.

Sammendrag

I denne rapporten forsøker vi å gripe begrepet «relasjonskapital», definert som foretakenes deltagelse i nettverk, og vi undersøker om slik nettverksaktivitet kan utgjøre en form for immateriell kapital.

Vi benytter verktøy for nettverksanalyse på norske registerdata for styremedlemskap og viser at styrenettverk er en realitet, ved at styrer i foretak i stor grad danner nettverk og klynger i nettverk. Dette synes særlig utbredt blant små og mellomstore foretak (SMB).

Vi undersøker også hvordan styrenettverket i Norge har utviklet seg fra vårt første år (2004) til siste år i dataene (2017). Vi ser at nettverket har vokst betraktelig gjennom analyseperioden, det vil si at det ble både flere styremedlemmer, flere foretak som ble bundet sammen og flere styrer per styremedlem.

Foruten å være forbindelse mellom foretak kan styrenettverk også studeres som nettverk av personer. Når vi ser på nettverket med individer som noder finner vi at under-nettverket av kvinnelige styremedlemmer er tettere enn nettverk bestående av kun menn. Dette kan delvis forklares med at det er betydelig flere menn som er medlem av bedriftsstyrer enn kvinner.

I kapittel 5 studerer vi styrenettverket med fokus på forbindelsen mellom foretak, altså der foretakene er definert som noder. Med forbehold om at vi ikke kan si noe sikkert om virkemåten og kausalitet, finner vi at det er en klar sammenheng mellom at foretak har sterk nettverkstilknytning og å ha høy FoU-aktivitet. Sterkest er sammenhengen blant små og mellomstore foretak.

Våre resultater kan gi støtte for hypotesen at informasjonsflyt og/eller overføring av andre ressurser via styrenettverk kan gi økt FoU-innsats i foretak.

Contents

1. Introduction

Work on the measurement of intangibles has focused on broadening the conceptualization of what constitutes a capital investment, developing measures of intangibles at the macro level and more recently also at the micro level for individual enterprises. The latter is a central part of the GLOBALINTO project, which is a research cooperation between Statistics Norway and European collaborators which is financed by the European commission under the Horizon 2020 programme. The aim for GLOBALINTO is “to provide new measures of intangible assets at the firm level, filling an important gap in measurement which has restricted statistical production, micro-based analysis and evidence-based policymaking”.

Intangible assets are normally classified into three categories:

- Organizational capital
- Broad research and development assets
- Information and communication assets (ICT).

Organizational assets are accumulated through investments in management and marketing activities, while R&D assets are accumulated through the technical activities of the enterprise. Networks are related to organizational capital.

Networks, if they exist, can be regarded as pathways through which information is shared (Larcker, 2013) or resources are being transferred. For example, board members with access to information about new technologies or market opportunities can represent a value to the enterprise on their own. Wincent et al. (2010) define the degree of board interlocking as a part of enterprises’ board capital (relational capital), along with human capital, and find that both have a positive influence on innovation in small and medium-sized enterprises (SMEs). Besides influencing an enterprise’s R&D and innovation, well-connected board members may also provide access to scarce financial resources through corporate network ties (Gygax et al., 2017).

Skilful individuals may also be appointed as board-members after suggestion from external investors, who themselves are part of business networks. The existence of board networks and their role as information channels is documented in Akbas et al (2016), who finds that sophisticated investors like short sellers, option traders and financial institutions are more informed when they trade stocks of enterprises with well-connected board members.

Boards are the central feature of a company, representing the interests of its shareholders. Of course, the idea that board networks may be important for an enterprise’s R&D policy or other managing decisions presuppose that besides their supervisory role, boards also have an advisory role (Schwartz-Ziv and Weisbach, 2013).

The social network literature provides well defined and standardized measures describing networks and nodes in networks, see for example Borgatti et al (2013). Attempts to measure the effect of network participation on firm performance is of course more faceted. One approach looks at enterprise involvement in innovation or R&D spending. Some studies with this perspective look at sharing of technology that takes place within strategic alliances (Gomes-Casseres et. al., 2006) or through business group affiliations (Belenzon and Berkowitz, 2010), both leading to more innovation activity. It should perhaps be no surprise that such formal networks like contracted cooperation and mutual ownership (business groups) are associated with coordinated R&D spending and innovation, but can more informal networks also play a role as information channels?

Balsmeier et al (2014) find that having an outside director from other enterprises in the board leads to increased patenting, while Oh and Barker (2018) find that CEOs that also have positions as board members of other enterprises adapt the R&D intensity of these enterprises in the R&D efforts of their own enterprise. Helmers et al. (2017) focus on possible information transmission through board interlocks, using a difference-in-differences approach on a “natural experiment” that appeared from a corporate governance reform in India. They find that the treatment group, which experienced an exogenous increase in their network size, showed increased R&D and increased patenting. Similar, using a sample of listed enterprises (the S&P 1500 index), Chuluun et al (2017) finds that those listed enterprises with more extensive and central networks have more innovation.

The term relational capital refers to the value to the single enterprise[1] of interlocks to other enterprises via board networks and is regarded as a part of the enterprise’s intellectual capital. However, being an intangible asset, it is not easy to measure and even more difficult to appraise. Networks may be important for R&D intensity, innovation activities and access to financial capital.

SMEs are much in focus for public policy both at the national level as well as in the context of EU and OECD, as they are regarded as a key factor in the process of innovation through business dynamics. A prominent example is the OECDs DynEmp project[2]. The fact that data on board membership for SMEs have so far not been available in many countries, to the same extent as data for listed enterprises, also demonstrates the importance of the results that we present in this report, which builds on data including SMEs.

While the ultimate task in our context would be to estimate the value of networks as an intangible asset, our ambition in this report is restricted to the task of identification and characterization of board networks, using Norwegian register data. The main purpose of the report is to describe the formation of board networks in Norway over a long period. However, we also want to provide a basis for future research on possible effects of board networks, both for the single enterprise and to the society. With this task in mind, we present the results from an analysis of how R&D spending is related to network participation, based on register data for board membership matched with the Norwegian R&D survey.

We describe networking among Norwegian enterprises and how the board network evolves over a period of 14 years, starting in 2004. Our focus is partly on describing the networks and how they evolve, partly on size characteristics of the participating enterprises that form the nodes in the network.

Most of the literature on board networks employs data for publicly listed enterprises, where information on board membership and CEOs are publicly available. Consequently, the results from these works will not be representative for most enterprises – and certainly not for SMEs[3]. Our register data, that covers the entire population of Norwegian enterprises, allows us to use advanced network analysis and calculate centrality measures like the Degree centrality and Eigenvector centrality on data for enterprises regardless their size, localization and industry.

[1] And possibly its social value, through positive externalities.
[2] Criscuolo et al, 2014.
[3] Wincent et al (2010) includes SMEs, but uses enterprises’ self-reported data and must rely on naive measures of interlocking like the total number of links an enterprise has to others.

2. Data

2.1. Dataset

All data used in this analysis is gathered from Statistics Norway. The main data is from the national role register, which contains information on the professional role of actors in Norwegian enterprises, where data on individual identification number and enterprises' organization number were used to map board members. We also used data from the national shareholder register and from the national accounting register to trim data and filter out relevant persons and relevant enterprises.

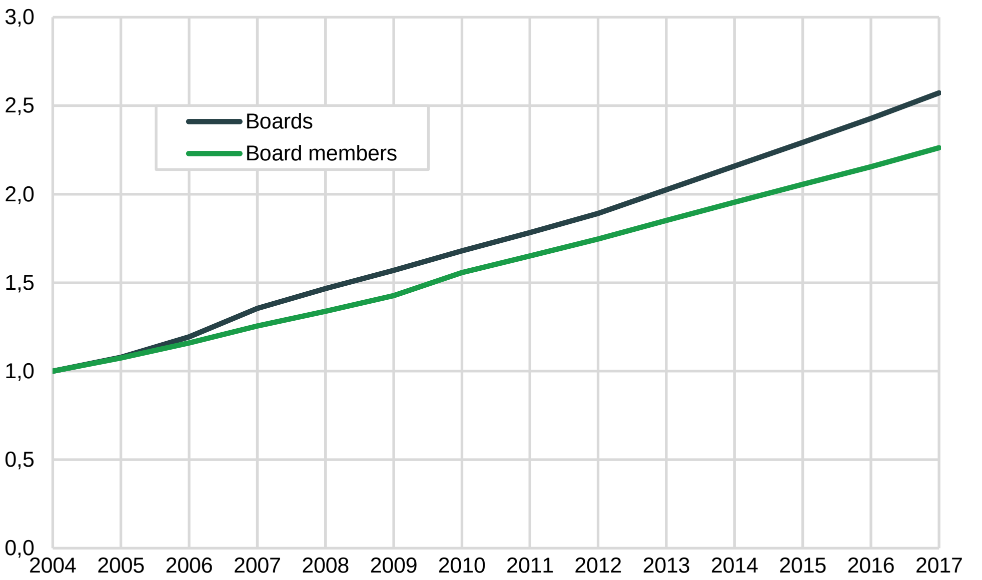
From Table 2.1 we see that the network expands during our observational period. Given that the enterprise population increases it is not surprising that both the number of connected boards and board members grows. At the same time the former grows faster than the latter (jf. Figure 2.1), implying that number of boards per board member is increasing over time.

Table 2.1 Board network sample size before trimming

Year	Boards and board members	Boards	Board members	Growth rate boards	Growth rate board members
2004	597,485	263,845	355,340	-	-
2005	647,651	284,489	381,704	0 .078	0 .074
2006	713,915	314,824	412,301	0 .107	0 .080
2007	798,842	357,261	446,032	0 .135	0 .082
2008	866,618	386,964	475,871	0 .083	0 .067
2009	931,583	414,357	507,366	0 .071	0 .066
2010	1,010,507	443,307	553,320	0 .070	0 .091
2011	1,076,375	470,649	586,872	0 .062	0 .061
2012	1,143,665	499,090	620,659	0 .060	0 .058
2013	1,218,967	534,385	657,716	0 .071	0 .060
2014	1,295,701	569,423	694,653	0 .066	0 .056
2015	1,371,427	604,785	730,634	0 .062	0 .052
2016	1,446,956	640,778	765,964	0 .060	0 .048
2017	1,528,566	678,748	803,930	0 .059	0 .050

Source: Statistics Norway, the role register.

Figure 2.1 Growth in number of board members and boards. Index 2004=1



Source: Statistics Norway.

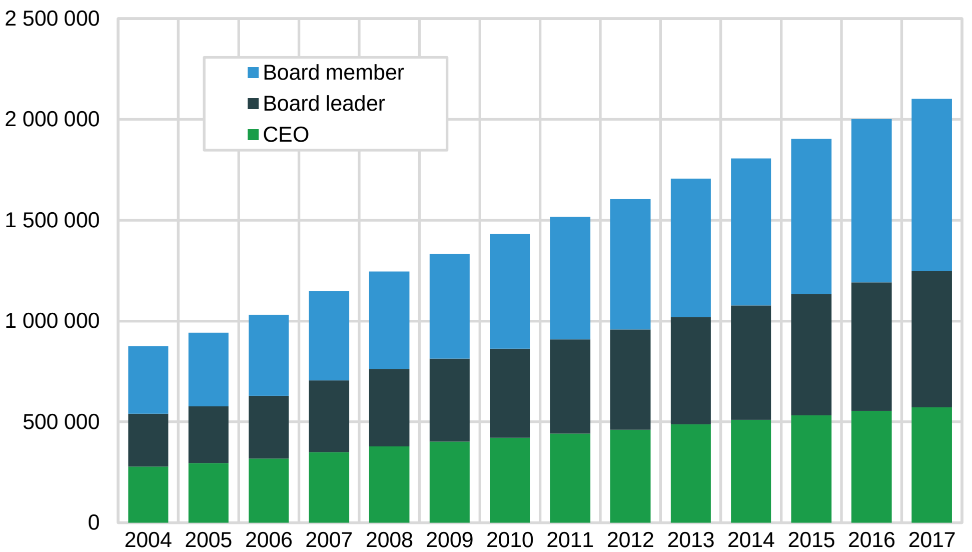
2.2. Data trimming

A typical problem when collecting social network data from secondary sources like databases, is high dimensionality. Due to the networking structure, all relevant data must be analysed together in large matrices, or else there is a risk of distorting important network structures. All choices of reducing dimensionality is therefore done with the research question in mind – whether networks contribute to firm performance.

1. Because the question entails whether board composition affects firm performance in systematic and significant ways, we include only board members and board leaders from the role register. We exclude CEOs, as these have different roles than board leaders and board members and would thus carry information and assert influence in an unsystematic manner compared to the other two types of roles. This means that in total 6,003,216 CEOs are removed from the analysis and 8,220,252 board members and 6,428,006 board leaders are included. The distribution of CEOs, board members and board leaders over the period of our analysis are shown in Figure 2.2.

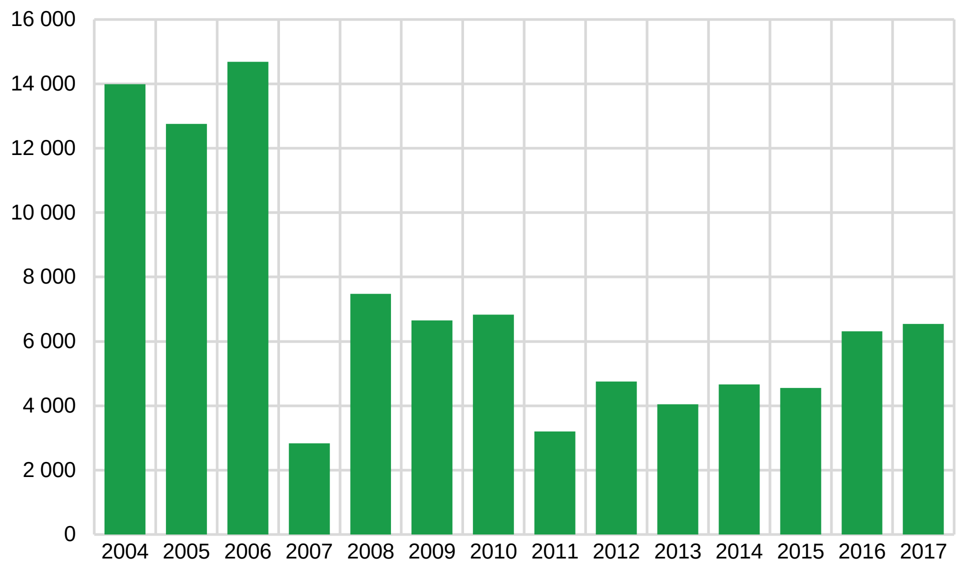
2. Using the accounting register, we filter out inactive enterprises. Inactive enterprises are defined as not having any salary costs *and* no revenues or expenses in the accounting year. To fill in those enterprises that are missing from the accounting register, we also use the national business register as a secondary source, where no turnover and no employees in the given period means that the enterprise is inactive. Figure 2.3 shows the number of enterprises that is taken out of the analysis each year due to inactivity. In total, 99,291 observations were removed from the analysis in this step.

Figure 2.2 Role distribution



[1] CEOs were removed from the analysis.
Source: Statistics Norway.

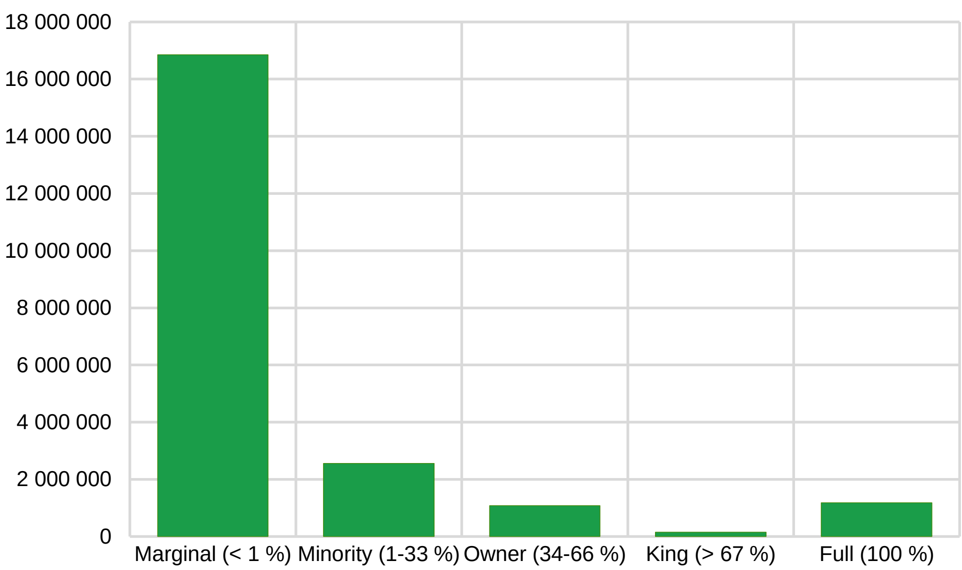
Figure 2.3 Number of enterprises removed from analysis due to inactivity



Source: Statistics Norway.

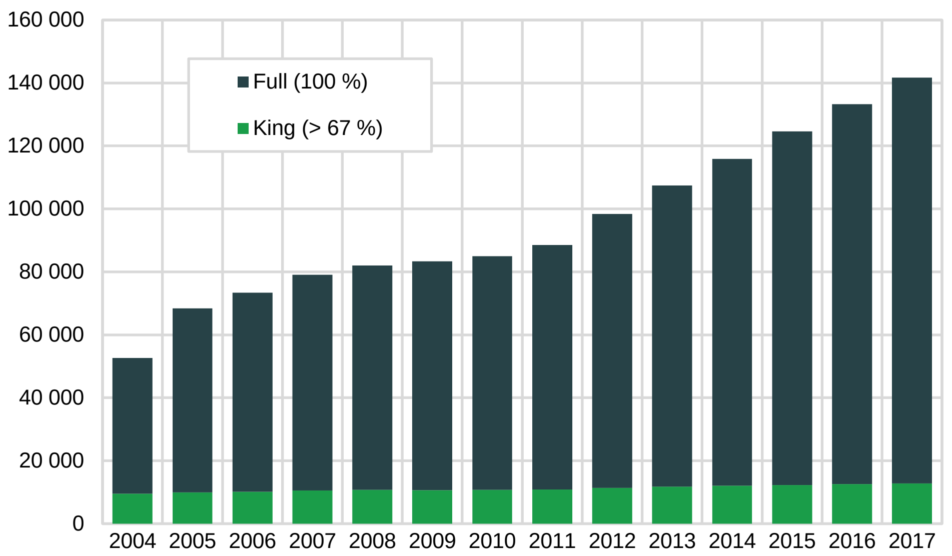
3. Using ownership data from the shareholder register, we filter out shareholders who control more than 67 % of the shares in the enterprise. This is done because we are interested in subtle spillovers and the ability to influence via independent board members, not the ability to overrule and make a unanimous decision on behalf of the board. By excluding majority owners with full control, we eliminate coordinated actions within formal business groups, i.e., within enterprises under full control of one and the same the same individual, from spillover effects within our network of independent enterprises. Especially the ability to capture information in one enterprise and bring it over to another enterprise is in our focus – that is, we are interested in the "small messengers". As shown in Figure 2.4, most shareholders left in our data have less than 1 percent ownership in the enterprise. Kings (67 - 99%) and full owners (100%) are filtered out of the analysis, in total amounting to 156,145 kings and 1,177,572 full owners for all years. As shown in Figure 2.5.

Figure 2.4 Ownership distribution



Source: Statistics Norway.

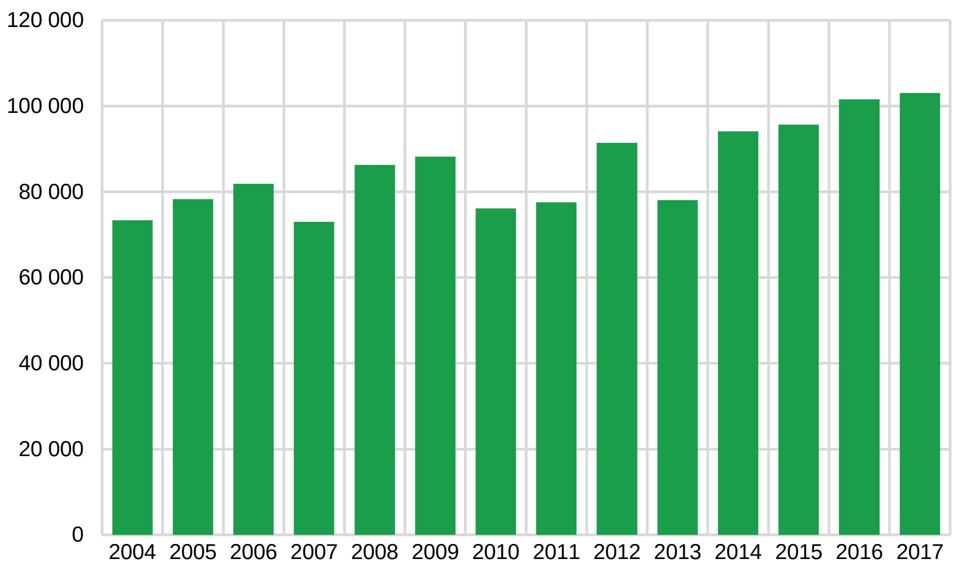
Figure 2.5 Number of owners removed from analysis due to ownership shares



Source: Statistics Norway.

4. We also exclude enterprises that have no ties to other enterprises in the sample. This means that board members who are only located in one board are removed from the data. The reason is that high dimensionality is an issue when working with network structures. Thus, given that we would not lose any information on tie-strength by removing the boards that are not connected, these enterprises are removed. This entails a removal of in total of 1 198 548 year-wise enterprises, as shown in Figure 2.6.

Figure 2.6 Number of isolated nodes removed



Source: Statistics Norway.

After the four-step trimming explained above, the data spans from 2004 to 2017 and contains in total 2,115,108 observations – on average 151,097 observations each year (cf. Table 2.2).

Table 2.2 Sample size after trimming

Year	Boards and board members	Boards	Board members	Growth rate boards	Growth rate board members
2004	137,753	74,339	43,382	-	-
2005	125,634	71,621	40,952	0.037	-0.056
2006	135,860	77,390	43,985	0.081	0.074
2007	129,044	75,608	41,550	-0.023	-0.055
2008	152,060	87,433	48,025	0.156	0.156
2009	158,442	91,129	49,636	0.042	0.034
2010	142,731	83,346	44,557	-0.085	-0.102
2011	144,265	83,785	45,165	0.005	0.014
2012	167,385	95,711	51,590	0.142	0.142
2013	143,218	83,610	44,603	-0.126	-0.135
2014	165,824	96,214	51,699	0.151	0.159
2015	161,839	94,938	51,182	-0.013	-0.010
2016	175,276	103,523	55,022	0.090	0.075
2017	175,777	104,417	54,773	0.009	-0.005

Source: Statistics Norway.

2.3. Memory storage

Social network analyses take a lot of memory and machine power. For our analysis, we used a server with 1 TB RAM and 64 CPU cores. This allowed us to store and work with network matrices in the short-term memory. The alternative would be to create separate codes for batching, which would have been a time-consuming process. Unfortunately, we did not have access to a computer with a strong GPU, which meant that each matrix was slow to generate.

A solution to the issue of machine power, was to modify the matrices so that they took up less space. What we were dealing with, was a set of sparse matrices, which are matrices defined by the fact that they contain very few non-zero elements. A zero in the matrix would imply no tie between boards or board members, which was the norm rather than the exception in our case. Sparse matrices represented by a 2D array leads to a lot of waste in memory storage. Thus, we converted the matrix into a sparse matrix and stored it in Compressed Sparse Column format. This led each matrix to take up significantly less space in R, thus allowing more efficient matrix generation and modification.

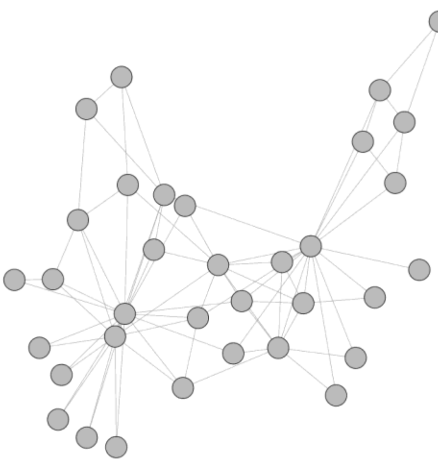
3. Methodology

3.1. Social network analysis

Social networks are interesting to study because many outcomes are dependent on collective structures and relations between actors. One actor may not be able to influence an outcome alone, but when his resources are placed within a certain network structure, he may become more influential. Firm performance may be dependent on whether boards are connected through common board members who are very well informed about trends in the market, or who possesses insider knowledge.

Network structures are primarily composed of two units; nodes (also called verticles) and edges (also called ties or links). Nodes are often individuals, such as board members, but they can also be collectivities, such as the boards themselves. Some nodes are connected through a common edge, which links the nodes together. In this case, we say that the node is adjacent. Chains of ties between nodes creates a connected web that we refer to as a network. In Figure 3.1, nodes are the circles and edges are the lines connecting them. Analysts of networks study either the node level, the dyad (group of two) level or the whole network (Borgatti, Everett & Johnson 2013: 2). This report will present information on both node and network level for Norwegian boards from 2004 to 2017.

Figure 3.1 Example of a network



Source: Zachary (1977).

3.2. Network structure

There are two possible ways to represent network structures mathematically; using either graphs or matrices. The term “graph” comes from graph theory and does not refer to a diagram, but a mathematical object denoting a structure where some pairs of objects are related. An example of a visually illustrated graph is given in Figure 3.1. Another option is to view the data as an adjacency matrix, where a zero in the elements indicates that two nodes are not linked, while a number indicates that the nodes are linked.

The data can be either one-mode or two-mode. In one-mode data, we observe the links between the nodes directly, for example through communication, friendship or family ties between individuals. In two-mode data (also called affiliation data), we have two sets of actors, and the connections are observed between these two types, not within them. Examples include students who attend a class, an activist attending a rally network (Borgatti et al., 2013p. 231). As is typical for network data collected from secondary sources, we have two-mode data for our analysis – in particular, board members sitting in different or common boards. This means that a tie cannot be interpreted as a definite sign of interaction, but an opportunity or increased likelihood of interaction. Data in this structure are represented as an affiliation matrix, consisting of individuals and the organizations they are affiliated to.

Two-mode data can be analysed in two different ways. One possibility is to convert the matrix into one-mode data by multiplying the affiliation matrix by its transpose. In this case, the affiliation matrix becomes an adjacency matrix. This is the strategy chosen in this report. Boards are thus linked by sharing a common board member, and board members are linked by sitting in the same boards. With this type of structure, only one type of actor is analysed as nodes in the network at the time. This gives a weighted adjacency matrix where the weights indicate either how many board members that belong to a specific board (if boards are nodes), or how many boards that board members participate in (if board members are nodes)[4]. These weights can subsequently be interpreted as the strength of a tie.

Converting two-mode data to one-mode data is the most conventional method for analysing affiliation matrices, but there is some dispute on how much information that is lost in analysing the data this way. Two boards can for example be equally strongly linked, but through very different board members. The second option is thus to maintain the two-mode data and create a bipartite network. Here, one converts the affiliation matrix to a large adjacency matrix where both rows and columns consist of both boards and board members. There are, however, a few difficulties with this approach because edges can only occur between the groups, not within the groups. In bipartite networks, boards cannot be linked to other boards, and board members cannot be linked to other board members. This means that one cannot find any cliques, since any edge would be linked indirectly through several nodes, and one would need to use modified centralization scores, since standard measures of centrality assume that all actors could in principle be connected to each other (Borgatti et al., p. 240). Everett and Borgatti (2013) have shown that using one-mode data when analysing networks does not necessarily have to entail a loss of data, as long as both conversions of the matrix – having both board and board members as nodes – are assessed.

[4] To control for differences in board size and the frequency at which board members sit in different boards, we could have used a normalized score, for example by dividing the sums of the rows and columns in the affiliation matrix by their square roots (Borgatti et al., p. 238). A normalization score has not been used in this report for two reasons. First, due to the data trimming noted above, these differences have already been adjusted systematically. Second, normalization would obscure the network influence reflected by being a large board, or having an active board member, which are characteristics that may be important to improve firm performance, regardless of whether they reflect grouped skewedness in the data.

4. Descriptive statistics

4.1. Network level

Table 4.1 shows four measures that characterize the board network over time.[5] The network characteristics are based on the data as described in Section 2. Thus, since isolated nodes were removed, the measures describe the segment of boards that are linked, either weakly or extensively, to other boards. It describes the part of the business world in Norway that is socially integrated.

Table 4.1 Network characteristics

Year	Average degree	Shortest average path	Transitivity	Centralization
2004	6.778	8.697	0.830	1.133
2005	6.065	9.222	0.858	1.062
2006	6.361	9.241	0.863	1.254
2007	6.358	9.443	0.830	0.825
2008	7.382	9.087	0.882	1.610
2009	7.734	9.104	0.883	1.575
2010	7.451	9.410	0.879	1.378
2011	7.808	9.483	0.898	1.592
2012	8.354	9.227	0.909	1.889
2013	7.834	9.601	0.903	1.679
2014	7.926	9.521	0.878	1.461
2015	7.551	9.870	0.866	1.368
2016	7.869	9.901	0.880	1.571
2017	7.979	10.049	0.884	1.615

Source: Statistics Norway.

The first measure – average degree – measures the cohesion of the network. Strongly cohesive networks are tangled, meaning that many of the nodes in the network have ties to other nodes. Density is the typical measure used to describe cohesiveness, but in large networks, density numbers can become extremely low. The reason for this is that density reflects the proportion of ties in the network compared to all possible ties. Thus, if there are many nodes that could theoretically have tie, but does not in reality have a tie, density is reduced to a very low number. This type of structure is typical for large networks where it is not obvious that the nodes should be linked, as in this case, where boards are, in fact, rarely linked through a common board member. Average degree is generally a better measure in large and untangled networks. The average number of ties of the nodes has risen since 2004, indicating that in general, enterprises in the network have gotten more board members in common.

Second, shortest average path measures the length of the path between two nodes with the minimum number of edges in between. It is also called the geodesic distance. In our case, the graph is weighted by the number of people connecting a board, so that the shortest average path is the minimum sum of weighted edges between two nodes. It can be interpreted as an index of the time taken for information to travel from one node to another, provided that the information always travels via the shortest path. This measure has also risen since 2004, indicating that the nodes in the increased by time network (jf. Table 2.2) have become less connected, so that information travel now slower.

Third, the transitivity measure reflects the clustering of the network. It measures the probability that the nodes which a node is connected to, also are connected among themselves. In doing this, it looks for triadic configurations – nodes organized in a triangle. Then it takes the number of transitive dyads and divides it by the sum of the number of transitive and intransitive dyads. This measure is

[5] All measures in this chapter are calculated for entirely connected enterprises, i.e. the network without isolated nodes. The analysis in chapter 5 is provided for an extended sample of enterprises that also comprises not connected nodes with zero network measures for them.

sometimes called the “clustering coefficient”, because it reflects the extent of small groups in the network. The clustering coefficient has increased somewhat, i.e. from 83 percent in 2005 to 88 percent in 2017. In general, this measure has varied little in the period of our analysis, i.e. between 83 and 90 percent.

Fourth, the centralization measure shows the starshaped-ness of the network. In very centralized networks, one or a few nodes are central to the structure, having many ties to other nodes, which typically have fewer ties. This centralization is based on each node’s degree. It then takes the node with the highest number of edges, and subtracts all other nodes’ edge number from this, finally summing the numbers up. To illustrate, if the max number of edges for a node in the network was 10, and the other nodes in the network had 3, 4 and 8 edges respectively, the centralization score would be: (10-3) + (10-4) + (10-8) = 15. To make the comparison of the measures easier, they have also been normalized by dividing the values by 10,000,000. Centralization has increased throughout the period of the analysis, showing that the network structure has increasingly been dominated by a few enterprises.

Overall, the network statistics tell a story of a network that has become more populated by board members who have several seats in different boards. Boards appear to be more widely spread throughout the network, as indicated by the increasing shortest path and transitivity. At the same time, more ties reach one single enterprise or cluster in the end of the 2010s than the beginning of the 2000s.

4.2. Node level

Node level measures are used to describe individual nodes in a network. They typically refer to centrality in some form, which is a property of a node’s position in a network. Central nodes are often interpreted as having an advantage by its position in a network. However, this advantage might take many forms. A node could for example be central by having access to a lot of information or being the sole source of information that goes from one actor to the next. Table 4.2 shows the average node statistics for each year.

Degree centrality is, first and foremost, a node property. It measures the number of ties that a node has. A high degree centrality means that a board is connected to many other boards by sharing board members. As shown, boards tend to have more ties to boards in 2017 than they did in the early 2000s.

Betweenness centrality is a measure on how often the node falls between the shortest path between two other nodes. It reflects how large a proportion of the networks’ shortest paths from one node to another that passes between the focal node. A high betweenness score for a board means that it often falls between the shortest path between two other nodes. The measure is often interpreted as the ability of a node to control the flow of information or playing a gatekeeping role. It has remained relatively stable throughout the period of our analysis, indicating that no enterprises have become particularly focal in information sharing.

Table 4.2 Node characteristics

Year	Average degree	Average betweenness	Average eigenvector centrality
2004	6.778	92297	0.002091
2005	6.065	78172	0.002104
2006	6.361	81938	0.002050
2007	6.358	70509	0.000764
2008	7.382	87579	0.001851
2009	7.734	91872	0.001792
2010	7.451	78108	0.000996
2011	7.808	77801	0.001923
2012	8.354	96410	0.001770
2013	7.834	70435	0.002068
2014	7.926	88683	0.001501
2015	7.551	86439	0.000374
2016	7.869	95181	0.001282
2017	7.979	92372	0.001262

Source: Statistics Norway

Eigenvector centrality is a measure of the node's centrality in comparison to how central its adjacent nodes are. In the case where board A is connected to another board B that has a lot of ties further, board A receives a high eigenvector score. Thus, a board with a high eigenvector could be understood as a board who is connected to many well-connected boards. The eigenvector score has been reduced somewhat since the 2000s, but not markedly. In general, the eigenvector score varies a lot throughout the period of the analysis. This is because this score is sensitive to whether board members shift between boards. Since the score is calculated on whether a board member has ties to other well-connected boards, the score is multiplied extensively when well-connected board members are part of the same board. This has a significant implication on the eigenvector score (Wilmet, 2017, p. 51).

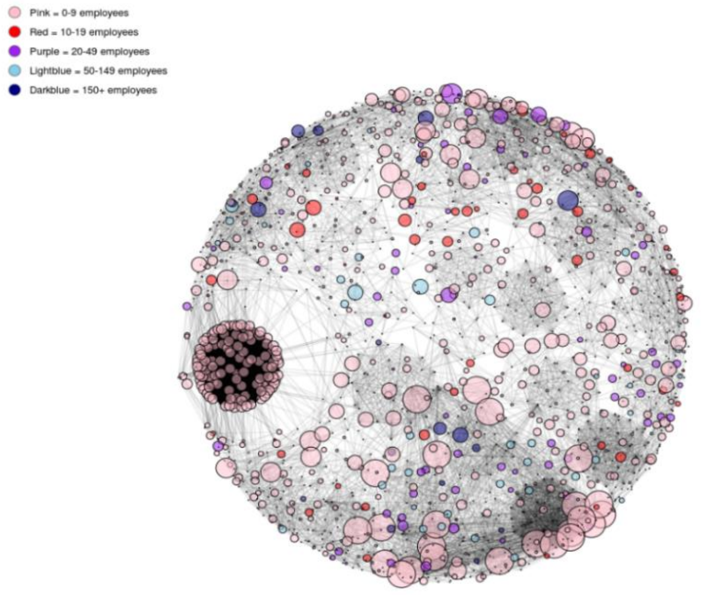
4.3. Visualizations

Networks are especially well suited for visual representations. In the visualizations below, the dots refer to the nodes in the data, meaning the enterprises. Lines are edges connecting the enterprises. The thickness of the edges is the same for all enterprises, a choice made in order to make the plot tidier and more interpretable. The colours of the nodes indicate enterprise size, while the size of the nodes indicates the size of the boards – that is, how many board members there are in the board. We present her plots for board networks in 2005 and 2017, i.e. the first (complete) year and in the final year of the role register.[6]

It is important to note that these graphs do not give a complete representation of the network structure. Nodes with less than three degrees and two edges were removed before generating the plot. This was a necessary step be able to generate the plot at all, and to make the plot interpretable. Thus, the graphs show only the most well-connected boards. The minimum number of degrees in 2005 was 1, and the maximum was 159. The minimum number of edges was 1, and maximum was 13. This means that of 74,437 enterprises, only 1794 meet our criteria and are represented in the graph for 2005 (jf. Figure 4.1.). In 2017, the minimum and maximum number of nodes was the same as in 2005, while the minimum number of edges was 1 and the maximum was 16. In 2017, 2471 out of 103,922 enterprises are represented in the graph (jf. Figure 4.2)

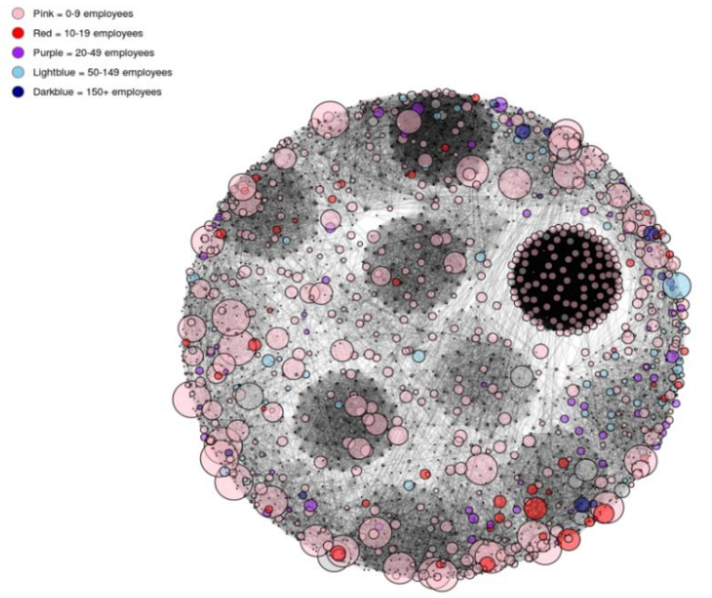
[6] While the data from the role register are available since 2004, the first-year data are incomplete.

Figure 4.1 Network structure for selected nodes, 2005



[1] The node size indicates size of the board.
[2] The Kamada Kawai algorithm is used.
[3] Only nodes with at least 3 degrees and 2 edges were included.
Data source: Statistics Norway.

Figure 4.2 Network structure for selected nodes, 2017



[1] The node size indicates size of the board.
[2] The Kamada Kawai algorithm is used.
[3] Only nodes with at least 3 degrees and 2 edges were included.
Data source: Statistics Norway.

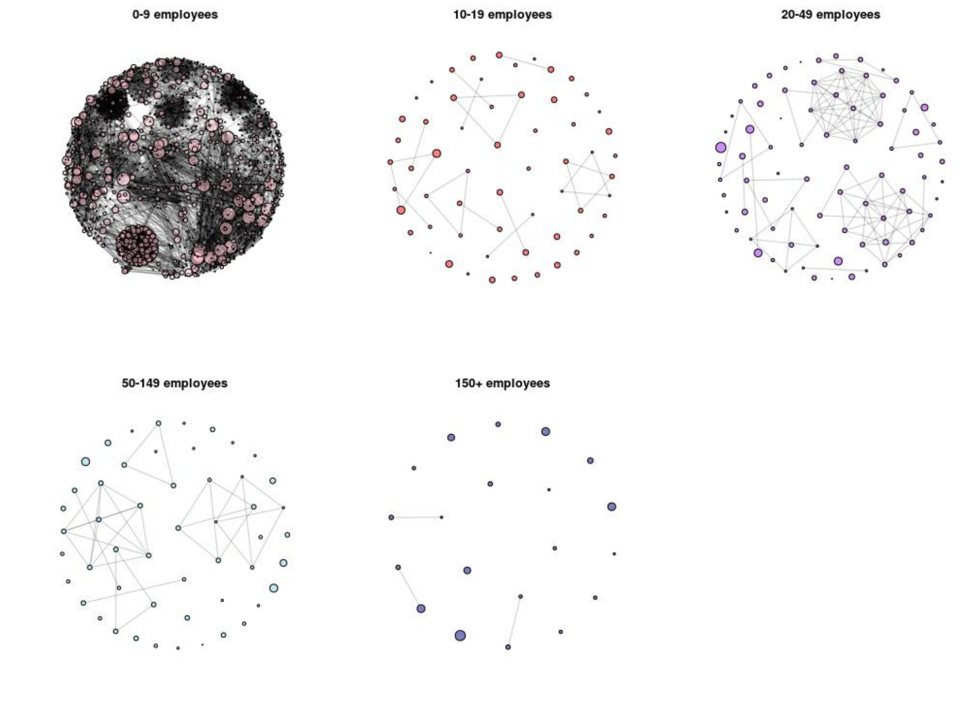
The graphs utilize the Kamada-Kawai algorithm for the layout of the plot. This layout has been designed to perform well in large, unconnected networks (Kamada and Kawai, 1989). In this layout, large weights on edges results in larger distances. Thus, enterprises that are connected by having many board members in common, are located far from each other in this layout. The alternative, perhaps more intuitive, would be to employ the Fruchterman and Reingold algorithm (Fruchterman and Reingold, 1991). However, this layout proved to be difficult to read and almost impossible to interpret. Thus, the Kamada-Kawai algorithm is chosen.

Comparing Figure 4.1 and Figure 4.2 provide much the same information as the network characteristics reported in Table 4.1, even though the figures do not show the complete network structure. There are more lumps in the 2017 network, as indicated by the increasing transitivity in Table 4.1.

The graphs also show that the networks are mostly comprised of small enterprises. These small enterprises often form clusters, where board members of small enterprises frequently sit in the same boards. Some of these small enterprises with 1-9 employees have rather small boards, but this group also has the largest boards in the sample. Interestingly, from the graphs in Figure 4.1 and Figure 4.2, the largest enterprises in terms of employees do not appear to have a strikingly central position.

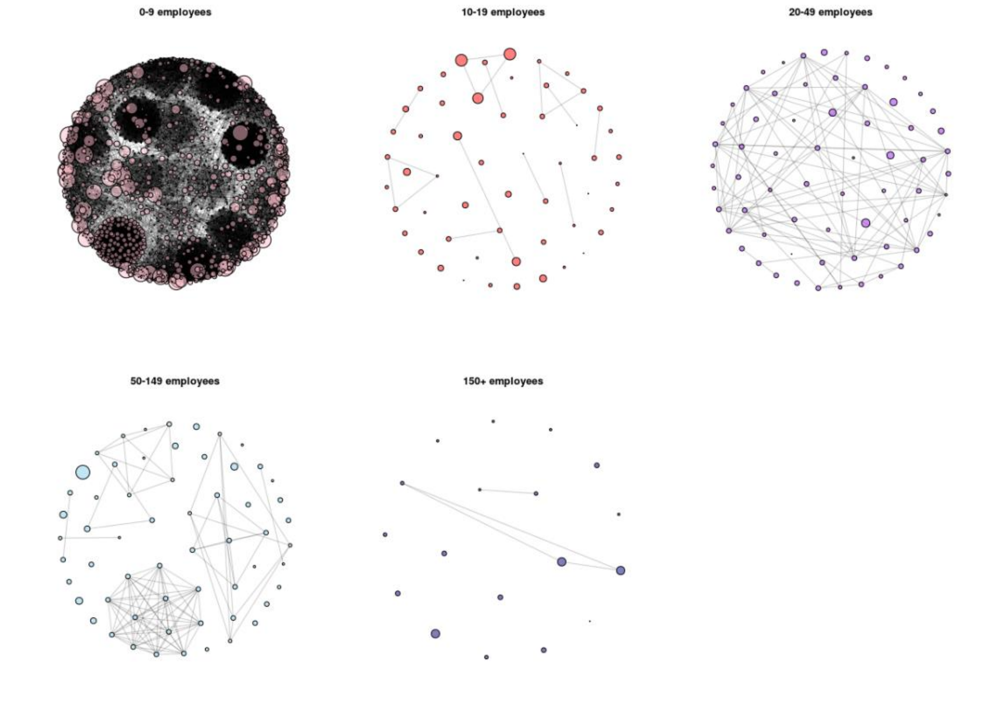
It is also interesting to look at subsets of the network to see if there are any differences in the network structure between size groups and between years. These graphs are shown in Figure 4.3 and Figure 4.4 for 2005 and 2017 respectively, where the networks are grouped by enterprise size. The lumpy structure in the smallest size group becomes particularly apparent with this decomposition.

Figure 4.3 Network structure by size group, 2005



[1] The Kamada Kawai algorithm is used.
[2] Only nodes with at least 3 degrees and 2 edges were included.
Data source: Statistics Norway.

Figure 4.4 Network structure by size group, 2017



[1] The Kamada Kawai algorithm is used.
[2] Only nodes with at least 3 degrees and 2 edges were included.
Data source: Statistics Norway.

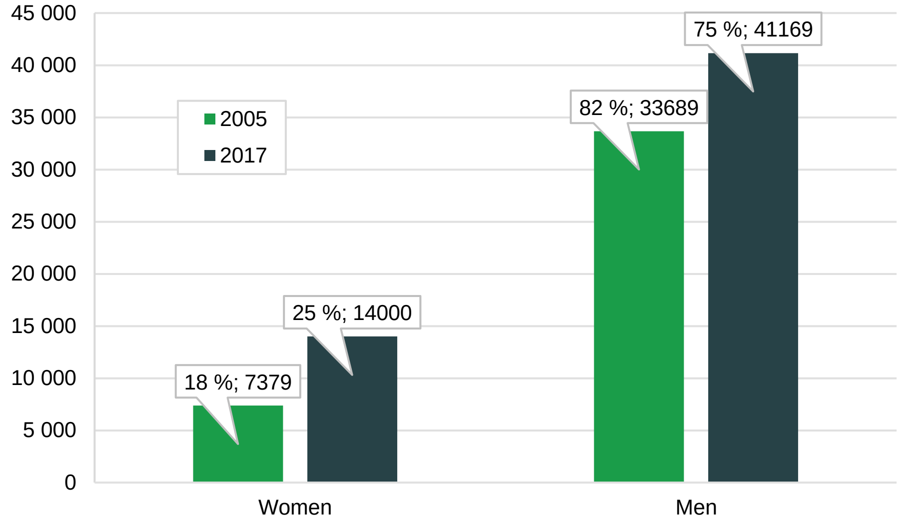
4.4. Board members by gender

So far, we have focused on enterprises within the network and treated board members as links. However, another important aspect of the network is the board members, and their connectedness through their common board membership. This illustrates a different, but related way of looking at the network: what drives the formation of networks?[7] Social class? Education? One relevant question to ask when focusing on board members, is whether men or women are more well-connected within board networks. This is a question we briefly turn to in this section.

As shown in Figure 4.5, there were notably more men than women in the board network in 2005 and 2017. However, the share of women in the board network increased from 18 per cent in 2005 to 25 per cent in 2017. Table 4.3 and Table 4.4 show density measures for men and women in 2005 and 2017 respectively. Since density can be interpreted as the probability that a tie exists between any set of nodes, it is no surprise that the density is very low in these large networks. Nevertheless, density is higher for women in both periods, showing that women may have higher capability to be connected by using more of their potential networking than men. However, this could also possibly be an effect of women's network being smaller than men's network. The capability of man and woman to be connected was also higher than capability of man and man to be connected in 2005. These densities become almost equal in 2017. The fact that all densities decreased by the time is due to a huge increase in the network size, and hence, in the number of all possible connections compared to the increase in observed connections.

[7] Hermalin and Weisbach (1988)

Figure 4.5 Distribution of the number of men and women among board members, 2005 and 2017



Source: Statistics Norway

Table 4.3 Density among gender groups of board members, 2005.

	Men	Women
Men	10.8E-05	-
Women	13.2E-05	58.5E-05

Source: Statistics Norway

Table 4.4 Density among gender groups of board members, 2017.

	Men	Women
Men	8.1E-05	-
Women	8.2E-05	27.5E-05

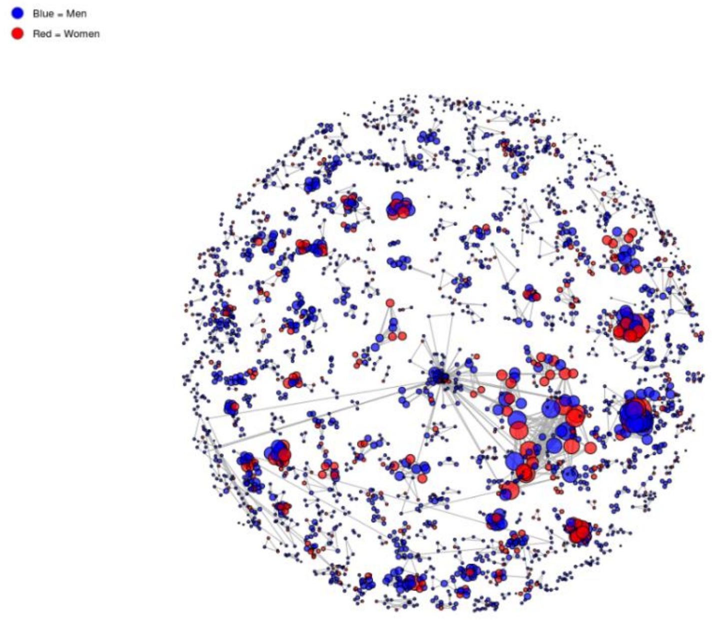
Source: Statistics Norway

Figure 4.6Figure 4.7 graph the network structure for board members by their gender in 2005 and 2017 respectively. Here, the dots indicate people who are sitting in boards, and the links are ties between people who sit in the same boards. This plot allowed for a more inclusive selection, since there were fewer board members in the sample. Isolates, meaning nodes with 0 degrees, were removed, and the nodes required 2 edges to be included, so that each of the board members in this plot sit in at least two boards together. Again, the Kamada-Kawai algorithm is used, so that nodes that lie far from each other have board members that sit in many boards together. The size of the node is their degree, meaning that larger nodes have more ties to other nodes. The colours indicate gender, where blue is for men and red is for women.

We see that the networks are populated by mostly blue nodes, but the difference regarding size is not striking. Many of the well-connected female nodes have more degrees, indicating that some women tend to sit in many boards. The selection of the edge criteria may also skew these plot representations. As we saw, women have higher density measures, meaning that they often share a board membership with someone else. Thus, when people who are only connected through one common board membership but connected to other boards are excluded, several men might lose their place in the plot.

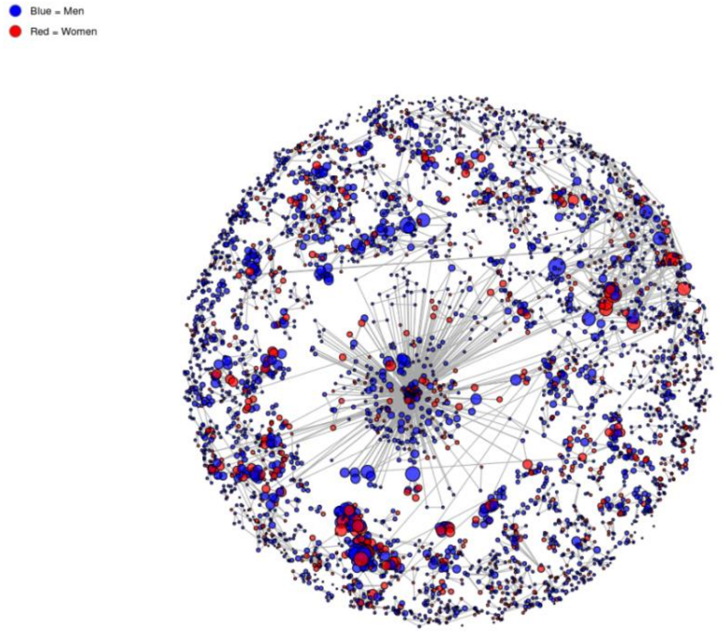
The network of well-connected board members has become more populated and more connected from 2005 to 2017. In 2005, there were many small clusters, and several of these had no tie to each other. In 2017, one clear cluster of board members exists, and persons in this cluster have many ties far out in the network, meaning that they are connected strongly to other actors by having several common board memberships.

Figure 4.6 Men and women network structure, 2005



[1] The node size indicates size of the board.
[2] The Kamada Kawai algorithm is used.
[3] Only nodes with at least 1 degree and 2 edges were included.
Source: Statistics Norway

Figure 4.7 Men and women network structure, 2017



[1] The node size indicates size of the board.
[2] The Kamada Kawai algorithm is used.
[3] Only nodes with at least 1 degree and 2 edges were included.
Data Source: Statistics Norway

5. Application: Centrality and R&D activity

As discussed in the introduction, board members with access to information about new technologies or market opportunities can represent a value to the enterprise by being a part of enterprises' relational capital. Along with an enterprise's knowledge and human capital, it can then have a positive influence on R&D and innovation activities in the enterprise. While the impact of knowledge and human capital on R&D input and output is well studied, the evidence on the impact of enterprises' relational capital is still scarce. The examples mentioned in the introduction have in common that even if they find a positive correlation between network centrality measures and R&D activity, we cannot see that neither of them identifies a causal relationship. Of course, this is not an easy task, nor is this the ambition in our report which first and foremost has a descriptive purpose. Besides, an a priori assumption that a causal relation may exist, through network as pathways for information and transfer of resources is not implausible.

Below, we present an empirical analysis based on our network data and data on enterprises' R&D spending. Our research question is whether R&D spending is related to network participation. Like previous studies, we cannot identify exactly which process that drives the results. For example, we cannot say for sure whether networking lead to increased R&D through transfer of knowledge or if there is some unobserved factor that simultaneously affects both an enterprise's centrality in the network and its R&D. However, we still introduce some improvements in the methodology compared to previous studies, namely for accounting sample selection bias. This is an important issue that should be considered when analysing enterprise R&D behaviour. Only a fraction of the enterprise population invests in R&D, whereas many enterprises are not engaged in R&D activities at all. Moreover, enterprises that already are engaged in R&D activities have a larger probability of doing R&D again and innovating than other enterprises do. Hence, regressing R&D intensity on different network measures ignoring selection problems may lead to seriously biased estimates. Most studies presented in the introduction do not take selection problem into account, while our results show that selection parameters are significant and cannot be ignored.

5.1. Model for R&D input decision

Here we describe a model for enterprise *i*'s decision to engage in R&D activities in period *t*. First the enterprise decides whether to start to invest in R&D in the given period; if it decides to invest, the enterprise then sets the amount of R&D investments. This statement of the problem can be modelled with a standard sample selection model (see Heckman, 1979):

$$rd_{it} = \begin{cases} 1 & \text{if } rd^*_{it} = x^{rd}_{it}\alpha_1 + e_{it} > c \\ 0 & \text{else} \end{cases}, \tag{1}$$

where rd_{it} is the observed binary endogenous variable equal to zero for non-R&D and one for R&D-performing enterprises, rd^*_{it} is a corresponding latent variable that expresses some decision criterion, such that an enterprise decides to invests in R&D if rd^*_{it} is above a certain threshold *c*, x^{rd}_{it} is a vector of enterprise characteristics (e.g. size, age, international orientation etc., and a constant term), α_1 is the associated coefficient vector, and e_{it} is an error term.

Once an enterprise has decided to engage in R&D activities, it must set the amount of resources devoted to R&D investments. Analogous to the previous equation and in line with the strand of literature using the so-called CDM model (Crepon et al., 1998)[8], the latent R&D intensity of enterprise *i* in a given period *t*, r_{it}^{*}, is represented as a function of another set of enterprise characteristics, x_{it}^{r}:

(2) $$r_{it}^{*} = x_{it}^{r}\alpha_2 + \varepsilon_{it} ,$$

where α_2 is the associated coefficient vector, and ε_{it} is an error term. The observed R&D intensity, *r*, is then equal to:

(3) $$r_{it} = \begin{cases} r_{it}^{*} & \text{if } rd_{it} = 1 \\ 0 & \text{else} \end{cases} .$$

The pair of random disturbances e_{it} and ε_{it} is assumed to be jointly i.i.d. normally distributed, with zero mean and covariance matrix given by

(4) $$\begin{pmatrix} 1 & \rho\sigma_{\varepsilon} \\ \rho\sigma_{\varepsilon} & \sigma_{\varepsilon}^{2} \end{pmatrix},$$

where σ_e and σ_{ε} are the standard errors of e_{it} and ε_{it}, $\sigma_e = 1$ by standardisation, and ρ is their correlation coefficient. This model can be estimated by maximum likelihood.

An econometric specification for enterprise R&D input decision contains then two R&D equations corresponding to the theoretical model (1)–(4):

(5) $$rd_{it}^{*} = x_{it}^{rd}\alpha_1 + e_{it} ,$$

(6) $$r_{it}^{*} = x_{it}^{r}\alpha_2 + \varepsilon_{it} ,$$

where the first equation models the propensity that an enterprise with characteristics x_{it}^{rd} engages in R&D activities. It is estimated for the whole sample of enterprises. The second equation focuses only on enterprises with positive R&D investment, *R >0*, and is employed to study how the R&D intensity of the enterprise, r_{it}^{*}, is affected by a set of enterprise characteristics x_{it}^{r}. In our case both x_{it}^{rd} and x_{it}^{r} include the measure on how well-connected the enterprise is through the board network.

5.2. Additional data sources and descriptive statistics

For the analysis, we use a rich enterprise-level panel data set based on annual R&D surveys collected by Statistics Norway. The enterprises included in the surveys constitute a large and representative sample of the Norwegian private sector. The

[8] The standard version of the CDM model is a structural model that studies the following interrelated stages of the innovation chain: the choice by an enterprise of whether or not to engage in R&D; the amount of resources it decides to invest in R&D; the effects of these R&D investments on innovation output; and the impact of innovation output on the productivity of the enterprise. In this report we test only the possible impact of being well-connected through the board networks on the enterprise's R&D decision retaining the deeper analysis for further investigation.

enterprises with 10–50 employees are selected using a stratified sampling method based on industry classification (SIC 2007 codes) and enterprise size, whereas all enterprises with more than 50 employees are included. These data are then supplemented with information on board network measures from the dataset described above. For the analysis, we also include the "isolated enterprises", i.e. those enterprises that are not connected to others through the board network and set all network measures for them to be equal to zero. Excluding enterprises with incomplete information on key variables (very few observations) or with extreme observations for R&D intensity (1 percent from each side), we obtain an unbalanced panel of 34,398 observations on 7,813 enterprises.

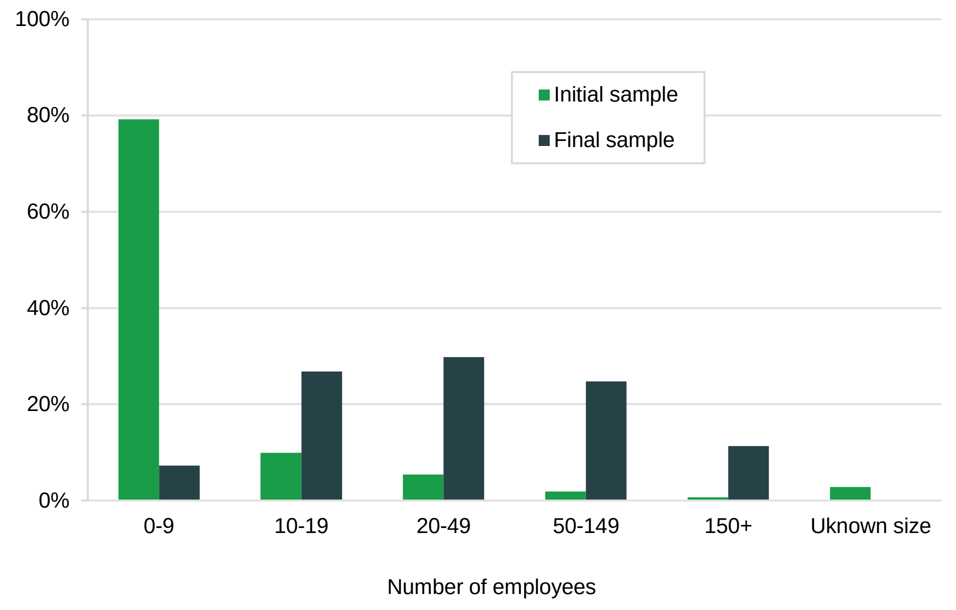
We use the following enterprise characteristics in the analysis:

- *Centrality indicator (eigenv_top10):* a dummy variable indicating whether an enterprise is in the upper decile of the distribution of eigenvector centrality for the given year.
- *R&D investment (R)*: R&D investment as it is reported in the questionnaire, deflated by the R&D deflator used in the national accounts (using 2005 as a reference year).
- *R&D intensity (r):* is the R&D investment per employee, *R/L*, where *L* is the number of employees.
- *Positive R&D history (d_lag_rd)*: a dummy variable indicating whether an enterprise has carried out R&D in the previous year.
- *Enterprise size*: a set of dummy variables indicating the enterprise size, i.e. 0-9, 10-19, 20-49, 50-149 or 150 employees or more. The latter category (large enterprises) is the reference category. An alternative measure *log(L)*, where *L* is the number of employees, is used for checking the robustness of the results.
- *Enterprise age*: a set of dummy variables indicating the enterprise age, i.e. 0-2, 3-5, 6-9, 10-15 or 16 years old and older. The latter category (mature enterprises) is the reference category.
- *Enterprise industry*: a set of dummy variables indicating the enterprise industry at the industry group level. Manufacturing (code 10-33 in SIC2007) is the reference industry.
- *Year*: a set of time dummies indicating the year of the R&D survey; 2005 is the reference year.

Distribution of the final sample across enterprise size groups, age groups and industries are provided in Figure 5.1-Figure 5.3 respectively. From Figure 5.1 we observe that compared to the distribution of the initial sample of enterprises that is solely based on the role register and includes "isolated enterprises", the very smallest enterprises (with 0-9 employees) are poorly covered by R&D survey data. This is not a surprise given the survey structure. However, more than half of our final sample comprises small (10+) enterprises that gives us opportunity to study the impact of well-connectedness for this group with high confidence. The fact that data on board membership for SMEs have so far not been available in many countries also demonstrates the importance of the results that we present further.

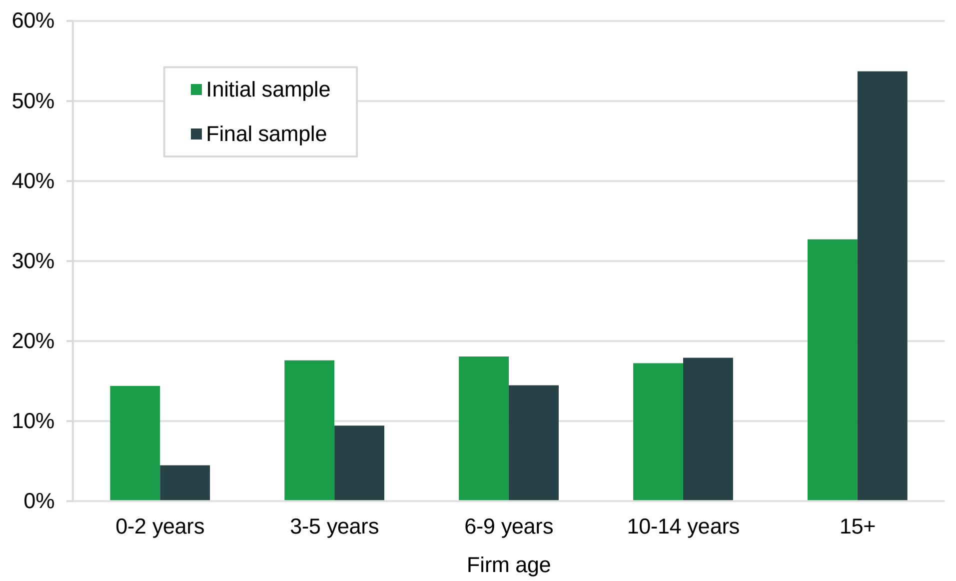
As to other enterprise characteristics, mature enterprises and enterprises in Manufacturing (10-33 in SIC 2007) are overrepresented in the final sample while young enterprises and enterprises in Other services (68-82 in SIC 2007) are underrepresented. Other age and industry groups are well-represented in the final sample (jf. Figure 5.2 and Figure 5.3).

Figure 5.1 Distribution of enterprises in initial and final sample by enterprise size



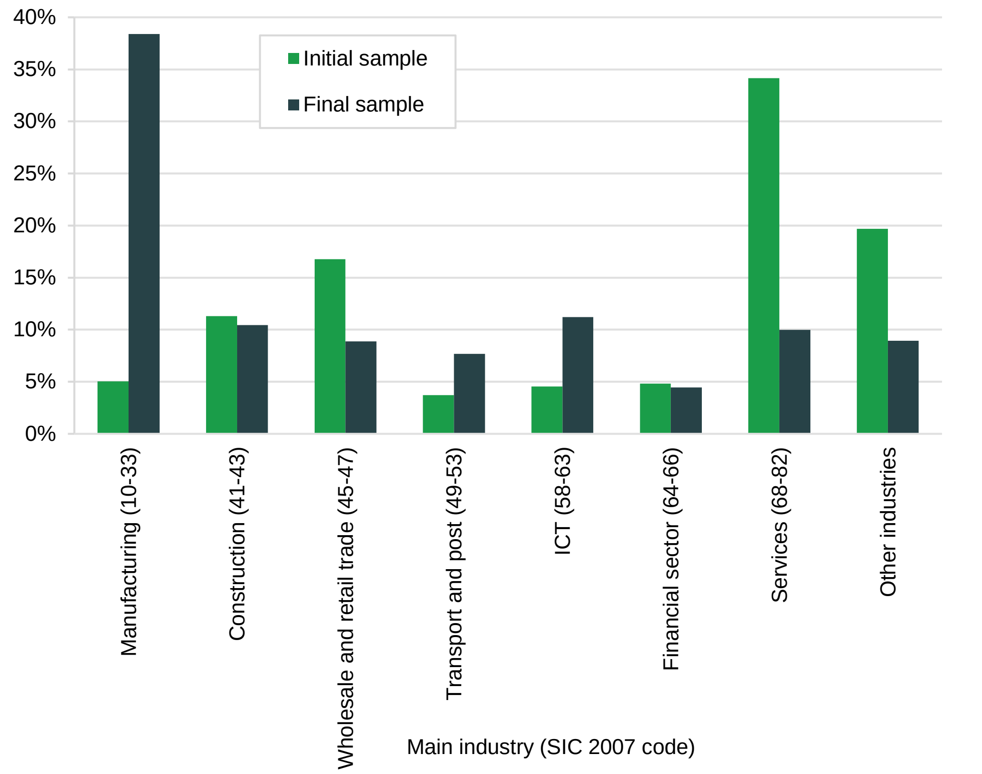
Source: Statistics Norway

Figure 5.2 Distribution of enterprise-year observations in initial and final sample by enterprise age[1]



[1] We use enterprise-year observations instead of enterprises for calculation of shares here since enterprise age changes from year to year and many enterprises are presented only in some years in the R&D survey during the observational period.
Source: Statistics Norway

Figure 5.3 Distribution of enterprises in initial and final sample by enterprise main industry



Source: Statistics Norway

We can see that mean degree centrality over all periods is increasing by enterprise size in both samples. The same is the case for betweenness centrality measure. These findings are expected since larger enterprises tend to have larger boards that potentially leads to higher degree and betweenness.

Mean eigenvalue centrality is also increasing by enterprise size in the final sample, while this is not the case in the initial sample. The main reason for such behaviour is that eigenvector centrality of a node is influenced by the eigenvector centrality of all the adjacent nodes, such that a node that is connected to other well-connected nodes will have a higher eigenvector centrality. This node will in its turn increase the eigenvector centrality of its adjacent nodes. This feedback triggers what can be called a snowball effect. That implies that even a very small board can have a high eigenvalue centrality if any of the board members is sitting in other boards with high eigenvalue centrality. This measure is very volatile among small enterprises in the initial sample. By having relatively few of the smallest enterprises in the final sample we obtain more stable mean values for eigenvalue centrality.

As to other enterprise characteristics, more mature enterprises have higher values of centrality measures than young enterprises. That is not surprising since it takes time to establish new relationships. Enterprises in Financial service activities (64-66 in SIC 2007) are leading in all centrality measures.
Table 5.1 presents summary statistics of our centrality measures for Norwegian board network by enterprise size, age and industry both in the initial and the final sample.

We can see that mean degree centrality over all periods is increasing by enterprise size in both samples. The same is the case for betweenness centrality measure. These findings are expected since larger enterprises tend to have larger boards that potentially leads to higher degree and betweenness.

Mean eigenvalue centrality is also increasing by enterprise size in the final sample, while this is not the case in the initial sample. The main reason for such behaviour is that eigenvector centrality of a node is influenced by the eigenvector centrality of all the adjacent nodes, such that a node that is connected to other well-connected nodes will have a higher eigenvector centrality. This node will in its turn increase the eigenvector centrality of its adjacent nodes. This feedback triggers what can be called a snowball effect. That implies that even a very small board can have a high eigenvalue centrality if any of the board members is sitting in other boards with high eigenvalue centrality. This measure is very volatile among small enterprises in the initial sample. By having relatively few of the smallest enterprises in the final sample we obtain more stable mean values for eigenvalue centrality.

As to other enterprise characteristics, more mature enterprises have higher values of centrality measures than young enterprises. That is not surprising since it takes time to establish new relationships. Enterprises in Financial service activities (64-66 in SIC 2007) are leading in all centrality measures.

Table 5.1 Centrality measures by sample and enterprise characteristics. Mean values

		Initial sample			Final sample	
Enterprise characteristics	Degree	Eigenvalue	Between-ness	Degree	Eigenvalue	Between-ness
Enterprise size						
0-9 empl. (S_0_9)	4.43	0.000997	41695	3.99	6.66E-07	95610
10-19 empl. (S_10_19)	4.53	0.000758	65725	4.65	5.84E-06	117051
20-49 empl. (S_20_49)	7.34	0.001438	115790	6.14	1.13E-06	149675
50-149 empl. (S_50_149)	8.62	0.000082	222987	8.02	0.000040	240174
150+ empl. (S_150plus)	9.64	0.000094	439764	9.78	0.000051	476611
Unknown	6.50	0.001494	59886	-	-	-
Enterprise age						
0-2 years	4.33	0.000338	33977	6.53	2.47E-07	147899
3-5 years	5.06	0.000681	48333	6.78	1.54E-06	198683
6-9 years	5.47	0.001621	56319	6.84	1.68E-06	219814
10-14 years	4.72	0.001052	57690	6.95	7.72E-06	236556
15+ years	4.33	0.001027	63657	6.57	0.000034	207581
Enterprise industry (SIC 2007 code)						
Manufacturing (10-33)	3.86	3.80E-06	60684	6.20	0.000014	139556
Construction (41-43)	4.46	0.000115	39536	6.07	8.53E-09	133563
Wholesale and retail trade (45-47)	2.95	4.59E-06	20420	5.56	9.96E-08	108666
Transport and post (49-53)	6.44	0.000032	74023	8.52	1.68E-06	232952
ICT (58-63)	3.32	0.000016	73043	6.54	3.85E-06	248437
Financial services (64-66)	6.11	0.002645	107025	9.21	0.000186	729734
Other services (68-82)	5.76	0.000468	48826	5.73	0.000060	229327
Other industries	4.46	0.003505	80938	9.06	3.42E-06	336508

Source: Statistics Norway

5.3. Identification strategy and empirical results

One important issue that should be considered when analysing enterprise R&D intensity is the possibility of sample selection bias. Only a fraction of the enterprise population invests in R&D, whereas many enterprises in the sample are not engaged in R&D activities at all (only 33 per cent of the observations in our final sample have positive R&D values). In line with the many CDM empirical studies, we correct for the selection bias by estimating (5) and (6) as a system of equations by maximum likelihood, assuming that the error terms in (5) and (6) are bivariate normal with zero mean and covariance matrix as specified in equation (4). In the literature, this model is often referred to as a Heckman selection model (see Heckman, 1979) or type II Tobit model (see Amemiya, 1984).

For identification of such a model, the vector x_{it}^{rd} in equation (5) should contain at least one variable that is not in the vector x_{it}^{r} in equation (6). Nevertheless, most previous works in the CDM literature use the same explanatory variables in both equations. The main reason for this practice is that it is difficult to find the factors explaining an enterprise's likelihood of engaging in R&D that are not related to the amount of resources the enterprise decides to invest in R&D. Following Rybalka (2015), we use here a dummy variable for the enterprise's previous R&D investment (whether an enterprise had an R&D activity in the previous year) as an exclusion restriction.[9] The main results are presented in column (1)-(2) of Table 5.2. In addition to this exclusion restriction we use the optional identification by functional form to check the robustness of the results. We then include employment (log) and employment squared (log) instead of set of size dummy variables, cf. column (3)-(4) of Table 5.2.

Table 5.2 Estimation results for impact of network well-connectedness on enterprise R&D input decision

Variables[1]	(1)	(2)	(3)	(4)
Dependent variable: R&D expenditures per employee (log)				
eigenv_top10	0.048		0.071*	
eigenv_top10 x S_0_9		0.116		0.033
eigenv_top10 x S_10_19		0.165**		0.168**
eigenv_top10 x S_20_49		0.246***		0.331***
eigenv_top10 x S_50_149		-0.062		-0.121*
eigenv_top10 x S_150plus		-0.004		0.084
S_0_9	2.473***	2.436***		
S_10_19	1.858***	1.814***		
S_20_49	1.259***	1.194***		
S_50_149	0.385***	0.396***		
Employment (log)			-1.603***	-1.528***
Employment squared (log)			0.115***	0.108***
Dependent variable: Dummy for R&D>0				
eigenv_top10	0.045**		0.038*	
eigenv_top10 x S_0_9		-0.022		-0.094
eigenv_top10 x S_10_19		0.164**		0.172***
eigenv_top10 x S_20_49		0.090*		0.078*
eigenv_top10 x S_50_149		-0.034		-0.027
eigenv_top10 x S_150plus		0.094**		0.068*
S_0_9	-0.505***	-0.469***		
S_10_19	-0.332***	-0.332***		
S_20_49	-0.277***	-0.268***		
S_50_149	-0.191***	-0.148***		
Employment (log)			0.034	0.078
Employment squared (log)			0.008	0.004
Exclusion restriction:				
Positive R&D history	2.504***	2.503***	2.498***	2.497***
Chi-square for selection	606.71	607.00	651.75	646.33
Correlation coefficient rho	-.556***	-.557***	-.570***	-.567***
Log likelihood	-37383.08	-37362.27	-37289.27	-37248.29
Number of obs. (uncensored)	34398 (15415)	34398 (15415)	34398 (15415)	34398 (15415)

[1] All regressions include a constant, dummies for enterprise age, industry, and time dummies. Reference group: year 2005, Wholesale and retail trade industry (45-47 in SIC 2007), mature enterprises (15 years old or older). Models (3)–(4) differ by using employment (log) and employment squared (log) instead of the set of size dummies. All models are estimated by maximum likelihood using the Heckman procedure in Stata.
*** p<0.01, ** p<0.05, * p<0.1

The results in Table 5.2 support the presence of selection bias with a highly significant estimate of the correlation coefficient, i.e., ρ, around -0.56 in all model specifications. As expected, the R&D investment history variable has a high positive impact on the propensity to invest in R&D, indicating the extent of persistency in the enterprises' R&D policy. We also observe that coefficients for enterprise size indicators are increasing with enterprise size implying that larger enterprises invest more often in R&D than smaller enterprises. However, R&D intensity is decreasing with enterprise size implying that if small enterprises invest in R&D, their investment per employee is higher than in the large enterprises.

[9] On the one hand, enterprises that have previous R&D experience have a higher probability of engaging in R&D activities in the given period since both R&D and innovation have a high degree of persistency (jf. Peters, 2009). On the other hand, it is not obvious that having R&D experience implies higher R&D intensity in the given period (it can happen that "new" R&D investors, or enterprises that took a break from investing in R&D, invest more intensively in R&D in the given period than enterprises that invest continuously). The correlation between the *Positive R&D history* variable and the dummy for positive R&D in the given year is 0.80, while the correlation with the R&D intensity variable (r_{it}^*) is much lower and equal to 0.27.

The key variable of our interest in this analysis is an indicator of well-connectedness, i.e. eigenv_top10. It indicates whether the enterprise is among enterprises in the upper decile of distribution of eigenvalue centrality values. We choose eigenvalue centrality from the three earlier presented centrality measures because we believe that it reflects best the potential for faster knowledge spill-overs and information flows in the network. This variable is positively correlated with both the probability that the enterprise is engaged in R&D and the R&D intensity. The variable is more significant in the selection equation than in the equation for R&D intensity (cf. columns (1) and (3) in Table 5.2). However, when we consider the impact of being well-connected by enterprise size (cf. columns (2) and (4) in Table 5.2), we obtain positive and highly significant results for SMEs (10-49 employees). This result is robust for both model specification and implies that being well-connected through board network has higher impact on R&D activity of an SME.

6. Conclusions

In this report we attempt to seize the term relational capital and we examine whether (or to what extent) board networks may constitute a part of enterprises' intangible capital. To serve as an intangible asset, two criteria must be met: 1) A network (and eventually sub-networks) between board members via shared positions must actually exist, and 2) participation in such a network must have a value; to the single enterprise and thus to the society (directly, by increasing value added in the enterprises involved and indirectly through positive externalities).

In chapter 4 we use network analysis tools on Norwegian data and demonstrate that it indeed exists a board network with interconnected nodes. This network also contains clusters or sub-networks, particularly among SMEs.

We also examine how the network evolves from our starting point in 2004 to 2017 and finds that the network expands as the enterprise population increases. Of course, an expansion of the network increases the number of nodes and all possible connections between them. The number of observed connections does not increase to the same extent, leading to lower density in the network.

Compared to previous studies that rely on data for publicly listed companies, we can use register data for the entire population of Norwegian enterprises. This enables us to include SME's, which are believed to be of importance for innovation and economic growth. Very interesting, we find that this category of enterprises indeed is network-forming, with a lumpy structure.

We will point at several possible research topics related to the formation of board networks that are not addressed here. One is the determinants of individual network participation through board membership. We show that gender, as an example, affects the density of networks, women having higher network density. However, this could possibly be an effect of women's network being smaller. Other approaches focusing on individuals as network nodes could be to examine the background of board members compared to others. Have they been attending the same school or university? Possible effects of social class and gender? Same municipality?

Looking at the network with nodes defined by enterprises, we present in chapter 5 an analysis of how R&D spending is related to the enterprises' network connectedness, measured by centrality measures. On the proviso that we do not know much about the mechanisms, or the nature of causal relationships, we find that high network connectedness does have a positive impact on R&D spending. Most interesting, we find that this correspondence is particularly strong among SMEs.

To conclude; yes, the board network exists, and the strength of enterprises' network connection is associated with higher R&D spending, which is important for innovation and economic growth. Hence it should be counted for as an important part of enterprises' intangibles.

References

Akbas, F., F. Meschke and M. B. Wintoki (2016): Director networks and informed traders, *Journal of Accounting and Economics* 62, 1–23.

Amemiya, T. (1984): Tobit models: a survey, *Journal of Econometrics*, 24, 3–62.

Balsmeier, B., A. Buchwald and J. Stiebale (2014): Outside directors on the board and innovative firm performance, Research Policy 43(10), 1800–1815.

Belenzon, S., and T. Berkovitz (2010): Innovation in business groups, Management Science, 56, 519–535.

Borgatti, S. P. Everett, M. G. and Johnson, J. C. (2013): *Analyzing Social Networks,* Sage Publications Ltd.: London.

Chuluun, T., A. Prevost and A. Upadhyay (2017): Firm network structure and Innovation, *Journal of Corporate Finance*, 44, 193–214.

Crépon, B., E. Duguet, and J. Mairesse (1998): Research, innovation and productivity: an econometric analysis at the firm level, *Economics of Innovation and New Technology*, 7, 115–158.

Criscuolo, C., P.N. Gal and C. Menon (2014): The Dynamics of employment Growth: New Evidence from 18 Countries, OECD Science, Technology and Industry Policy Papers no. 14, OECD.

Everett, M. G. and Borgatti, S. P. (2013): "The dual projection approach for 2-mode networks", *Social Networks*, 35, 204–210.

Frutcherman, T. M. J. and Reingold, E. M. (1991): Graph Drawing by Force-directed Placement. *Software – Practice and Experience,* 21(11), 1129–1164.

Gomes-Casseres, B., A.B. Jaffe and J. Hagedoorn (2006); Do alliances promote knowledge flows? *Journal of Financial Economics*, 80(1), 5–33.

Gygax, A. and Hazledine, M. and M.J. Spencer (2017): Are Directors Really Irrelevant to Capital Structure Choice? Available at SSRN: https://ssrn.com/abstract=2876221

Helmers, C., M. Patnam and P.R. Rau (2017): Do board interlocks increase innovation? Evidence from a corporate governance reform in India, *Journal of banking and Finance* 80, 51–70.

Heckman, J. J. (1979): Sample selection bias as a specification error, *Econometrica*, 47, 153–161.

Hermalin, B.E., and M.S. Weisbach (1988): The Determinants of Board Composition, *The RAND Journal of Economics*, 19(4), 589–606.

Kamada, T. and Kawai, S. (1989): An Algorithm for Drawing General Undirected Graphs, *Information Processing Letters*, 31(1), 7–15.

Larcker, D.F., E.C. So, and C.C.Y Wang (2013): Boardroom Centrality and Firm Performance, *Journal of Accounting & Economics*, 55(2-3), 225–250.

Oh, W.-Y. and V.L. Barker (2018): Not All Ties Are Equal: Outside Directorship and Strategic Imitation in R&D, *Journal of Management* 44(4), 1312–1337.

Peters, B. (2009): Persistence of innovation: Stylised facts and panel data evidence, *Journal of Technology Transfer*, 34, 223–43.

Rybalka M. (2015): The Innovative Input Mix: Assessing the importance of R&D and ICT investments for firm performance in manufacturing and services, Discussion paper 801, Statistics Norway, Oslo.

Wilmet, A. (2017): Analysis of Board Networks in Belgium, *Master thesis.* University of Gent.

Wincent, J., S. Anokhin and D. Örtqvist (2010): Does network board capital matter? A study of innovative performance in strategic SME networks, *Journal of business research* 63, 265–275.

Zachary, W. W. (1977): “An Information Flow Model for Conflict and Fission in Small Groups”, *Journal of Anthropological Research* 33 (4), 452–473.

Schwartz-Ziv, M. and M. Weisbach (2013): What do boards really do? Evidence from minutes of board meetings, *Journal of Financial Economics* 108, 349–366.

List of figures

List of tables

ISBN 978-82-587-1024-7 (printed)
ISBN 978-82-587-1025-4 (electronic)
ISSN 0806-2056