THE EFFECT OF THE NETWORK POSITION OF A VENTURE CAPITALIST ON START-UP INNOVATIVENESS

– A QUANTITATIVE COMPARATIVE STUDY –

Illustration by Rhonald Blommestijn for Financieel Dagblad. Used with permission.

Author	*Sibert van Raaij – s1038206*
Course	*Master's Thesis Business Administration*
1st Supervisor	*Dr. Maurice de Rochemont*
2nd Examiner	*Professor dr. ir. Nanne Migchels*

Master's thesis – Innovation and Entrepreneurship

Author	*Sibert van Raaij*
S-number	*1038206*
1st Supervisor	*Dr. Maurice de Rochemont*
2nd Examiner	*Professor dr. ir. Nanne Migchels*
Date	*13 – 6 – 2021*
University	*Radboud University*
Master	*Business Administration*
Specialisation	*Innovation and Entrepreneurship*

Abstract

The purpose of this thesis is to empirically test how the network position of a venture capitalist (VC) influences the innovative performance of their portfolio companies in high-tech and low-tech industries. The obtained data was accessed through Crunchbase databases and consists of 157 pharmaceutical (high-tech) start-ups and 69 retail (low-tech) start-ups, separated into two samples by means of their technology intensity. Two linear regression analyses were executed and compared to test the main hypotheses. Findings show that start-ups in low-tech industries report higher innovative performance when collaborating with VCs with *strong ties* to other VCs. In addition, start-ups collaborating with VCs with higher number of *weak ties* report negative effects on their innovative performance. This implies that start-ups are more innovative when collaborating with VCs with strong ties. Significance of all mentioned effects have only been reported in the low-tech industry sample. Contribution of this thesis is to decrease ongoing entrepreneurial confusion by shedding light on the theoretical debate over network position effects. In addition, it provides entrepreneurs with an understanding of what implications selecting a VC has for their innovative performance.

Table of Contents

1. Introduction

1.1. The context of the research topic

Innovative start-ups have contributed significantly to global economic growth, job creation, and technological breakthroughs in the twenty-first century (Audretsch, Falck, Heblich, & Lederer, 2011; Baumol, 2004). Nevertheless, start-ups have higher failure rates compared to established firms, creating uncertainty for investors about the viability of start-ups (Audretsch et al., 2011). This causes start-ups to experience difficulty acquiring sufficient funding to grow into a solid and well-functioning venture (Winborg & Landström, 2001). This uncertainty also withholds formal banks from lending money to start-ups (De Clercq, Fried, Lehtonen, & Sapienza, 2006). Financial distress of start-ups can be satisfied through funding by a venture capitalist (VC), which is an investment type that funds start-ups with high growth potential, but also with high risks associated. The added value of venture capital funding is studied (Colombo & Grilli, 2010; De Clercq et al., 2006; Pratch, 2005) and it is suggested that VCs do not only invest in a monetary manner. VCs also bring managerial skills and their network to the table, a coaching or professionalization role (Bottazzi, Da Rin, & Hellmann, 2008). This thesis focuses in particular on the role of a VC's network on the further innovative performance of a VC-backed start-up.

Social Network Theory is taken as point of departure, which is a branch of literature seeking to explain and analysing the mechanics, interactions, and relations in network structures (Borgatti & Halgin, 2011). Within this field, two theories have been predominant: *The Strength of Weak Ties* (SWT) by Mark Granovetter (1973) and the *Structural Holes Theory* (SH) by Ronald Burt (1992). Due to their importance to this thesis, both theories will be clarified in the theoretical framework. In economic and managerial literature, Social Network Theory is used to explain the difference in firm performance. From this perspective, it is claimed that the firms' embeddedness in a network (its capacity to build and utilize social capital) affects the performance and innovativeness of the firm (Gulati, Nohria, & Zaheer, 2000). In other words, the superior network position of a firm matters and enhances its performance. Other research supports these claims and concludes that various techniques are present in social networks that improve start-up performance, such as access to valuable knowledge and resources (Burt, 1992). In addition to Burt (1992), it is suggested by Uzzi (1996) that diffusion of information and resources can be stimulated by a network.

The abovementioned relation also holds for networks of start-ups (Uzzi, 1997). Start-ups who are backed by VCs outperform start-ups who are not (e.g., Hochberg, Ljungqvist, & Lu, 2007; Rosenbusch, Brinckmann, & Muller, 2013; Sun, Zhao, & Sun, 2020). Other research (Engel, 2004) suggests that this stems from deviating characteristics of VCs such as size, experience, and most interesting to this thesis, the regional focus and position in its network. In other words, start-ups backed by VCs with an emphasis on their region and local network, perform better. The findings suggest that having a superior network position in the region tend to be relevant for growth of the portfolio firms of that VC (Engel, 2004). Argumentation for these findings is that information about investment opportunities often circulate in close proximity of the VC's geographical location (Sorenson & Stuart, 2001). Collectively, these studies indicate that the network position of a VC is a key explaining factor for innovation of the VC-backed start-up.

1.2. Relevance of the research topic

Many scholars focus on the relation between the network of a VC and their investing performance, i.e., the performance of their backed firms (Bellavitis, Filatotchev, & Souitaris, 2016; De Clercq et al., 2006; Hochberg et al., 2007). In addition, much literature is devoted to the exploration of a network structure, which is associated with strong and weak ties, structural holes and metrics such as betweenness and network-centrality (Audretsch et al., 2011; Borgatti & Halgin, 2011). The results emphasize the relevance of a central position in the network and its positive effect on the performance of the investments of VCs (Hochberg et al., 2007). This means that the way a VC is positioned in their network influences the innovativeness of VC-backed start-ups. Hence, it matters for a start-up's innovativeness which VC they choose to collaborate with.

Despite this relationship, contradictions are present about what network position is seen as a stimulant for innovativeness in different situations (Baum & Rowley, 2008; Bertrand-Cloodt, Hagedoorn, & Van Kranenburg, 2011; Granovetter, 1973). One explanation for these contradictions could be the effect of *technology intensity* of industries. Reasons for this presumption are that market mechanisms and innovation systems deviate across industries of different technology intensities (Hirsch-Kreinsen, 2008). In addition, the way knowledge is generated and shared is also assumed to differ across industries of different technology intensities (Hirsch-Kreinsen, 2008). Since networks contribute to the diffusion of knowledge (Xue, Dang, Shi, & Gu, 2019), it could be that they are affected by technology intensity as well. The Organisation for Economic Co-operation and Development, OECD, defines technology intensity of industries according to the ratio of R&D expenditures to the added value of the industry (OECD, 2016). This indicator is used to categorise industries from low-tech to high-tech. Due to the credibility of the OECD, many scholars use the classification (Zawislak, Fracasso, & Tello-Gamarra, 2018). Hecker (2005) uses the STEM-framework to classify industries by the concentration of knowledge- and technology-complex workers in an industry. It is argued that the concentration of Science, Technology, Engineering, and Mathematics workers in an industry say something about the complexity of the knowledge, processes, and technologies used (Hecker, 2005).

Scholars studied industry samples from different technology categories and came to different conclusions on whether strong ties, weak ties or a balanced mix of ties stimulates innovation. On the one hand, some researchers state that firms in high-tech industries engage in more long-term partnerships than firms in low-tech industries do (Baum & Rowley, 2008). On the other hand, various researchers contradict this view and state that higher numbers of weak ties and less similar partners, are positively related to the innovativeness of a firm (Bertrand-Cloodt et al., 2011). Most researchers examine high- and low-tech industries separately (e.g., Colombo & Grilli, 2010). The relevance of studying this relationship simultaneously in both high- and low-tech industries is that it enables comparison. It also makes it possible to explain why the relation differs in different technology categories. By testing the difference between high- and low-tech industries of a VC's network position, this thesis aims to contribute and help conclude the debate. Entrepreneurs could be provided with helpful insights by gaining an understanding of the beneficial influence of a VC's network. For example, on how venture capital could enhance the innovativeness of the start-up, and thereby improve the likelihood of survival.

1.3. Problem formulation

There is a lack of understanding of the relationship between a VC's network position and the innovativeness of the start-up. This touches upon the debate between perspectives on weak and strong ties (Baum & Rowley, 2008; Bertrand-Cloodt et al., 2011; Granovetter, 1973). From a Social Network Theory perspective, it is uncertain what the added value of a collaboration with a VC is based on this debate. For example, the general body of literature might support the assumption that VC networks have a positive impact on the innovativeness of start-ups (Abell & Nisar, 2007). However, this relationship might not hold for a particular industry, as can be suggested from the research of Wang, Zhao, Li, & Li (2015). These contradicting conclusions could result in confusion for entrepreneurs about if it matters which VC they choose to collaborate with. In addition, it can create confusion about what network position (strong versus weak ties) positively affects the innovative performance of start-ups. This might lead to negative finance decisions and threatens survival for entrepreneurs in these industries. This thesis tries to shed light onto this gap in the literature.

1.4. Research goal and research questions

The purpose of this thesis is to test whether the relationship between network position of a VC and the innovativeness of VC-backed start-ups is equal or different in high-tech and low-tech industries. Furthermore, a practical objective of this thesis is to give insights into what VC network position is suitable for start-up in both high- and low-tech industries. With this insight, entrepreneurs get more understanding of the ambiguous relationship and are enabled to choose VCs that fit with the start-up.

The aforementioned problem can be translated into the following research question: *How does the effect of the venture capitalist's network position on the innovativeness of VC-backed start-ups differ between high- and low-tech industries?*

The concepts mentioned in this question will be elaborated in chapter 2: Theoretical framework. To adequately answer the research question, the following sub questions have been formulated:

1. *What is the relationship between the venture capitalist's network position and the innovativeness of the VC-backed start-up in high technology intense industries?*
2. *What is the relationship between the venture capitalist's network position and the innovativeness of the VC-backed start-up in low technology intense industries?*

The remainder of this thesis is structured as follows. The next section will contain the theoretical framework, in which key concepts are explained and the relationship between them is elaborated a priori to the analysis. In the section thereafter, the methodology will be elaborated, including an explanation of the samples and variable measures as well as the data analysis method. The subsequent section will present the results per sub question and the corresponding statistical hypotheses. The final two sections will cover the discussion and conclusion of the findings and end with implications for theory and practice, future research suggestions and limitations of the research.

2. Theoretical framework

This section aims to explain the relevant theoretical concepts through previous literature and set the scope of this thesis. Thereafter, both the conceptual and statistical relationships between the concepts are justified and visually represented in a conceptual model at the end of this section.

In the previous chapter a problem statement, research goal and research questions were formulated. In order to answer the questions sufficiently, the concepts and definition will be explained and embedded in previous literature.

2.1. A start-up's innovativeness

The first concept is innovativeness within start-ups. Due to its large breadth, it is not surprising that the topic innovation has been researched regularly throughout the years. Topics are covered ranging from mechanisms of innovation (Garud, Tuertscher & Van de Ven, 2013), diffusion of innovation (Rogers, 1962) to innovation as industry destruction (Tripsas, 1997). The definition of innovation by Rogers (2003) is: *". . .an idea, practice or object that is perceived as new by an individual or group [or organisation]"* (p.12). However, this thesis is interested in the innovativeness – the perception of newness – of a start-up, which often does not yet possess products or services. Therefore, only the part of Rogers' definition about an idea is appropriate in this context. A more suitable definition of innovation comes from the research of Parida, Pesämaa, Wincent, & Westerberg (2016): *"a tendency to engage in and support new ideas, novelty, and experimentation that lead to developing new products, services, and technologies."* (p. 98). This definition will be used in this thesis because it focuses on both novelty and the consequences of innovativeness, such as development of new products, technologies and later on intellectual property. If an idea has novelty and sufficient distinctiveness, it can be protected by means of intellectual property rights. This sequentially can signal legitimacy and credibility to stakeholders (e.g., potential investors) (Miozzo & DiVito, 2016). The definition given above also has merit regarding the concept of absorptive capacity of a firm, which refers to *"the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends"* (Cohen & Levinthal, 1990, p. 1). Innovativeness of a firm is thus assumed to be dependent on the capacity to obtain and utilize knowledge from external factors, exposing the relevance of one's network position (Caseiro & Coelho, 2019). However, a lack of useful information and financial capital are not rare issues among start-ups and are main reasons for the high failure rate of start-ups (Mikle, 2020). Evidently, start-ups want to avoid failure, and one could ask how a start-up can develop a more profound network and stimulate innovativeness. One possibility is to acquire funding from VCs, as it has various benefits aside from the monetary funding. These will be elaborated in the next section.

2.2. Venture capital

The second concept is venture capital, which is a type of private equity investing in high potential start-ups in the seed, start-up and expansion stage (Florida & Kenney, 1988; Jeng & Wells, 2000). De Clercq et al. (2006) developed an extensive comparison between venture capital and other forms of private equity, such as business angels. The latter are wealthy private investors that often invest in start-ups based on personal affiliation or interests (De Clercq et al., 2006). Because they are private investors, business angels tend not to invest large amounts of money, often between €50.000 and €100.000 (De Clercq et al., 2006). In contrast, VCs invest capital from outside fund providers and often in joint operation with other VCs – called VC syndicates. Therefore, they have a notably bigger investment size, often between €2 million and €10 million (De Clercq et al., 2006; Van Osnabrugge, 2000).

Venture capital has various advantages over other forms of external private equity. Firstly, VC-backed start-ups have superior past performance over non-VC-backed start-ups (Florida & Kenney, 1988; Hochberg et al., 2007). VCs often take a formal seat in the firm as advisor or board member, making the relationship between the VC and the start-up more formal and frequent (De Clercq et al., 2006; Sapienza, 1992). Secondly, the formal character of a VC makes it subject to strict regulations and protocols from governmental institutions, stabilizing the quality and reliability of the VC (DeClercq et al., 2006). Thirdly, entrepreneurs can get access to the network and resources of the VC. The degree to which this third advantage is genuinely beneficial in high- and low-tech industries is the topic of this thesis and is assumed to be dependent on the network position of the VC. The work of Burt (1992), Gulati et al. (2000) and Uzzi (1996; 1997) all support this assumption. These studies achieve the same bottom-line conclusion: the position of a firm in the network affects one's performance. Hence, it is through the network that resources can be acquired, information can be shared and absorbed. As start-ups actively use the network of their VC (Bottazzi et al., 2008), the question arises what the effect of the VC's network position is on the innovative performance of that start-up.

2.3. The relationship between VCs and a start-up's innovativeness: A VC's network position

In order to understand the meaning of the concept of a VC's network position, it is necessary to take a closer look at Social Network Theory. As mentioned in the introduction, *The Strength of Weak Ties* (SWT) developed by Mark Granovetter (1973) and *Structural Holes Theory* (SH) by Ronald Burt (1992) are two predominant theories. To date, SWT and SH have been cited on Google Scholar over 58,630 times and 28,850 times respectively. Hence, the importance of these theories as foundation of new studies in for example economics, management and organizational design. As this thesis argues the existence of a relationship between two parties in a network, Social Network Theory is relevant to this thesis.

In his SWT, Granovetter (1973) starts by distinguishing two relationships in a social network, namely strong ties (e.g., family and friends) and weak ties (e.g., acquaintances) between nodes. Borgatti & Halgin (2011) reviewed SWT and came to two main premises. The first premise is that the social worlds of strong ties are likely to overlap, and thus contain not much new information for each other. The second premise is that there are ties connecting two nodes to each other through a third tie, the *bridging tie* (Granovetter, 1973, as cited in Borgatti & Halgin, 2011). Without bridging ties, it is argued that social systems become isolated (Kadushin, 2012). Combining the premises indicates that strong ties are unlikely to be the bridging tie (Granovetter, 1973, as cited in Borgatti & Halgin, 2011; Kadushin, 2012). Reason being that strong ties are mutually known and thus not providing new linkages to otherwise distant nodes. The bridging tie is more likely to be a weak tie since it provides new information that potentially can be the source of innovation (Borgatti & Halgin, 2011).

The second dominant theory in Social Network Theory is Burt's *Structural Holes Theory* (1992), in which the focus in not on connectedness between nodes but on the mere lack of it. A structural hole can be seen as a network position in which the central node is connected with others, who are only linked to each other through that central node (Burt, 1992; Burt, 2000). The theory has the premise that this situation can create (economical) advantage for the connecting node since its contacts are not connected without the connecting node. The value of structural holes can be determined by the number of *non-redundant ties*, i.e., ties that provide you with novel information which cannot come from other ties in the network. Burt (1992) states that having non-redundant ties is crucial for innovation. Having too little non-redundant ties can cause a central node to be isolated from information from more distant parts of a network and miss out on innovation opportunities.

When comparing both theories it becomes apparent that they are alike, however departing from a different perspective. Granovetter (1973) sees weak ties as crucial due to their bridging function, whereas Burt (1992; 2000) sees non-redundant ties as key factor due to the relative uniqueness of the information. As Borgatti & Halgin (2011) and Kilduff (2010) argue, whether the ties are called bridges or non-redundant, the concept and the outcome is the same, namely access to novel information.

2.3.1. The network of VCs

There is evidence from network theories (e.g., Burt, 1992; Gulati et al., 2000; Park & Bae, 2017; Uzzi, 1996; Uzzi, 1997) that support the relationship between network position and innovative performance. These findings also hold for start-ups that are backed by VCs (Abell & Nisar, 2007; Hochberg et al., 2007). The value of a VC's network is manifested in useful ties that otherwise could not have been utilized, for example ties with other VCs in investment syndicates (Ferrary & Granovetter, 2009). These could be used to acquire additional resources, human capital in the form of higher management members or social capital in the form of knowledge and information. Furthermore, to cope with information asymmetry, VCs utilize their network in order to validate information (Gompers, Gornall, Kaplan, & Strebulaev, 2020; Shane & Cable, 2002; Xue et al., 2019) on a start-up's past performance, legitimacy, experience or trustworthiness.

As elaborated above, the network position of a VC has a relationship with the innovative performance of VC-backed start-ups. Scholars provide grounds for argumentation that not only the VC's network position matters, but that it also could depend on the technological intensity of an industry. The effect of network position on VCs could be stronger in high technology-intense industries (Sun et al., 2020) and high knowledge-intense industries (Wang et al., 2015).

The effect of network position across high-tech and low-tech industries is the main focus of this thesis. Therefore, specific argumentation will be provided for the different effects of technology intensity of industries.

2.4. Network position in industry context: High- and Low-tech industries

Since this thesis makes the distinction between high- and low-tech industries, explanation of the network positions in both industries is relevant. As outlined in the introduction, technology intensity is defined as the ratio of R&D expenditures to added value of the industry and can be classified from high-tech to low-tech industries (OECD, 2016). In addition, this thesis looks at the concentration of STEM-workers, as it represents the focus on complex knowledge of an industry (Hecker, 2005). Due to the distinct nature of both industry types, different effects of network position are assumed to be present. While conducting literature review on this matter, contradictions have been noticed on which ties best stimulate acquisition of novel information (Baum & Rowley, 2008; Bertrand-Cloodt et al., 2011).

2.4.1. Network position in high-tech industries

On the one hand, firms in high-tech industries are thought to prefer strong ties instead of weak ties. It is argued that strong ties and strong long-term relationships based on mutual trust, provide crucial value for start-ups (Baum & Rowley, 2008). Baum & Rowley (2008) further argue that firms in high-tech industries tend to create more partnerships than firms in low-tech industries. Since high-tech industries are characterised by high levels of uncertainty (De Carolis, 2010), firms have to find stable partners in their network. Strong ties with other firms could give stability and give access to resources, knowledge, economies of scale or inter- and intra-industry network ties. In addition, DiMaggio & Powell (1983) argue that firms apply mimicking processes with the goal to copy successful ideas and products. It is in such way that firms try to avoid investing in unnecessary or high-risk projects. A second argument states that high-tech research and development projects have a longer timespan than in low-tech (Hirsch-Kreinsen, 2008). A longer research phase means that more money, effort and risk is involved. Firms involved develop specific knowledge of the project. This makes them of strategic value and can make it harder and riskier to switch, hence a possible favour for sticking with the strong tie instead of seeking elsewhere. Due to the high pace of technological development, firms in high-tech industries invest greatly in R&D capabilities (Hirsch-Kreinsen, 2008). This makes innovation-related information of key value for high-tech firms. Hence, firms are reluctant to share information with others but their trusted strong ties. This mechanism sheds light on the important role of having strong ties for organisations. However, Baum & Rowley (2008) do place some nuance regarding the findings. They state that it does not mean that the advantages of partnerships in these industries are a given fact, but that it is just a more occurring phenomena in high-tech industries than in low-tech industries.

On the other hand, researchers argue the importance of having weak ties in high-tech industries. For instance, Granovetter (1973) warns for overembeddedness. This happens when actors only have very strong ties, thus are likely to have overlapping social worlds, and do not look for opportunities beyond the strong tie partnerships. Overembeddedness could result in isolation, short-sightedness and inertia to new information. These findings are in line with Burt (1992), who states that higher numbers of weak ties provide a bridging function to distant networks. The distant networks, not accessed by anyone in the close circle, can give access to novel information beyond the own network (Bertrand-Cloodt et al., 2011). Inversely, it is also stated that embeddedness decreases turnover, but thereby stimulating firms to seek opportunities, be creative and find new partnerships (Hagedoorn & Frankort, 2008). However, it is likely that after years of strong collaboration between two nodes, the options for new partners or acquiring novel knowledge are limited (Baum & Rowley, 2008). Hirsch-Kreinsen (2008) also argue that high-tech firms have various information sources from in- and outside the industry, indicating the importance of weak ties to bridge the distance to outside the industry. This argument also has merit, as high-tech industries are subject to very rapid changes in processes, demand or technologies. Thereby, making it valuable to have many weak ties and acquire bits of novel information from various corners of a (distant) network.

2.4.2. Network position in low-tech industries

Within low-tech industries this debate is lively as well. The optimal tie strength in low-tech industries is equivocal. Compared to high-tech industries, the lower-tech branches are much more mature in terms of years of existence and are less subject to radical innovations (Hirsch-Kreinsen, 2008). Therefore, it is less uncertain that the status quo will be distorted. In other words, the uncertainty of the innovation paths that low-tech firms follow differ from high-tech firms since the nature of innovation differs (Hirsch-Kreinsen, 2008).

Some researchers state that low-tech firms need strong ties in order to optimally utilize innovation-related opportunities within their network. A first argument for this is that the innovation systems of low-tech and high-tech industries are often interdependent (Hirsch-Kreinsen, Jacobson, & Robertson, 2006; Pavitt, 1984; Reguera-Alvarado & Bravo, 2017). Therefore, low-tech firms need strong ties to cooperate with high-tech firms arguing a strong reciprocal relationship between low-tech and high-tech firms (Pavitt, 1984). The ambiguity of this relationship creates difficulty to ascribe the innovative performance to one actor since the origin of innovation is often hard to determine. For instance, innovation in a high-tech industry can come from other links in the value chain, such as the low-tech suppliers which initiate innovations. After initiating the innovation, it is furtherly developed by high-tech firms (Hansen & Winther, 2011; Pavitt, 1984). Having such strong ties with actors in high-tech industries can give access to complex knowledge needed to implement innovations in one's own low-tech industry. This is harder to be achieved with a VC with weak ties as high-tech firms can be reluctant to share knowledge with weak ties because of lack of mutual trust (Parmentola, Ferretti, & Panetti, 2020). A second important argument favouring strong ties is embedded in the nature of the low-tech industry. The stable nature of most low-tech industries enables firms to establish long-term partnerships and build long-term trust. These partnerships can evolve in joint operations and developing innovations together (Hirsch-Kreinsen, 2015).

Other researchers contradict these conclusions and state that innovativeness in low-tech industries should be ascribed to the weak ties in a network. As will be elaborated in the hypotheses section, low-tech industries rely more on practical knowledge obtained from "ongoing operational processes" (Hirsch-Kreinsen, 2008, p.27). Less complex knowledge requires less strong ties in order to understand and share since it is easier to codify (Hansen, 1999). This follows the same principle as tacit knowledge being harder to directly diffuse and converse without socialization and externalization processes to make knowledge explicit and codifiable (Nonaka & Takeuchi, 1995). It also aligns to the STEM-framework of Hecker (2005), where he states that less concentration of STEM-workers equals less complex and science-based knowledge. From this perspective, it means that start-ups in stable low-tech industries could potentially be better off with VCs that have many weak ties. Reason being that the novel information is less hard to understand from a distance in the network. The argument of Pavitt (1984) could also be used in the reversed sense. The core argument is that innovation is not solely ascribable to either high- or low-tech industries, but that it is interdependent. This could indicate that low-tech firms seek outside their industry to generate novel information. Since the knowledge in low-tech industries is less complex, it should be understandable without having strong ties and overlapping knowledge domains. In this sense, weak ties can function as bridging ties to gather less complex information in similar low-tech industries. This argumentation circles back to Granovetter's paper *The Strength of Weak Ties* (1973).

As can be noted, there is little consensus on whether strong or weak ties are the leading force for innovative performance of start-ups in low-tech industries. Both perspectives have compelling arguments and papers of both sides have considerable merit. Hence, it is often the conclusion that the innovative performance is stimulated by a balanced mix of both strong and weak ties (Bertrand-Cloodt et al., 2011).

2.5. Conceptual model and hypotheses

The focus of this thesis is on the difference between high-tech and low-tech industries. Therefore, the general relationship between a VCs network position and innovative performance of VC-backed start-ups will not be tested. In addition, this general relationship is already widely researched (Burt, 1992; Gulati et al., 2000; Park & Bae, 2017; Uzzi, 1996; Uzzi, 1997). However, it has been added as a positive relationship in the conceptual model as it is part of the context.

Up until now in this thesis, network position has been formulated as an ambiguous and normative dichotomy of good or bad. However, a network position can be calculated based on different mathematical scores, called *centrality*. As discussed in §2.3, ties can be reciprocal, meaning that the flow of for example information goes from A to B, but also back from B to A. Kadushin (2012) explains that ties can also be non-reciprocal. In other words, the number of ties you diffuse does not necessarily have to be equal to the number of ties that you assimilate. Generally, in the case that the number of direct ties a node has toward it exceed the number of direct ties it diffuses, a high *Degree Centrality* is noted. This is also referred to as "popularity" (Kadushin, 2012). Thus, a high degree centrality indicates a network position with many direct strong ties with the ego at the centre (Wang et al., 2015). This is the same as person A being popular because a lot of people know A, but A does not know all of them. While analysing a network, one can also look at the *Betweenness Centrality* of a node, referring to the degree to which information must pass others to get to a certain point (Freeman, 1978). Betweenness is not concerned with the shortest path, but rather with the most efficient path where traffic only has to pass nodes once (Borgatti, 2005). Nodes with high betweenness scores are located on the most proper path that connects clusters together (Zhang & Luo, 2017). This means that information has to go through them in order to reach the other cluster in the most efficient way. This level of control can give advantages since you are likely to acquire information very quickly and you are in control of to whom the information is diffused. They are connected to many, but far away, weak ties. Wang et al. (2015) state that in general, a high score on network centrality indicates a higher score on innovative performance of firms. This relationship is reported to be stronger in smaller firms than in bigger firms. Therefore, both degree centrality (presence of strong ties) and betweenness centrality (presence of weak ties) are assumed to positively relate with innovativeness of the VC-backed start-up, in line with research discussed in previous sections.

2.5.1. Network position in high-tech industries

The first sub question is concerned with the effect of a high technology intensity on the relationship between the independent variable "Network position" and the dependent variable "Innovativeness of start-up".

As outlined, the importance of one's network position is expected to differ across types of industries. Research of De Carolis (2010) states that high-tech industries have more rapid technological developments and require firms to respond faster to changes to keep technology at a state-of-the-art level. Schilling (2011) adds the notion that high-tech industries rely more on tacit, difficult to codify knowledge than low-tech industries. This implicit and complex knowledge requires more strong ties to be transferred (Zhang & Wang, 2013). In addition, the effect of strong ties is more significant in high-tech start-ups (Zhang & Wang, 2013). Thus, having multiple strongly tied R&D partners encourages information flow, creation and sharing of knowledge through trusted relationships (Bertrand-Cloodt et al., 2011). In addition, Delgado-Verde, Emilio Navas-López, Cruz-González, & Amores-Salvadó (2011) find a positive relationship between the implicitness of knowledge and the importance of a network in high-tech markets. Thus, pointing towards the importance of having strong ties in complex environments with high implicit knowledge being transferred, which is often a characteristic for high-tech industries. For a start-up in a high-tech industry, access to the network of a VC could impart the needed knowledge and resources to be innovative on the long term (Delgado-Verde et al., 2011; Parida et al., 2016; Schilling, 2011). In environments with long R&D-projects and high investments prior to generating profit, owning patents that offers protection to mimicking activities can be beneficial. Since information in these industries can be the difference between start-up survival or death, knowledge will only be shared with strongly trusted partners. Therefore, having strong ties (i.e., Degree Centrality) rather than having many weak ties (i.e., Betweenness Centrality) in a network is expected to lead to improved innovative performance. The abovementioned analyses constitute as base for H_{1a}:

Hypothesis 1a. *In high tech industries, strong ties in the network of a venture capitalist are positively related to the innovative performance of the VC-backed start-up.*

However, some researchers argue that implying weak ties in high-tech industries as inferior is oversimplified and incautious. As explained, solely focusing on strong ties can give rise to overembeddedness and create start-ups to only look for information in their strongly tied network (Baum & Rowley, 2008; Granovetter, 1973). Overembeddedness can therefore cause firms to not seek knowledge outside their network with like-minded actors. In addition, in the same article as used previously, Bertrand-Cloodt et al. (2011) also emphasizes that weak ties stimulate the innovative performance of high-tech firms. In favour of the weak tie perspective, Bertrand-Cloodt et al. (2011) conclude that "*the less organisationally intertwined companies and their partners are and the more companies have weaker ties to other companies that are not well-connected to similar partners, the higher their innovation performance*" (p.1026). This indicates that having a VC with relations that are not very similar to itself, can spark innovative performance. In essence supporting the *Strength of Weak Ties* paper from Granovetter (1973). In addition, as high-tech industries tend to move fast and are subject to a high innovation pace, having access to many different information sources in- and outside the industry can be crucial. Based on the previous analysis of various scholars, H_{1b} is formulated to test the weak tie-perspective in a high-tech industry sample.

Hypothesis 1b. *In high tech industries, weak ties in the network of a venture capitalist are positively related to the innovative performance of the VC-backed start-up.*

2.5.2. Network position in low-tech industries

The second sub question is concerned with the effect of a VC's network position on the innovative performance of VC-backed start-ups, in low technologically intense industries. Low-tech industries are characterised by a general lack of internal R&D capacity, the lack of a specific knowledge base and the unstructured innovation processes (Hirsch-Kreinsen, 2008). Low-tech industry firms are partly reliant on others in order to spark innovation. The debate however is on whether these firms need strong or weak ties to do so.

In low-tech industries, knowledge is less complex and therefore in order to be transferred and understood it doesn't rely dominantly on strong ties (Hansen, 1999). Some scholars favour the perspective of weak ties being more valuable in low-tech industries (Zhang & Wang, 2013).

As VC, having many but weaker contacts (i.e., a high betweenness score) enables VC-backed start-ups in low-tech industries to tap into multiple external knowledge networks and gain a wide variety of insights to stimulate innovativeness (Abbasiharofteh & Dyba, 2018). This is possible since knowledge is less dependent on knowledge intense workers and scientifical research, but rather on "ongoing operational processes" (Hirsch-Kreinsen, 2008, p. 27). In low-tech industries it is thought that having weak ties with other more distant networks or industries is important, as R&D innovation in the low-tech industry often does not come from within (Hirsch-Kreinsen, 2015). Hence, it is needed to gain insights from innovation-lively industries and implement it in one's own industry (Hirsch-Kreinsen, 2015; Pavitt, 1984). For example, the retail industry took insights from the software publishing industry on communicating and selling products through online means and implemented it in the retail industry. Hirsch-Kreinsen et al. (2006) state that SMEs in low-tech industries make up for the limited R&D expenditure and innovative power in other ways. Low-tech firms use practical knowledge, key capabilities and establishing contact with actors in distant fields or industries, as explained above. Hence, the assumed importance of weak ties in low-tech industries. Therefore, H_{2a} is formulated as follows:

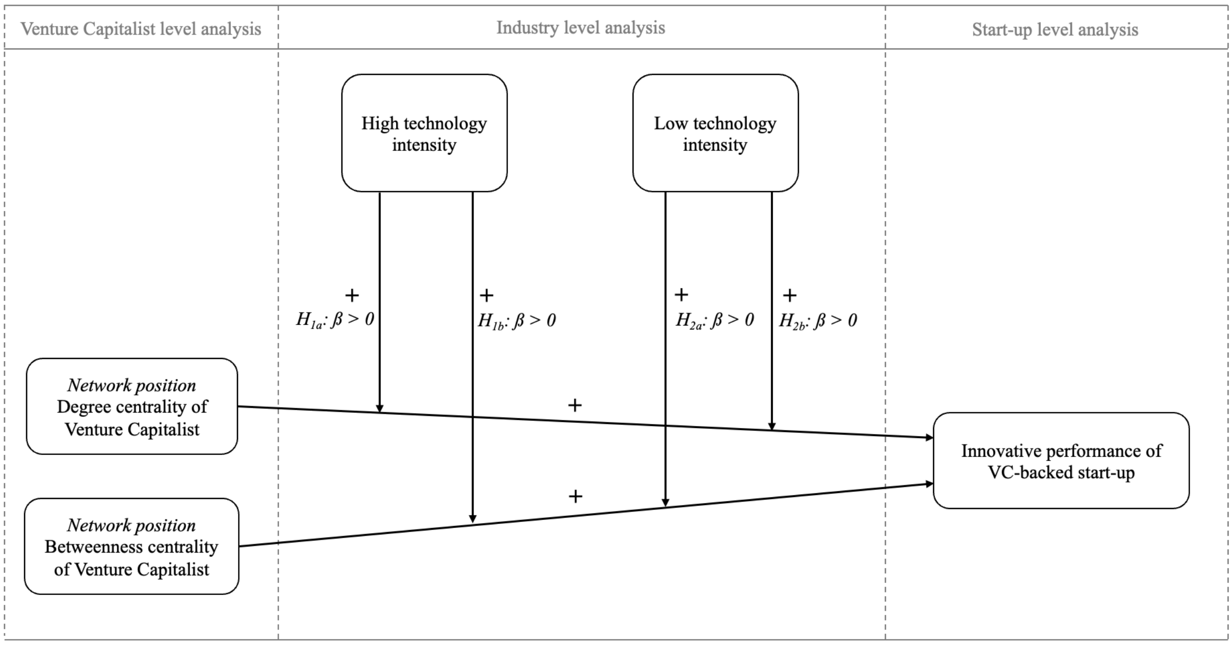
> Hypothesis 2a. *In low-tech industries, weak ties in the network of a venture capitalist are positively related to the innovative performance of the VC-backed start-up.*

Nonetheless, since the debate in low-tech industries is lively, strong ties are thought to be important as well. Pavitt (1984) argued the existence of an interdependence relationship of knowledge exchange between high- and low-tech actors. Hirsch-Kreinsen (2008) adds that having a tight connection to firms in R&D-intensive industries is crucial for the ability to innovate. This "tight coupling" (Hirsch-Kreinsen, 2008, p.34) between low- and high-tech firms can be interpreted as strong ties in the sense that they are reciprocal and long-term. The knowledge being shared could enclose models, prototypes or new product or process technologies (Hirsch-Kreinsen, 2015). According to Som (2012), low-tech industry firms can only be granted access to this knowledge through a close and reciprocal relationship. By means of H_{2b} the expected effect of strong ties on start-ups in low-tech industries is tested.

> Hypothesis 2b. *In low-tech industries, strong ties in the network of a venture capitalist are positively related to the innovative performance of the VC-backed start-up.*

These hypotheses can be visualized in the following conceptual model (Figure 1).

Figure 1. *Conceptual model and proposed hypotheses*



3. Methodological framework

The third chapter of this thesis is focused on elaborating the procedures which are followed to collect, prepare and analyse the data. Choices regarding type of research, variable measures, sample selection and analysis will be justified. Lastly, validity, reliability, generalizability and research ethics are discussed.

3.1. Type of research

In order to properly test the hypotheses and answer the (sub) questions, this thesis makes use of quantitative research methods. According to Babbie (2010), quantitative research methods focus on objective measurements and the statistical, mathematical, or numerical analysis of data collected. It is concerned with searching and analysing relationships between different variables. Two aspects are especially important in quantitative research: *numerical data* and *mathematically based methods* (Sukamolson, 2007). For mathematical analysis, the studied data is required to be of numerical nature in order to be properly analysed. Whereas qualitative research methods often do not use numerical data and therefore can't be fully statistically substantiated (Sukamolson, 2007). Quantitative research has the advantage that it enables the researcher to efficiently study larger groups or larger datasets and make generalizations about a group beyond the studied sample (Swanson & Holton, 2005).

Since this thesis has the goal to compare measures of high-tech industries and low-tech industries, a comparative research will be conducted. This type of research has the goal to identify a relationship between two or more groups. As can be seen from the theoretical framework, scholars already invested great effort in this topic and therefore exploratory research is not the best suited approach for this thesis. Instead, an explanatory approach will be taken as this thesis tries to *explain* why the innovative performance of start-ups might differ across VC networks in high- and low-tech industries.

3.2. Data source, measures and samples

3.2.1. Data source

The dataset that is used for this thesis consists of samples from a larger dataset accessed through a database from Crunchbase. Crunchbase is a data source, created in 2007, mostly focused on funding of innovative start-ups and performance (Dalle, Den Besten, & Menon, 2017).

Crunchbase is suited for this thesis since the majority of data in the database concerns venture capital funding from a wide variety of industries (Dalle et al., 2017, p.14). It is important to note that US based start-ups are overrepresented in the database, 34.8%, while the second country is the UK with only 6.2% (Ferrati & Muffatto, 2020). This could harm the generalizability of the results as the country of origin is not equally distributed. This should be taken in mind when interpreting results based on the entire database. The validity of the data included in Crunchbase, however, is checked daily by data scientists, AI-technology and the investors of whom data is included (Ferrati & Muffatto, 2020). The Crunchbase database is structured in a way that all scores on the variables are available per start-up, which enables comparative analysis between different industries of start-ups, i.e., the samples.

The centrality measures are calculated within programming software Python and NetworkX. NetworkX is a Python software module focused on network analysis. It rearranges datasets in such a way that enables the user to create, manipulate and study the structure of networks (NetworkX, 2020). In the following section these formulas from NetworkX are elaborated.

3.2.2. Variable measures

3.2.2.1. Dependent Variable

The dependent variable within this thesis is Innovative Performance of Start-ups that are backed by VCs. Different measures for innovativeness exist in research, distinguished by the input- (e.g., R&D budget) or output-orientation of the measure (innovations, patents, trademarks). Previous research showed the success and accuracy of using output-oriented measures such as the number of patents as proxy for innovativeness (Crosby, 2000; Owen-Smith & Powell, 2004; Rothaermel & Hess, 2007). Therefore, the primary proxies for determining the innovativeness in this thesis are the *Number of patents* and *Number of trademarks* a start-up owns. Patents enable start-ups to protect their innovations, ideas, and technologies in the form of intellectual property. Trademarks are an adequate way for firms to protect their brands, products, or strategies and to indicate innovativeness (Mendonça, Pereira, & Godinho, 2004). In addition, measuring patents to determine innovativeness is a highly accurate and up-to-date proxy since the average time lag between the invention and application for a patent is only 2-3 months (Rothaermel & Hess, 2007). Besides the accuracy of this proxy, patents are also an important legitimacy signal to external investors of technological novelty and expertise (Rothaermel & Hess, 2007). This enhances the possibility for future investment in the start-up. The proxy is measured using the Crunchbase data add-on *IPQwery* which measures the patents officially granted and trademarks officially registered to the start-up (Crunchbase, 2018, January 19). Within the Crunchbase database, *Number of patents* and *Number of trademarks* are of metric measurement level as they met the criteria of (1) having an indisputable order, (2) fixed units of measurements, and (3) the value “0” means an absence of the unit of measurement.

Since a regression model only accepts a single dependent variable, the two measures are combined into a new dependent variable called Innovative Performance of Start-ups. This variable is the sum of the scores on Patents granted and Trademarks registered. The equation is as follows:

Innovative Performance of Start-ups = Trademarks Registered + Patents Granted

3.2.2.2. Independent Variable

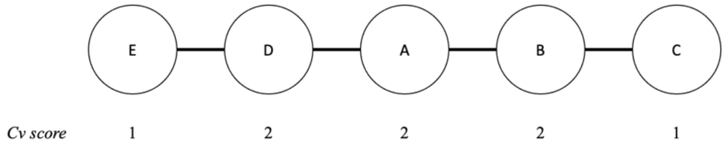
The independent variable is Network Position. As explained in the theoretical framework, a suitable proxy for network position of the VC is the centrality measure. Centrality has many variations, all concerned with other aspects of one’s network position. Mainly used are Degree Centrality, Betweenness Centrality, Closeness Centrality, and Eigenvector Centrality.

In the dataset, the Degree and Betweenness centrality measures of the first five investors are shown. In this thesis, the average score of these investors is used to determine the network position of the VC or VCs. So, if only three investors are present, then the average score is calculated based on three and not on five scores.

As theoretically substantiated in §2.5., *Degree Centrality* can function as proxy for VCs having strong ties in their network, also shown by research of Valente, Coronges, Lakon, & Costenbader (2008). Research on relations among centrality measures report high positive correlation between Degree and Eigenvector centrality, making combining them redundant (Batool & Niazi, 2014). Due to the widespread use, this thesis will use Degree centrality as proxy for strong ties in the VC’s network. Within NetworkX, Degree centrality (Cv) is calculated as the number of direct ties a node has (Figure 2), meaning all the nodes that are directly connected to the ego (NetworkX Developers, 2020). For example, with five nodes in a row, the outer two nodes only have 1 direct tie, whereas the middle three nodes all have 2 direct ties.

Figure 2. *Formula and simplified example of Degree Centrality* (derived from NetworkX Developers, 2020; Zhang & Luo, 2017)

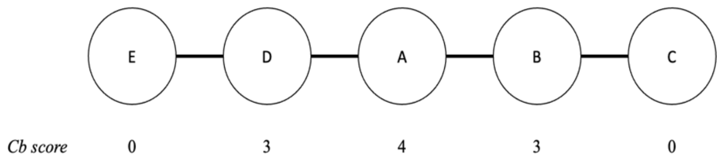
$$Cv = \frac{\mathrm{d}v}{(N-1)}$$



Similarly, in §2.5. Betweenness centrality has been theoretically linked to the presence of weak ties in a VC's network. Therefore, this thesis uses *Betweenness centrality* as proxy for weak ties. According to the same research of Valente et al. (2008), betweenness centrality did not noticeably correlate with any of the other centrality measures. As such, generating insights which cannot be gained from another measure (Batool & Niazi, 2014). Because betweenness centrality refers to the degree to which information has to pass other nodes, it also measures the path through which the most information flows. Having a high betweenness centrality (Cb) score means that you are on a path of two or more nodes trying to reach each other. The same example is depicted in figure 3 below, but now with betweenness scores. Node D is on three paths: E to A, E to B and E to C. Evidently, node E plays no part in passing information through, and thus has a Cb score of 0.

Figure 3. *Formula and simplified example of Betweenness Centrality* (derived from NetworkX Developers, 2020; Zhang & Luo, 2017)

$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$



Both centrality measures are of metric measurement level, as all criteria are met: (1) presence of a ranked order; (2) there is a fixed distance between each measurement unit; and (3) there is a natural point of zero where zero means "nothing". Both centrality measures are calculated by a data scientist for the entire database from Crunchbase. The centrality scores of a single VC are calculated in relation to the entire network of all VCs in the Crunchbase database.

3.2.2.3. Moderating Variable: Technology intensity

Classification of whether industries are classified as high- or low-tech can be proxied by the measure of *R&D intensity* (OECD, 2016). This measure is often used for ranking economic activities in various industries in either low-tech, medium-low-tech, medium-tech, medium-high-tech and high-tech industries. Almost all classifications only include manufacturing industries as high-tech service industries lack concrete data (Goldschlag & Miranda, 2019). However, the OECD classification used in this thesis, accounted for this and was able to include service industries as well. The intensity of an industry is measured as R&D expenditures of an industry as % of the Gross Value Added (GVA). In order to account for country specific differences such as purchasing power, the OECD (2016) uses *purchasing power parity* to properly compare two or more countries despite differences. The formula is therefore the following (Figure 4), where R&D intensity is measured in industry *i* and GVA is measured in country *c* (OECD, 2016).

Figure 4. *Formula of R&D intensity.* Derived from OECD (2016)

$$\left(\frac{R\&D}{GVA}\right)_i = \frac{\sum_c R\&D_{ci}}{\sum_c GVA_{ci}} = \sum_c \frac{R\&D_{ci}}{GVA_{ci}} \frac{GVA_{ci}}{\sum_c GVA_{ci}}$$

Furthermore, company activities are clustered according to the ISIC-hierarchy at 2 (in some cases 3) digit-level. This means that for example, from the ISIC code 32 (Other Manufacturing) only ISIC code 325 (Manufacture of Medical and Dental Instruments) is included in the medium-high R&D intensity cluster (OECD, 2016). This enables an industry analysis at a deeper level, enhancing the accuracy of the industries included.

However, only focusing on R&D intensity when classifying industries can lead to very rigid classifications. Therefore, this thesis also uses the STEM-framework of Hecker (2005) to determine the technology intensity of industries.

The moderating variable is dichotomous and of categorical measurement level as it will be transformed to a dummy variable. A score of “0” will mean low-tech and a score of “1” will mean high-tech. The data for the R&D intensity scores is retrieved from the *OECD Taxonomy of Economic Activities Based on R&D Intensity* from 2016 and the STEM framework of Hecker (2005), summarized in Appendix A.

3.2.2.4. Control variables

It is expected that the discussed variables do not explain the entire relationship. In order to understand the tested relationship, it is important to rule out other possible effects. The first control variable is *Firm size*. In order to control for this variable, the number of employees of the start-up is used to measure firm size, as is common in research alike (Delgado-Verde et al., 2011; Engel, 2004; Rothaermel & Hess, 2007). Controlling for firm size is important because bigger firms often have bigger pools of resources which can be utilized to reinforce innovation.

The data for this control variable is available in the Crunchbase database and is divided into seven groups ranging from 0 to 10001+ employees. In this thesis, these groups are divided into three groups labelled as *Low*, *Medium* and *High number of Employees* dummy variables. The group boundaries are respectively start-ups with ≤ 175 employees; between 176 and 750 employees; and lastly more than 751 employees (Table 1). By dividing the control variable in three categories, the effect on Innovative Performance of Start-ups can also be distinguished between small, medium and big start-ups.

Table 1. *Group sizes of dummy categories for control variable Size of start-up*

	Low-tech		High-tech	
	N	Percentage	*N*	Percentage
Low number of employees (≤175)	174	71%	140	53%
Medium number of employees (176 - 750)	49	20%	62	23%
High number of employees (≥751)	21	9%	63	24%
Total	244	100%	265	100%

The second control variable is *Firm age*, measured in year of foundation of the start-ups. Also often used as control variable in entrepreneurial innovation research (Delgado-Verde et al., 2011; Engel, 2004; Rothaermel & Hess, 2007). Controlling for firm age is important as older start-ups had more time to do research, establish a network, or gain legitimacy to signal potential investors. The used Crunchbase database provides access to this data.

The third control variable is *Number of investors*, measured in the total number of investors and lead investors a particular start-up has. Start-ups with more investors are expected to bring more knowledge. However not used very often, many researchers use equivalents of this control variable such as number of portfolio companies (Engel, 2004) or number of founders (Colombo & Grilli, 2010). In research on venture capital syndication, the number of investors is often an important variable, such as in Lerner (1994) or Hochberg et al. (2007). The distinction between number of investors and number of lead investors has also been made. Reason being that it is interesting to analyse if the effect on innovative performance differs between having a normal investor or having dominant investor(s). Table 2 below shows a summary of all concepts, their accompanying variables, its measures and previous research using the measure.

Table 2. *Research Variable Operationalization*

Concept	Variable	Measure	Previous research using measure
Dependent Variable			
Innovative Performance of Start-up	Innovativeness	Number of Granted Patents	Crosby, 2000; Mendonça et al., 2004;
		Number of Owned Trademarks	Owen-Smith & Powell, 2004; Rothaermel & Hess, 2007
Independent Variable			
Network Position Venture Capitalist	Network Centrality	Degree Centrality	Abell & Nisar, 2007; Hochberg et al., 2007; Valente et al., 2008
		Betweenness Centrality	Abell & Nisar, 2007; Batool & Niazi, 2014; Hagedoorn & Frankort, 2008; Hochberg et al., 2007
Moderating Variable			
Technology Intensity	R&D Intensity	R&D Expenditure to Gross Added Value - Ratio	Goldschlag & Miranda, 2019 OECD, 2016;
Control Variables			
	Firm Size	Number of employees	Delgado-Verde et al., 2011; Devigne et al., 2011; Engel, 2004;
	Firm Age	Years since foundation	Rothaermel & Hess, 2007
	Number of venture capitalists	Number of investors	Colombo & Grilli, 2011; Engel, 2004; Hochberg et al., 2007
		Number of Lead investors	Lerner, 1994

3.2.3. Research samples

In order to select adequate industries, this thesis used the *method of differences* for case selection. According to Esser & Vliegenthart (2017), this method focuses on comparing cases or samples that are similar, except on aspects that are studied. In this thesis the differentiating factor is the technology intensity. This method is appropriate in comparative research as it allows the researcher to select samples that have the highest opportunity of finding a relationship (Esser & Vliegenhart, 2017). To select observations within the two samples, this thesis used *purposive sampling* which is part of the non-probability sampling techniques (Taherdoost, 2016). This means that the researcher selects cases based on set criterion. In this thesis, all criteria are theoretically substantiated. The drawback of using purposive sampling is that it is prone to researcher bias as the samples are selected based on the researcher's judgement. In order to account for this bias, sample selection is done through objectified and previous theoretically used criteria (Table 3). Often, scholars use one or two criteria for classifying the technology intensity of industries, such as Zawislak et al. (2018). Firstly, as elaborated in the section before, R&D-expenditures is one method used for selecting cases. Industries with ratios over 24,0 are considered high-tech and industries with ratios under 0,38 are considered low-tech (Appendix A).

Secondly, the concentration of STEM-employment in an industry can also be used to label industries. The STEM employment methodology is developed by Hecker (2005) and identifies the high-tech industries based on the concentration of Science, Technology, Engineering, and Mathematics (STEM) workers in an industry. The group with the highest concentration of STEM-workers has at least five times the average STEM-workers in employment (Goldschlag & Miranda, 2019). The industry selected for this thesis should fall within the top group, whereas it is required that the low-tech group should not be included in the STEM-classification, as only high-tech industries are included in the framework. Table 3 and Appendix A show that the samples met the requirements, and which previous literature also used or examined these criteria. No case from the middle three classes has been selected since the boundaries of the classes are unambiguous and context dependent (Goldschlag & Miranda, 2019).

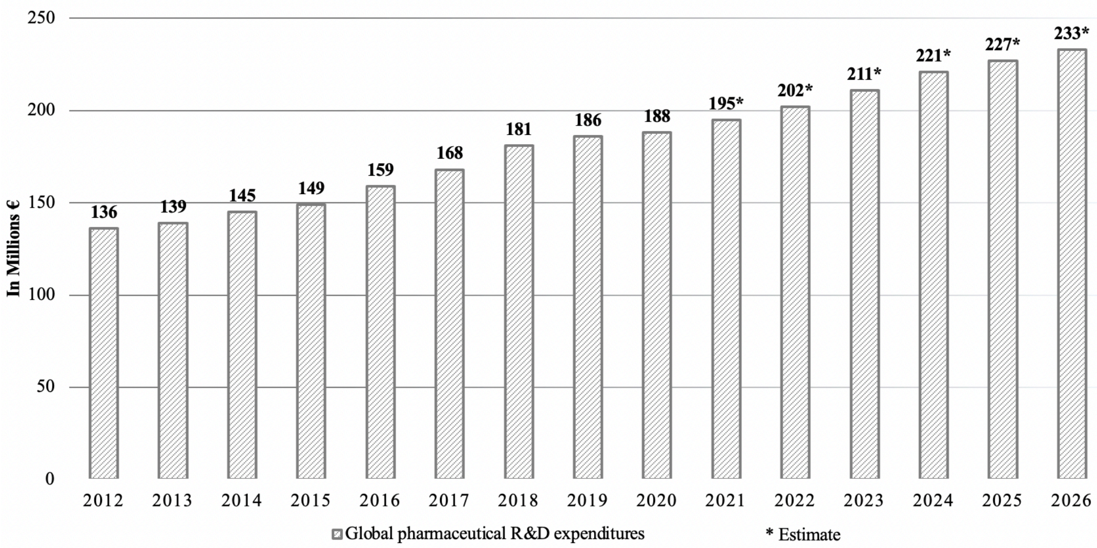
Table 3. *Framework for selecting industries representing the research samples*

Selection criteria	Detailed criteria for the research sample	Previous research	Outcome
1 R&D-expenditure of the industry	Low-tech sample: ratio of expenditures on Research and Development to Gross Added Value within the industry should not exceed 0,38%. See Appendix A. High-tech sample: ratio of expenditures on Research and Development to Gross Added Value within the industry should not fall below 24,0%. See Appendix A.	Goldschlag & Miranda, 2019; Hirsch-Kreinsen, 2008; OECD, 2016; Zawislak, Fracasso, & Tello-Gamarra, 2018	Retail industry scores 0,28%, thus meets criteria. Pharmaceutical industry scores 27,98%, thus meets criteria.
2 Concentration of STEM-workers	Low-tech sample: should not be included in STEM level I, II or III from Hecker (2005). See Appendix A. High-tech sample: should be included in level I in the STEM methodology. See Appendix A.	Goldschlag & Miranda, 2019; Haltiwanger, Hathaway, and Miranda, 2014; Hecker, 2005; Wolf & Terrell, 2016	Retail sample is not included in level I, II or III of the STEM framework. Pharmaceutical sample is included in level I of the STEM framework.

3.2.3.1. The high-tech industry sample in context: Pharmaceutical industry

VC-backed start-ups within the pharmaceutical industry (ISIC code 2100) have been selected as sample for the high-tech industry for a number of reasons. First, the pharmaceutical industry is classified as the near top of the high-tech industries within the OECD research (2016) and the STEM methodology (Hecker, 2005). It consists of a variety of organizations ranging from the manufacturing of biotech pharmaceuticals, antibiotics, vitamins, to vaccines and medical impregnated bandages (SICCODE, 2019). In 2019, the global revenue of this industry was estimated at $1,250,100,000,000 (or 1.25 trillion USD) (Statista, 2020a). The OECD calculated an R&D intensity of 27,98%, second highest of all analysed manufacturing activities. Second, according to Hecker (2005), the pharmaceutical industry is labelled as a level I high-tech industry, in line with the R&D-intensity classification. Third, patenting is proven to be of crucial importance for business success in the pharmaceutical industry (Rothaermel & Hess, 2007). Shown in figure 5, pharmaceutical firms invest money in development of new technologies at a large, and in the US unmatched, scale (Statista, 2020a; Statista, 2021).

Figure 5. *Global pharmaceutical R&D Expenditure 2012-2026*



Data derived from Statista (2020a).

The R&D expenditures are made far in advance of profit generation, creating high business risk and uncertainty (Grabowski, 2002). Therefore, the innovation can be patented for a set period of time in which the R&D costs have to be earned back. Absence of patents can seriously harm the viability of firm because of the free-rider advantage of competitors. The costs of imitating are miniscule compared to the development costs (Grabowski, 2002). Hence, the importance of patents in the pharmaceutical industry.

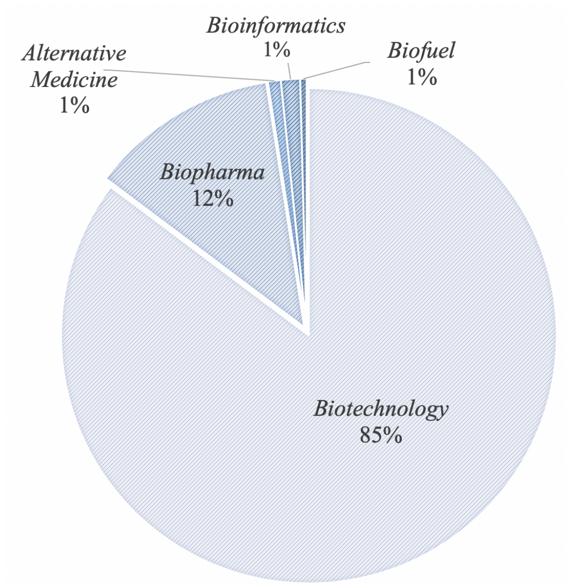
The research sample consist of 244 start-ups, with headquarters located in the United States, China, Switzerland, United Kingdom, Germany, and others with a share < 3% (Table 4). All start-ups are active in the pharmaceutical industry. Figure 6 shows the sub industries where the start-ups are active in. The vast majority, 85% is active in Biotechnology, whereas 12% is active in Biopharma. The former produces medicines based on biological organisms and the latter produces medicines based on chemical compositions. The other 3% is ascribable to biofuel, bioinformatics and alternative medicine manufacturers. The variety of sub-industries is relatively low, which limits the generalizability of the results.

In order to further ensure validity of the sample, various keywords such as “retail”, “artificial intelligence”, “consumer health”, “developers API”, or “cannabis” were checked for and if necessary, excluded from the sample group. This criterium was implemented so that there was no overlap between the sample to ensure validity. The average start-up size, measured in number of employees, is quite high for start-ups. One reason could be that due to substantial (R&D) investment in these start-ups, they are able to hire employees and grow at rapid pace. In addition, many pharmaceutical start-ups are either doing or are subject to acquisitions. This way merging the pools of employees together. A second reason could be that large incumbents start smaller business units that focus on a specific opportunity in the market, getting resources from the parent company (Chen, 2017).

Table 4. *Configuration high-tech industry research sample*

Country	Frequency	Percentage of Total
United States of America	159	65,16%
China	17	6,97%
Switzerland	10	4,10%
United Kingdom	10	4,10%
Germany	9	3,69%
Other	39	15,98%
Total	244	100,00%

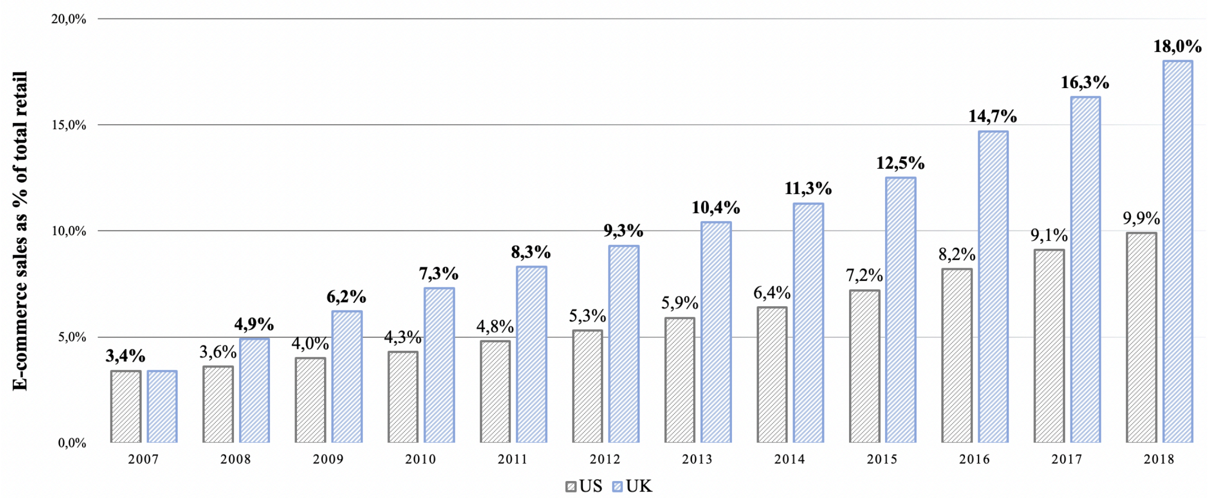
Figure 6. *Industries included in high-tech research sample*



3.2.3.2. The Low-tech industry in context: Retail industry

In order to study the role of low-tech start-ups, the retail industry (ISIC code 47) is selected. This industry consists of firms (re-)selling used or new products mainly targeted at the general public for personal or household consumption (United Nations, 2008). It embodies brick and mortar stores such as clothing stores, furniture stores, electronics stores, and supermarkets. With the rise of e-commerce, the era of digitalisation began and impacted the retail industry vastly, both by means of growth and number of stores, see figure 7 (Tolstoy, Nordman, Hånell, & Özbek, 2021). This enabled new types of organisations focusing more on self-service, convenience and platformisation (Tolstoy et al., 2021). Famous examples being Uber, AirBnB, and Just Eat Takeaway. Besides platform organisations, previously brick-and-mortar stores also transformed to selling in physical stores and online web-shops.

Figure 7. *E-commerce retail sales relative to total retail sales 2007-2018*



Data derived from Statista (2020b) and Office for National Statistics (2021).

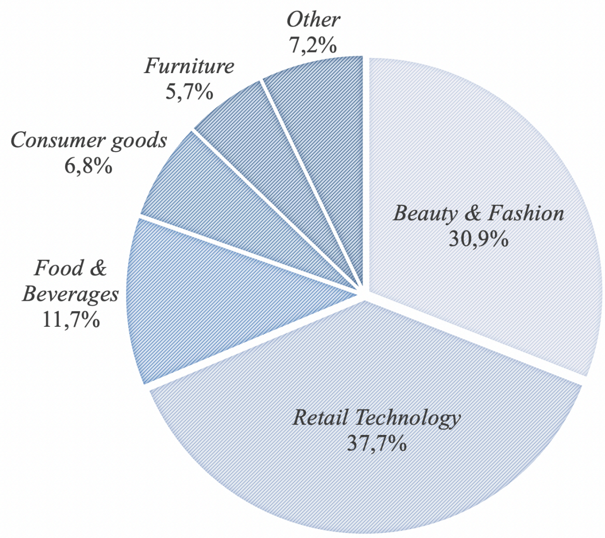
The retail industry is suitable as sample for low-tech industry for various reasons. Firstly, the retail industry is classified as a low-tech industry, with an OECD taxonomy score of 0,28% (OECD, 2016, see Appendix A). Secondly, the retail industry is not included in level I, II or III of the STEM-classifications, which excludes the industry from being a high-tech industry based on the concentration of STEM-workers. Thirdly, the retail industry is frequently subject to venture capital investment (The Economist, 2017). Due to the rise of e- and m-commerce, retailers shifted from being an intermediary to a model called "Direct-to-Customer", which called the attention of VCs (The Economist, 2017). Trademarks are especially suitable for indicating innovativeness in the low-tech sample (Mendonça et al., 2004). Product designs, brands, commercials or product differentiation can be subject to trademarks.

The sample consist of 265 start-ups active in the retail industry. The Crunchbase dataset used a very broad definition of "retail", including many start-ups not included in the ISIC 47 definition. In order to ensure that the sample solely consist of ISIC 47-included start-ups, the data has been cleaned thoroughly. Start-ups with keywords such as "Supply Chain Management", "Robotics", "Manufacturing", "Food processing", "Agri", "Delivery" and "Wholesale" were excluded as they represented activities not included in ISIC code 47. Table 5 shows the country distribution of the selected sample, reporting similar overrepresentation of US-based start-ups to the high-tech sample. Figure 8 shows the distribution of industries that are included in the low-tech research sample. The two biggest groups within this sample are Retail Technology (37.7%) and Beauty & Fashion (30.9%). Other sub-industries in which the start-ups are active are Food & Beverages (11.7%), Consumer Goods (6.8%), Furniture (5.7%) and other sub-industries (7.2%). It immediately becomes clear that this sample is more diverse that the high-tech sample. This should be taken into mind when interpreting the results, as they can be generalized across a higher variety of industries than the high-tech sample.

Table 5. *Configuration low-tech industry research sample*

Country	Frequency	Percentage of Total
United States of America	138	52,08%
United Kingdom	38	14,34%
India	13	4,91%
Canada	12	4,53%
France	11	4,15%
Germany	11	4,15%
China	10	3,77%
Other	32	12,08%
Total	265	100,00%

Figure 8. *Industries included in low-tech research sample*



3.3. Data analysis strategy

The intended data analysis procedure is regression analysis. According to Hair, Black, Babin, & Anderson (2018), as the dependence relationship between multiple metrically scaled variables is tested, a preferred data analysis is regression analysis. More specifically, a multiple regression analysis is conducted since there are more than one independent variables. In addition, this relationship is modified by a third categorical variable, technology intensity of the industry. Since two samples are drawn from the dataset, two separate multiple regression analyses will be conducted and after analysis compared to each other.

In the field of quantitative data analysis, various dependence techniques can be used such as regression analysis and (multivariate) variance analysis (Hair et al., 2018). Regression analysis has various advantages over variance analysis for this thesis. Firstly, it is an analysis method focused on dependence relationships of metric variables and is the most powerful and flexible statistical test (Allen, 2004). Secondly, it allows for accurate explanation of the different effects of the independent variables on the dependent variable (Allen, 2004). Thirdly, regression analysis provides information about the relevance of each variable on the dependent variable. This way it can become clear what the role of network position is and if it is stronger or weaker in high- and low-tech industries.

3.4. Validity and reliability

A valid and reliable study enhance the value of the results. In addition, it is important for future research that previous studies are conducted as valid and reliable as possible. According to Hair et al. (2018), validity is defined as the "*Extent to which a measure or set of measures correctly represents the concept of study*" (p. 3). It is concerned with how well the selected measures actually measure the concept it should measure. There are various validation processes to assess validity of the research. First, validity of the research can be established prior to regression analysis via several assumptions that indicate how well the variables fit the data. One of the assumptions that can give an indication of validity is the constant distribution of residuals. This assumption, among others, will be addressed in chapter 4 hereafter.

Secondly, goodness of fit of the regression models indicates how well the models fit the observed data (Field, 2017). The *adjusted* R^2 and *adjusted* R^2 *– change* measures indicate how well the model fits the data and therefore can be used to determine the goodness of fit. These measures indicate how much of the model's total variance is actually explained by the included variables (Field, 2017). A significant improvement of adjusted R^2 tells the researcher that the newly included variable helps to increase the explanatory power. In addition to the adjusted R^2, F-tests also conclude how well the model fits the data and if the model is useful for interpretation. Chapter four will address these measures that will conclude usefulness and validity of the regression models.

Thirdly, in order to ensure that the data samples have high validity, it is required to adhere to various criteria. First, each variable used as proxy for the concepts has been successfully used by a variety of peer-reviewed studies. Thus, the scales used in this thesis are validated by previous research (see previous table 2) and indicate sufficient construct validity. Second, elimination of start-ups that are not included in the ISIC-code also increased validity since deviations from what is desired to measure, are excluded. However, purposively selecting observations for sampling (known as purposive sampling) harms the sample's objectivity and generalisability (Taherdoost, 2016, p.23). In order to minimize sampling bias, this thesis used objective methodological approaches to select samples such as the OECD-taxonomy, STEM-framework and the method of difference. Despite purposive sampling, the used dataset and the samples in this thesis are overrepresented by US-based start-ups. This harms the generalizability of the findings as they cannot be directly extrapolated to other countries in the dataset. However, as this thesis does not focus on differences between countries, the impact is expected to be marginal.

Reliability of this thesis is partly based on the quality of the provided data from the Crunchbase dataset. The use of secondary data as primary source is beneficial as the validity and reliability are already pre-established. In this thesis through previous scholars, the data-collection agency, independent validators and an internal data science team of Crunchbase (Ferrati & Muffatto, 2020; Olabode, Bakare, & Olateju, 2018). According to research of Dalle et al. (2017), the datasets of Crunchbase have sufficient validity and reliability and is being used increasingly in publicised studies. In addition, reliability is also concerned with the degree to which data is distributed normally. Non-normally distributed data will behave differently every time the data is used, and therefore it can harm the reliability of the results. Using statistics software SPSS 27, normality of the IVs and DV has been tested. Since multiple variables were skewed and kurtotic, they had to be transformed (Table B1-B8, Appendix B). Despite feeling that the data is being manipulated, this procedure is very common in order to ensure normality (Field, 2017, p. 372-373). When normally distributed, the data can be interpreted as normal and reliable in the sense that the data will behave similarly when used at different moments by different researchers. Thus, enhancing the reliability of the research.

3.5. Research ethics

As master student at the Radboud University, the Code of Conduct on Scientific Practice has to be adhered to at all times (Vereniging van Universiteiten, 2014). This code of conduct requires researchers to be for example honest and ethical, professional, critical about self and others, and being respectful to those involved (Vereniging van Universiteiten, 2014). The learnings drawn from the Business Administration Pre-Master and the current Master trajectory help to adhere to this scientific behavioural code of conduct. Furthermore, the APA 7th edition guidelines from the American Psychological Association regarding citing and reporting are followed at all times.

In general, research ethics can mean being truthful, fair, wise and to prevent research misconduct. Pimple (2002) argues six key domains within responsible conduct of research (RCR), most important to this thesis being: scientific integrity and social responsibility. Scientific integrity is maintained through prevention of falsification, fabricating and plagiarism. This means giving honest credit to the rightful contributor and reporting truthful results without forging the results in any direction.

Another important aspect of research ethics is respondent consent (Smith, 2003). However, as this thesis makes use of secondary data, no physical interaction with respondents is conducted. Despite, respondents' privacy within the dataset should be guaranteed at all times. Company information and respondents' personal information will be fully anonymised. In addition, from the side of the researcher, it will be prevented that the dataset is shared with or leaked to non-relevant others. It is the researchers' social responsibility to protect information from being shared unnecessarily.

4. Results

Within this section, the results of the statistical analyses will be reported alongside the descriptive and frequency statistics. The data forms the centre of this thesis as it will proclaim whether the proposed hypotheses are supported or rejected.

This chapter will start off with the univariate statistics analyses. Second, the assumptions for executing the regression analysis will be covered for the low- and high-tech sample. Lastly, the actual regression analyses will be conducted, and the results will be reported.

4.1. Univariate analysis

To understand the possible relationship between variables, it is important to first understand the distribution and nature of the variables individually. Within this univariate analysis, the normality of all involved variables is checked. As elaborated in the validity section, transforming the variables can solve non-normality. Appendix B presents all normality checks for both the high-tech sample and the low-tech sample.

4.1.1. Network position VC

The first variable is Network Position and is of continuous nature (Chapter 3), therefore the univariate statistics will be presented as descriptive statistics. The transformations of the Degree Centrality and Betweenness Centrality measures are displayed in Appendix B. As can be seen, the square root transformation (SQRT) has the most profound skewness-kurtosis improvement compared to the original values (see Appendix B, table B1-B8). According to Field (2017), the variable which has the most improvement compared to the original should be selected for further analysis. In this case, the SQRT transformation will be used in further analysis in the low- and high-tech sample. The SQRT-transformation is powerful at solving data distributions that are (highly) right-skewed or positively skewed (Field, 2017). This is the case for most variables in this thesis. Therefore, the outcomes of the transformations make sense from a statistical perspective. In table 6 below, the original variable is shown alongside the best transformation for comparative matters.

Table 6. *Descriptive statistics of measures of variable Network Position*

	N	*Range*	*Minimum*	*Maximum*	*Mean*	*Std. Deviation*
High tech sample						
Degree centrality Average SQRT	244	.18	0	.18	.0650	.03734
Degree centrality Average	244	.033512	0	.0335152	.0056149	.0058690
Betweenness centrality Average SQRT	244	.09	0	.09	.0252	.01849
Betweenness centrality Average	244	.0082406	0	.0082406	.0009745	.0013794
Valid N (listwise)	244					
Low tech sample						
Degree centrality Average SQRT	263	.22	0	.22	.0681	.04971
Degree centrality Average	263	.0490720	0	.0490720	.0070954	.0090819
Betweenness centrality Average SQRT	262	.12	0	.12	.0281	.02312
Betweenness centrality Average	262	.0136744	0	.0136744	.0013193	.0020777
Valid N (listwise)	262					

4.1.2. Innovative performance start-ups

The second variable is Innovative Performance of Start-up, measured in patents and trademarks, all continuous level variables. In Appendix B, it is shown that the Ln-transformation is the most suitable transformation for the low-tech and the high-tech sample. Therefore, this variable is used in further analysis.

Table 7 below shows the descriptive statistics of this variable, divided into trademarks and patents. Start-ups in the high-tech sample have on average 1,94 trademarks registered and own 3,17 patents. In comparison, start-ups in the low-tech sample registered 1,82 trademarks and own 1,56 patents. This constitutes a difference of approximately 50% in owned patents, which is in line with widespread literature stating that patents are of more impact and therefore are more used in high-tech industries (Rothaermel & Hess; 2007 Hirsch-Kreinsen, 2008, chapter 2 and 3).

Table 7. *Descriptive statistics of variable Innovative Performance of Start-ups.*

	N	*Range*	*Minimum*	*Maximum*	*Mean*	*Std. Deviation*
High tech sample						
Trademarks Registered Ln	193	5.42	0	5.42	1.9406	1.220
Trademarks Registered	244	225	0	225	11.55562	22.268
Patents Granted Ln	221	6.9	0	6.9	3.1647	1.527
Patents Granted	244	995	0	995	56.508	106.886
Valid N (listwise)	173					
Low tech sample						
Trademarks Registered Ln	249	6.26	0	6.26	1.8246	1.34366
Trademarks Registered	265	523	0	523	16.61	42.241
Patents Granted Ln	77	5.58	0	5.58	1.5588	1.42010
Patents Granted	265	264	0	264	4.11	18.857
Valid N (listwise)	69					

4.1.3. Technology intensity

The third variable is technology intensity, a dichotomous categorical variable where "0" means low-tech and "1" means high-tech. Hence, the noteworthy statistics are the frequency statistics (Table 8). In order to maximize the accuracy of results from categorical variables, the group sizes should be equal (Hair et al., 2018). This means that the largest group must be < 1.5 bigger than the smallest group in order to have equal group sizes (Hair et al., 2018). The largest group (265) is only 1.09 bigger than the smallest group (244), thus making the groups equally sized.

Table 8. *Frequency statistics of variable Technology intensity*

	Frequency	
	0	1
High tech sample	0	244
Low tech sample	265	0

4.2. Assumptions of Regression Analysis

4.2.1. Linearity of the relationship

The first assumption of regression analysis is concerned with linearity. As regression analysis estimates a straight regression line in which residuals show the least deviation from the estimated line, linearity of the relationship is required (Osborne & Waters, 2002). This assumption is being tested via scatterplots in SPSS. Figure 9 represents the low-tech sample and figure 10 represents the high-tech sample.

Figure 9. *Scatterplot for linearity criterium assessment of low-tech sample.*

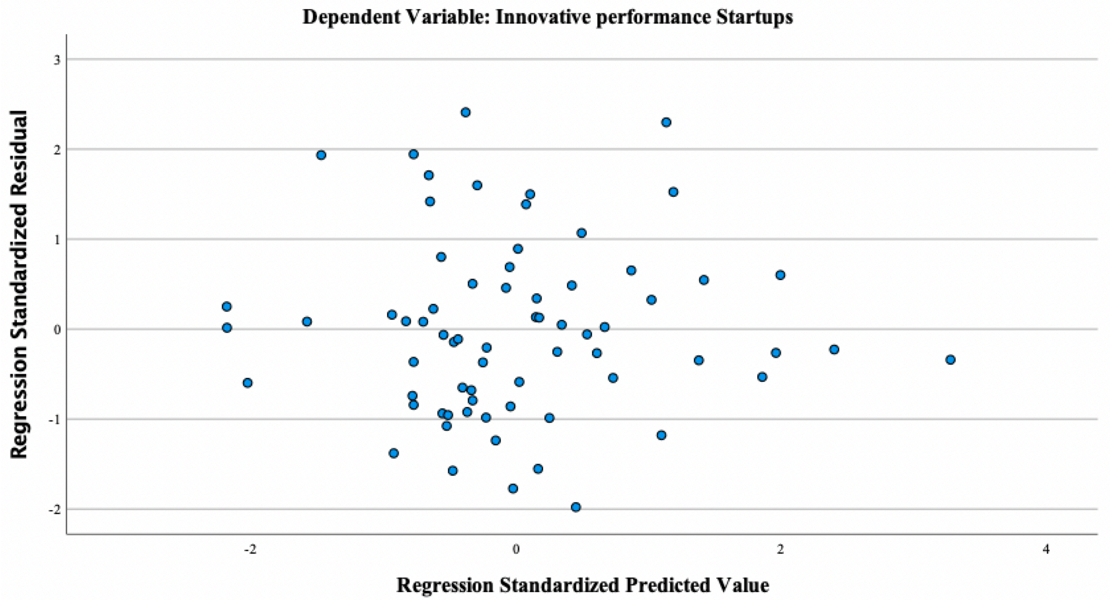
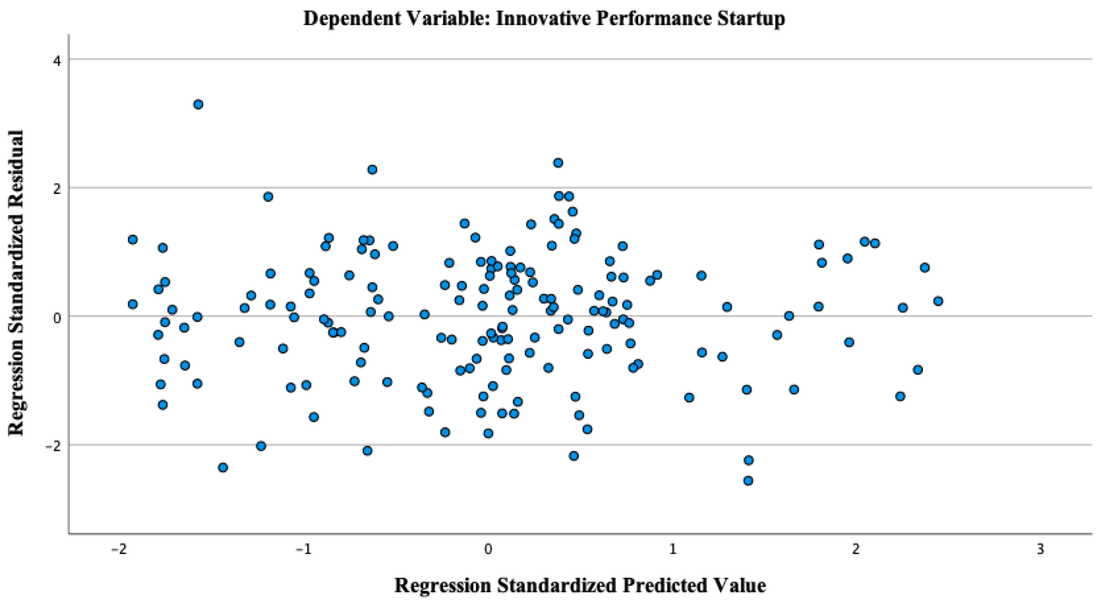


Figure 10. *Scatterplot for linearity criterium assessment of high-tech sample.*



The scatterplots are built in SPSS using the ZRESID and ZPRED values of the variable. According to Hair et al. (2018), a linear relationship between the dependent and independent variables is represented by dots showing no clear pattern. In other words, if all the positive and negative residual scores are evenly distributed along the null line, the relationship can be classified as linear. Figures 9 and 10 show that the far majority of residuals are distributed uniformly between the section from +2 to -2. The high- and low-tech sample do not show a clear pattern and therefore it can be assumed that the linearity assumption is sufficiently met.

4.2.2. Constant variance of residuals

The range of residuals should be constant, which is called homoscedasticity of the variance of residuals (Hair et al., 2018). The range of residuals for lower scores on the horizontal axis should be the same as the range of residuals for higher scores on the horizontal axis (Osborne & Waters, 2002). Heteroscedasticity is the opposite of homoscedasticity, which refers to dots presenting different patterns across the scores on the horizontal axis. In order to check this assumption, figures 9 and 10 are used again. For the low-tech sample, the range of residuals is approximately the same across the horizontal axis scores -2, 0 and +2 (Figure 9). Only few outliers are present on higher scores on the X-axis. It is not close to perfect homoscedasticity but there is no clear-cut shape visible as well. For the high-tech sample, the dots are behaving more like homoscedastic data since the range of the residuals in even across all scores on the horizontal axis (Figure 10).

According to Field (2017), there is not a real solution for heteroscedasticity, but it should be reported by the researcher and kept in mind while interpreting the results. In sum, the variables from both samples met the assumption, but for the low-tech research sample due diligence is advised since it is somewhere in between homo- and heteroscedastic patterns.

4.2.3. Independence of the error terms

The third assumption that is not very robust to violations is the independence assumption. The independence of the error terms indicates to what degree an estimated value relates to the value of another estimation (Hair et al., 2018). In other words, how much are the values determined by other values. The error term is a specific part of the variance that cannot be explained by the independent variables in the model (Field, 2017). This assumption is checked in SPSS in the *Residuals Statistics* table in the row *Std. Predicted Value* from the linear regression analysis output. In this table, the mean should be 0.000 and the standard deviation should be 1.000. As can be seen in Appendix C table C1-C2, the high- and low-tech sample report means of 0.000

and standard deviations of 1.000. This means that the error terms are independent and the dependence of residuals on it is mild, and the reliability of the estimations is sufficient, therefore meeting the third assumption.

4.2.4. Normality of residuals

It is crucial for proper regression analysis that the residuals of the scores are normally distributed. Non-normally distributed data can distort the reliability of results since outliers increase the chance of Type I or Type II errors and decreases the power of the analysis (Field, 2017; Osborne & Waters, 2002). Normality can be checked via PP-plots in SPSS. This is a probability plot that shows the cumulative scores compared to a perfect normally distributed trendline. The data is normally distributed when the residuals are closely to this line. As can be seen in figure C1-C6 of Appendix C, all variables show residuals in close proximity to the normality line. Therefore, all variables from the high- and low-tech sample have met the fourth assumption of regression analysis.

4.2.5. Multicollinearity

Multicollinearity is concerned with the prohibited correlation among multiple independent variables included in a model. In this thesis, the independent variables are *Average Degree Centrality* and *Average Betweenness Centrality*. In order to meet the multicollinearity criterium, the tolerance statistic should be > .100, which is the case for all IVs (Table 9). The VIF value, however, has an ambiguous threshold. Some researchers apply a preferred threshold of 4 (Hair et al., 2018), others use a threshold of 5 or a maximum of 10 (Kim, 2019). As can be seen in table 9, no VIF value exceeds 10 and thus meets this assumption.

High correlations between variables can also indicate multicollinearity issues and will be checked (Field, 2017) (Appendix D, table D4 and Appendix E, Table E4). The correlations matrices report high correlation between degree centrality and betweenness centrality ($\rho = .921$ and $\rho = .927$). In addition, number of lead investors and number of investors are also highly correlated. Lastly, in the low-tech sample, number of investors is highly correlated with number of lead investors ($\rho = .599$) and degree centrality ($\rho = .773$). High collinearity scores cause the β-slope to become unpredictable (Field, 2017). Therefore, caution should be applied while interpreting the results since the statistics show presence of multicollinearity.

The presence of multicollinearity between Degree and Betweenness Centrality is explainable as they are ought to measure a phenomenon of the same nature, namely network centrality.

Combining both variables is not an option since they measure two distinct aspects of a network. From the formulas of NetworkX in the methodology section, it can be seen that the two variables are defined clearly and represent two different aspects. In addition, previous research also uses degree and betweenness centrality, or even all five popular centrality measures as explained in chapter 3. Therefore, this thesis continues to keep using degree centrality and betweenness centrality separately.

Table 9. *Multicollinearity statistics of the low- and high-tech sample*

Independent variable	**Collinearity statistics**	
	Tolerance	VIF
Low-tech sample		
Average degree centrality	.123	8.119
Average betweenness centrality	.123	8.119
High-tech sample		
Average degree centrality	.149	6.706
Average betweenness centrality	.149	6.706

4.3. Results of the Regression Analysis

The regression analysis is conducted in order to test what the effect of a certain network position of a VC is on the innovative performance of start-ups. This effect is being tested in two samples representing a low-tech industry and a high-tech industry. In table 10, the results of the regression analysis are presented. Models 1 – 3 analyse the high-tech sample, and models 4 – 6 analyse the low-tech sample. For every trio, the first model solely consists of control variables, the second model adds the first independent variable “Degree Centrality”, and the last model includes the second independent variable “Betweenness Centrality”.

4.3.1. Hypothesis 1: High-tech industry model

The first hypothesis tests the effect of network position on innovative performance of VC-backed start-ups in a high-tech industry environment. The hypothesized expectation is that Degree centrality positively affects on the innovative performance of VC-backed start-ups.

4.3.1.1. Model summary and model fit

In the high-tech sample, three models were tested. Table 10 shows that none of the models have a high proportion of explained variance with an adjusted R^2 of .036, .029 and .023. Regarding model 3, this indicates that the model explains only 2.3% of all variance in the model. In addition, the F – value change from model 1 to model 2 and model 2 to model 3 is not significant (Appendix D, table D1). Looking at the F-values, model 1 is significant at $p \leq .10$, but model 2 and 3 are not useful for interpretation at a CI of 95% with $F(5, 151) = 2.154$, $p = .062$; $F(6, 150) = 1.786$, $p = .106$ and $F(7, 149) = 1.521$, $p = .164$, respectively (Appendix D, table D2). Nonetheless, the results will be reported in order to support or reject the hypotheses in this thesis for the sake of discussion of the results.

4.3.1.2. Hypothesis 1a

Table 10 shows the coefficients for the independent and control variables in the high-tech sample. The adjusted R^2-value decreases from 0.036 (3.6%) to 0.029 (2.9%) when the main effect of Degree Centrality is added. In addition, the F-value loses its significance, indicating a decreasing usefulness and an insufficient goodness of fit of model 2.

The main effect of Degree Centrality on the Innovative Performance of Start-ups is non-significant ($\beta = 1.515$, $t = .117$, $p = .907$), therefore hypothesis 1a is not supported. However, the hypothesized direction of the effect is in line with the results. When Degree Centrality is the sole IV, model 2, it also reports an insignificant effect ($\beta = .668$, $t = .127$, $p = .899$). However, the reported direction is also in line with the hypothesis. In model 3, it can be noted when including Betweenness Centrality, the effect of Degree Centrality is more than double as strong compared to the coefficient of model 2.

4.3.1.3. Hypothesis 1b

When Betweenness Centrality is added in model 3, it can be seen that the adjusted R^2 decreases from 0.029 (2.9%) to 0.023 (2.3%). This makes sense because the added variable is non-significant and adds less explaining power than would have been by random chance (Field, 2017). However, for the purpose of this thesis, the results will be reported and discussed. Looking at model 3, the main effect of Betweenness Centrality on the Innovative Performance of Start-ups with $\beta = -1.707$ is non-significant as well ($t = -.072$, $p = .943$). Hypothesis 1b proposed a positive effect of Betweenness Centrality. The β-slope of Betweenness Centrality is negative, and the reported direction is not in line with hypothesis 1b. Therefore, hypothesis 1b is not supported.

Table 10. *Multiple Regression models and coefficients for high-tech and low-tech industry samples*

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	High-tech industry			Low-tech industry		
(Constant)	5.689****	5.656****	5.646****	6.348****	5.197****	4.644****
	(6.836)	(6.475)	(6.355)	(6.350)	(4.489)	(3.989)
Age of Company	–.003	–.003	–.003	–.008	–.002	.002
	(–.299)	(–.281)	(–.265)	(–.646)	(–.149)	(.187)
Low number of Employees	–1.034	–1.034	–1.035	–2.676****	–2.508***	–1.923**
	(–1.520)	(–1.515)	(–1.511)	(–3.405)	(–3.240)	(–2.367)
Medium number of Employees	.230	.223	.221	–1.954**	–1.766**	–1.298
	(.316)	(.305)	(.300)	(–2.338)	(–2.144)	(–1.548)
Number of Investors	.030	.028	.027	.132***	.050	.038
	(.703)	(.585)	(.565)	(2.939)	(.796)	(.620)
Number of Lead investors	.079	.080	.080	–.565***	–.514**	–.462***
	(.606)	(.610)	(.609)	(–3.181)	(–2.924)	(–2.666)
Degree Centrality Average		.668	1.515		17.549*	48.219**
		(.127)	(.117)		(1.858)	(2.640)
Betweenness Centrality Average			–1.707			–77.643*
			(–.072)			(–1.944)
Diagnostics						
R^2	0.067	0.067	0.067	0.277	0.320	0.365
Adjusted R^2	0.036	0.029	0.023	0.211	0.244	0.281
N	156	156	156	60	60	60
Model degrees of freedom	5	6	7	5	6	7
F - value	2.154*	1.786	1.521	4.205***	4.236****	4.358****

* $p \leq .10$; ** $p \leq .05$; *** $p \leq .01$; **** $p \leq .001$; (Constant) = Innovative Performance of Start-ups; accompanying t-values in parentheses.

4.3.1.4. Control variables

Models 1 – 3 show non-significance for all control variables ($p \leq .05$). However, throughout all models, the dummy categories for the control variable Number of Employees show an inclined slope as the start-up size increases. In other words, as the start-up grows bigger, the effect on Innovative Performance of Start-ups also increases. In fact, when a start-up grows from low employee size to medium employee size, the effect shifts from negative (around -1.0 to positive (around .22). Visualisation of the effect of number of employees in the high-tech industry is shown in Appendix F, figure F1. However, the scores are insignificant and since the constant is highly significant, this effect is thought to be caused by other variables outside the model. The two control variables Number of Investors and Number of Lead Investors report insignificant and very marginal positive scores on Innovative Performance of Start-ups.

4.3.2. Hypothesis 2: Low-tech industry model

The last hypothesis suggests a positive relation between Betweenness centrality (H2a) and Innovative Performance of Start-ups and a positive relation between Degree centrality (H2b) and Innovative Performance of Start-ups. Models 4 – 6 are allocated to the second hypothesis.

4.3.2.1. Model summary and model fit

Consistent with the first hypothesis, three models are composed. Model 4 includes only the control variables. Model 5 adds Degree Centrality as independent variable, whereas model 6 includes both Degree Centrality and Betweenness Centrality as independent variables.

Table 10 shows that the composed models for the low-tech sample report significant adjusted R^2 scores of $R^2 = .211$ for model 4; $R^2 = .244$ for model 5, and $R^2 = .281$ for model 6. This means that respectively 21.1%, 24.4% and 28.1% of the variance is explained by the variables in the models. As, the F-value changes between models 4 – 6 are all significant (at $p \leq .05$, model 5 and 6 at $p \leq .10$), the third model will be used for further interpretation (Appendix E, table E1). Looking at the ANOVA-output (Appendix E, table E2), all three models are significant and can be interpreted as useful with $F(5, 55) = 4.205$, $p \leq .05$; $F(6, 54) = 4.236$, $p \leq .001$; and $F(7, 53) = 4.358$, $p \leq .001$.

4.3.2.2. Hypothesis 2a

The adjusted R^2 value increases from 0.244 (24.4%) in model 5 to 0.281 (28.1%) in model 6 when Betweenness Centrality is added. This indicates that the added variable is significant and improves the explanatory power of the overall model. In addition, an adjusted R^2 of 28.1% is interpreted as high. The main effect of Betweenness Centrality on Innovative Performance of Start-ups is significantly negative ($ß$ = -77.643, t = -1.944, $p \leq .10$) (Table 10). The $ß$-slope is the amount with which Innovative Performance of Start-ups will change when the independent variable increases one unit. For $H2_a$, this means that every unit Betweenness Centrality increases, Innovative Performance of Start-ups decreases with 77.643. The hypothesis is that Betweenness Centrality has a positive effect on Innovative Performance of Start-ups. Despite being significant, the reported result is in the opposite direction from the hypothesis. Hence, hypothesis 2a is not supported.

4.3.2.3. Hypothesis 2b

For this hypothesis, one has to look at model 4 and model 5 (Table 10). First, when Degree Centrality is added, the adjusted R^2 increases from 0.211 (21.1%) to 0.244 (24.4%). Similar to hypothesis 2a, this means that adding this variable improves the model and reports significant results. The main effect of Degree Centrality on Innovative Performance of Start-ups is positive with a regression slope of $ß$ = 48.219 and is significant (t = 2.640, $p < .05$). This indicates that every unit Degree Centrality goes up, Innovative Performance of Start-ups increases with 48.219. Hypothesis 2b proposed a positive $ß$-slope for Degree Centrality, which is in line with the reported results. As the result is significant and in line with the hypothesized direction, hypothesis 2b is supported.

4.3.2.4. Control variables

The low-tech sample reports significant results for several control variables. First, throughout models 4 – 6, an inclined $ß$ – slope can be noticed in the variable Number of Employees. The dummy Low number of employees has a significant negative effect in all three models. Dummy variable Medium number of employees has a significant and less negative effect in model 4 and model 5. Thus, an inclined regression effect is noticed from small to medium number of employees (Appendix F, figure F2). Post-hoc analysis of all dummy categories, including High number of employees, also shows an inclined effect that shifts from a negative to a positive effect from medium to high number of employees (Appendix F, figure F4). When adding Betweenness Centrality in model 6, the inclined slope is still present but not significant anymore. The implications of this finding will be touched upon in the discussion section.

Secondly, control variable Number of Lead Investors report a significant negative effect in model 6 ($ß$ = -.462, t = -2.666, $p \leq .01$). Indicating that start-ups that are backed by a higher number of lead investors, report lower scores on Innovative Performance of Start-ups.

4.3.3. Validity and reliability

After the analysis the validity and reliability can be assessed again. Low F-values and low *adjusted* R^2-value give indication that the high-tech models do not fit the data very well. However, F-values of the low-tech models are highly significant (at least $p \leq .01$) and report substantial explanatory power with *adjusted* R^2-values ranging from 21.1% to 28.1%. Reliability has been checked via the pre-analysis assumptions. In addition, the positive R^2-change values indicate positive improvements for the low-tech models. Lastly, the difference in sample size could affect the reliability and validity of the outcomes. Smaller samples have less statistical power and have less chance of finding true accurate results (Field, 2017). This should be taken into account when interpreting the results. However, both samples are still above the minimum required sample size of 5 observations per variable (5 observations * 8 variables = a minimum of 40 observations) or the minimum of 30 observations (Field, 2017).

4.3.4. Overview

Table 11 shows a complete overview of all tested hypotheses and its outcomes on (1) the hypothesized direction and (2) the significance level. It can be noted that two out of the four hypotheses are supported in terms of hypothesized direction. When looking at the significance level, one out of the four hypotheses are supported. Non-significant results cannot be generalized beyond the research samples and can be found based on mere luck.

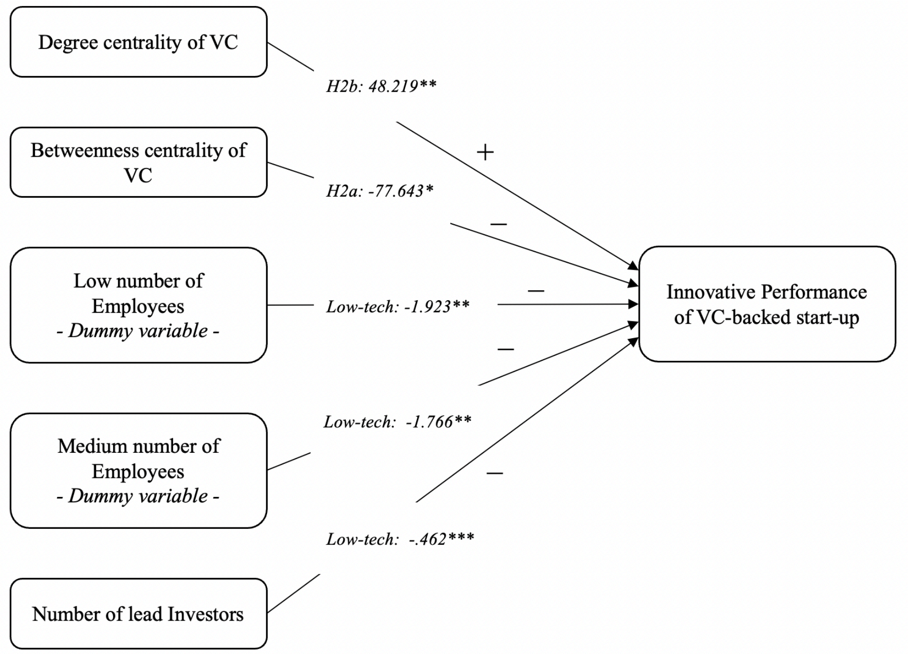
Table 11. *Overview of tested hypotheses*

Hypothesis	*Relationship*	*Direction*	*p-value*
H1a	Degree Centrality > positive > IPS[1]	Supported	Not significant
H1b	Betweenness Centrality > positive > IPS	Not supported	Not significant
H2a	Betweenness Centrality > positive > IPS	Not supported	Significant
H2b	Degree Centrality > positive > IPS	Supported	Significant

[1]Innovative Performance of Start-ups

Figure 11 shows the significant relationships of the independent variables on the dependent variable Innovative Performance of VC-backed start-ups.

Figure 11. *Overview of tested hypotheses*



5. Discussion and conclusion

The purpose of this thesis is to identify in what way the effect of a VC's network position on the innovative performance differs across high-tech and low-tech industries. Utilizing the previous research on Social Network Theory, this thesis develops two models on how network position affects innovative performance. The intend of this structure is to create equal models enabling comparison and exploring the differences between a high-tech and low-tech industry. In other words, this thesis aims to answer the following research question: *How does the effect of the network position of a venture capitalist on the innovativeness of VC-backed start-ups differ between high- and low-tech industries?*

5.1. Discussion

To begin with, the results of the high-tech industry will be discussed. Within the high-tech sample, a higher presence of strong ties in the network of VCs enhances the innovative performance of start-ups backed by that VC (Hypothesis 1a). This implies that VCs with a higher number of strong relations to other VCs in their network, can stimulate innovativeness of the start-up. However not significant, this is in line with the expectations derived from previous research (Abell & Nisar, 2007; Ferrary & Granovetter, 2009; Zhang & Wang, 2013). For example, Abell & Nisar (2007) find that the two most important network measures for explaining the performance of VC-backed companies are indegree and outdegree centrality (p.

931). Similar to this thesis they are degree centrality measures. This thesis adds the notion that this relationship also holds for the high-tech sample included in this thesis. However, the findings are not significant and therefore are not interpretable as such.

While strong ties have a positive effect, the presence of weak ties has the opposite effect. This implies that start-ups backed by VCs with high presence of weak ties are associated with negative effects on their innovative performance (Hypothesis 1b). Interestingly, these findings are not in concur with the widespread accepted works from Granovetter (1973) or Burt (1992). However, the results do find justification for the concept of information asymmetry. The social network of investors has the function of sharing information in order to validate it and increase investment performance (Xue et al., 2019). When ties are not strong enough to be based on trust, transparency and honesty, information asymmetry occurs (Shane & Cable, 2002). Start-ups that rely on VCs with many weak ties could be subject to wrongful interpretation of information and miss out on business opportunities. Meanwhile, start-ups with different VCs might seize these opportunities, gain a superior position and threaten the survival of the other start-up (i.e., the start-up with a weak tied-VC). Again, statements about the results should be interpreted with the nuance that they reported non-significance and high multicollinearity scores.

When looking at the findings in the low-tech sample, VCs with a high number of weak ties in their network have a negative impact on the innovative performance of start-ups backed by that VC (Hypothesis 2a). A possible explanation for this is that innovation in low-tech industries is often interdependent on high-tech firms or institutions (e.g., Hirsch-Kreinsen et al., 2011). However, a main barrier for this type of high- and low-tech partnerships could be the mutual lack of trust (Parmentola et al., 2020), which is hard to establish with only weak ties. The results are contrary to the hypothesized direction. Nonetheless, not surprising since the debate on weak or strong ties in low-tech industries is unconclusive and equivocal. The findings from H2a are in line with the strong tie-perspective of the debate (Abell & Nisar, 2007; Pavitt, 1984). The research of Abell & Nisar (2007) indicates that indirect relations, weak ties, are "much less important in the way venture capital industry is organized" (p. 931).

This could mean that strong ties are more important than weak ties in low-tech industries. This thesis discovered that VCs with high numbers of strong ties in their network are highly able to boost the innovative performance of start-ups in their portfolio (Hypothesis 2b). This implies that it is preferable for low-tech start-ups to seek investment at VCs with strong relationships based on trust and frequent contact instead of many loose relationships with business acquaintances. The proposition was made that low-tech firms seek innovations elsewhere and implement it backwards in their own industry. The findings are in line with research from Hirsch-Kreinsen (2008), stating that firms in low-tech industries can collaborate with firms in high-tech industries and jointly create innovations. However, to establish collaborative partnerships, strong ties based on trust are needed (Som, 2012). In turn, strong ties decrease the risk of using misinformation as basis for innovation (Shane & Cable, 2002). Hence, the importance for start-ups in low-tech industries to find VCs with many strong relationships with other VCs. One possible explanation could be due to the type of knowledge and the type of innovation in low-tech industries. As explained by Hirsch-Kreinsen (2008), knowledge is more company-specific, practically oriented and based on "ongoing operational processes" (p. 27). Consequently, innovations based on company specific knowledge are themselves often company specific. With highly company specific knowledge, analysing the potential of an innovation could be more difficult when relationships are less frequent, not based on trust and superficial.

Regarding the control variables, the variable Size of start-up was divided into three dummy variables: low, medium and high number of employees. In the low-tech industry, it can be stated that the size of a start-up has a significant effect on the innovative performance of that start-up. The effect is negative for small and medium start-ups but is positive for large start-ups. More specific, the bigger the start-up's employment pool grows, the less negative network position will impact innovative performance (Appendix F, figure F4). In the high-tech industry, the effect has a similar positive slope, but is insignificant. Since one of the dummy categories reports significant results, it can be said that the size of start-ups in low-tech industries impacts the innovative performance of the start-up. Exploiting economies of scale could be a reason why larger start-ups report higher innovative performance, as they can spread the costs of innovations over larger scale and yield more performance from it (Audretsch et al., 2011). Regarding the second control variable, Age of company, it can be stated that the older a start-up is, the higher their innovative performance. As this effect is very small and barely significant, its impact on the relationship can be questioned. Lastly, the significant negative effect of Number of Investors contradicts the expectation that more VCs lead to more innovative performance of the start-up. However, Wang & Wang (2012) argue that the entrepreneurs' incentives to work are reduces when the size of an investors syndicate in the company grows.

5.1.1. Comparison of the high- and low-tech industries and answering the research question

This thesis had the aim to execute a comparative study between a high-tech and a low-tech industry. As the results are reported and discussed separately, now a comparison can be made and an answer to the research question can be formulated. The research question is: *How does the effect of the venture capitalist's network position on the innovativeness of VC-backed start-ups differ between high- and low-tech industries?*

The answer on the research question is two folded.

Firstly, the effect of a VC's network position on the innovative performance of VC-backed start-ups differs across high- and low-tech industries. Table 12 below shows an overview of the most important differences and similarities. Start-ups in low-tech industries gain substantially more innovative performance than high-tech start-ups when attached to a VC with strong trust-based relationships. The inverse is true as well: start-ups in low-tech industries are also substantially more harmed by selecting a VC with an emphasis on having weak ties. In addition, the network position reports a highly significant effect in the low-tech industry sample but reports no significant effect in the high-tech industry. This difference could stem from their different technology intensities.

Secondly, the samples also report differences between high- and low-tech industries on whether start-ups should seek VCs with strong or weak ties. The low-tech industry sample shows that start-ups should collaborate with VCs that have strong relations with others and that do not depend on weak ties. Within the high-tech industry, network position does not seem to have a significant impact on the innovativeness of the VC-backed start-up. Thus, the network position of VCs in low-tech industries has a different effect on innovative performance of start-ups than the network position of high-tech industry VCs.

However, heavy nuance should be applied to this conclusion. Due to insignificant results of the high-tech sample, multicollinearity issues and unequal sample sizes, it should be concluded that the data is not sufficient to make an equal comparison. Based on this, the statements regarding the comparison of high- and low-tech industries should be interpreted with care. It can only be said for certain that results of the high-tech sample do not report significant effects for adequate comparison of the results between high-tech and low-tech industries. It could be that the theoretically expected effect of network position in high-tech industries in practice is present, but that the multicollinearity issues and sample size were insufficient to show this effect within this thesis.

Table 12. *Comparison between industries on key findings*

	Low-tech industry	*High-tech industry*
Preferred presence of ties in VCs network	Strong ties	Strong ties
Effect of IPS[1]	48.219	1.515$^{N.S.}$
Best suited size of the start-up	High number of employees[2]	High number of employees
Generalizable to other industries[3]	Yes	No

[1] Innovative Performance of Start-ups
[2] See post-hoc analysis in Appendix F
[3] at $p \leq 0.10$.

5.2. Conclusion

Start-ups in low-tech industries benefit more from VCs with strong relationships. It was also shown that the inverse effect is present in low-tech industries, since the negative effect is also stronger in the low-tech sample. In the high-tech industry sample, no significant effect of network position was found.

5.2.1. Contributions to the body of literature

The findings of this thesis contribute theoretically to the body of literature in various manners. First, the findings contribute to theory as it is an attempt to end the equivocal debate on whether start-ups need VCs with strong or weak ties. The findings favour the perspective of strong ties in the low-tech industries and are in line with the conclusions of for example Abell & Nisar (2007), Ferrary & Granovetter (2009), Pavitt (1984) and Zhang & Wang (2013). In addition, the findings also reject the perspective of weak ties as they appear to have a negative influence on the innovative performance of start-ups. Therefore, this research does not only support one perspective but also refutes the other (for the low-tech retail population). Hence, the debate could be a step closer to giving conclusive answers on the question what network position of a VC is most favourable for start-ups. It is important to state that this cannot be proclaimed for the high-tech industry, due to insignificant results. Secondly, the methodological approach of this thesis contributes to the body of literature due to its comparative nature. Often, research is designed to separately study high-tech start-ups (Bertrand-Cloodt et al., 2011; Colombo & Grilli, 2010) or low-tech (Abbasiharofteh & Dyba, 2018). Indisputably, comparative studies are executed similar to this thesis such as Hansen & Winther (2011) and Hirsch-Kreinsen et al.

(2006). However, such research is often focused on cross-border policies or innovation systems and less on a network perspective and venture capital. Instead, this thesis contributes to the body of literature and develops understanding on the entrepreneurial and venture capitalist level.

5.2.2. Practical implications

The practical implication of this thesis is that it builds entrepreneurial understanding of the consequences venture capital investment has on the start-up's innovative performance. Entrepreneurs reading literature on network theory will generally read that "networking" is something entrepreneurs should constantly participate in to stay ahead of competitors. This thesis aimed to bring nuance to these statements and help to formulate a realistic perspective on the effect of a VC's network on the entrepreneurial level. It is recommended to start-ups to appreciate technology intensity of the industry as a determining variable of the innovative performance. Especially as this thesis showed that different environments have a different effect on the innovativeness of start-ups. For low-tech start-ups it is recommended to seek VCs that have a network of strong ties. In addition, it seems to be the case that smaller start-ups gain less advantage from a strong tied VC than bigger start-ups do. The implication of this finding for entrepreneurs is that they should consider different types of external finance in different stages of the start-up.

By developing two models, this thesis provides clarity for entrepreneurs on how to consider what kind of VC should be chosen in different technology intense industries. The general assumed relation between a VC's network position and innovative performance of the VC-backed start-up might hold, but this relation and its strength are altered when industry technological intensity is factored in. In sum, understanding what consequences selecting VCs has and how it differs per environment is of utter importance for entrepreneurs as it can prevent their innovative performance from being harmed.

5.2.3. Theoretical recommendations

Without further investigation, it is impossible to determine if the innovative performance of VC-backed start-ups can also be explained by other factors then the VC's network position. It is possible that the scores on Innovative Performance of Start-ups are caused by other factors than Degree and Betweenness centrality measures. Future research is recommended to study the possible effect of network position on other performance parameters such as financial performance.

A second future research recommendation is to study and compare other industries from different technological intensity categories. Due to limited observation variety in the high-tech sample, results only say something about this sample. However, it could be that the findings differ across other high-tech industries such as the aeroplane and spacecraft manufacturing industry. In addition, research could also compare manufacturing industries (aeroplane and spacecraft) to non-manufacturing industries (e.g., software publishing). This enables intra-category comparisons and will shed light on how the effect of a VC's network position on the innovative performance of start-ups differs across other industries.

5.2.4. Societal implications

As illustrated in the introduction section, start-ups have contributed to job creation, economic growth and technological breakthroughs (Audretsch et al., 2011; Baumol, 2004). This thesis contributes to the survival chances of start-ups by studying how VCs can stimulate the performance of a start-up. When start-ups scale up, they can hire additional people and thereby contribute to the labour market and the people's personal development. In addition, start-ups stimulate the generation of knowledge that is needed to improve any thinkable aspect of society from health care to public transport. Understanding what effect venture capital might have on a start-ups innovative performance can improve the quality and quantity of innovations generated by start-ups and diffused into society.

5.2.5. Conclusion on the methodology

This study executed a comparative study in order to compare regression models from a high-tech industry and a low-tech industry. Despite non-significance of the high-tech models, regression analysis was an appropriate method for this thesis. Regression analysis allowed this research to identify the linear relation between all IVs on the DV separately. More importantly, it allowed for a comparison of the high-tech and low-tech regression models. Since the research question was concerned with the difference between two groups, the ability to compare models and results was highly important. In addition, the detailed and almost qualitative analysis of the selected industries was also an appropriate method. This way, a contextual layer was added onto the quantitative regression analysis by exploring the characteristics, history and relevance of the industries involved.

6. Limitations of the research and the findings

This thesis has potential limitations regarding the choice of research methodology and the generalization of the findings. These limitations will be addressed in the following critical reflection. First, due to overrepresentation of US based start-ups in the samples, findings can be biased towards start-ups based in the United States. The impact is not very large since the focus of this thesis is not on national differences.

Second, in order to ensure the quality of the samples and ensure adequate representation, the samples have been selected very carefully using methodological frameworks of the OECD (2016) and Hecker (2005). However, due to this specific sampling technique the samples are not random, and the results are harder to generalize beyond the industries used in the samples. The sampling choice was made to ensure that the samples really represented the high- and low-tech sample. However, this limits the generalizability of the results. In addition, a limitation of the results is that no adequate comparison can be made between high- and low-tech due to insignificant results. This limits the generalizability of the findings. However, results from the low-tech sample are generalizable and thus add value to the body of literature.

Third, this thesis uses Degree centrality to represent strong ties and Betweenness centrality to represent weak ties. However, some research also state that degree centrality can represent the presence of bridging ties (Baum & Rowley, 2008, p.192). This means that scholars are not conclusive about whether Degree centrality solely represents strong ties. It is possible that Degree centrality also represent the mere sum of ties a node has in a certain network. Nonetheless, Degree centrality has been chosen as measure for direct ties due to its wide use in previous research (Wang et al., 2015). Acknowledging the existence of this limitation enables future scholars to critically assess alternative measures to represent strong and weak ties in research.

Fourth, due to the Ln-transformation of the DV the sample size of the low-tech industry shrank from 265 and 244 observations to 69 and 157 usable observations. This did not harm the appropriateness of using regression analysis as sixty-nine observations are size wise still enough for regression analysis (see §4.3.3). However, unequal sample sizes can lower the statistical power, and this could result in unpredictable regression coefficients (see §4.3.3).

Lastly, time lag between the calculation of centrality measures ($t = 0$) and the moment patents or trademarks were granted (e.g., $t = -4$ yr.) exists. This means that the network position at the time of granting could have been different than the current value. This can limit the explanatory power of the found relation and should be taken into mind.

References

Abbasiharofteh, M., & Dyba, W. (2018). Structure and significance of knowledge networks in two low-tech clusters in Poland. Regional Studies, Regional Science, 5(1), 108–116. https://doi.org/10.1080/21681376.2018.1442244

Abell, P., & Nisar, T. (2007). Performance effects of venture capital firm networks. Management Decision, 45(5), 923–936. https://doi.org/10.1108/00251740710753729

Allen, M. (2004). *Understanding Regression Analysis*. Springer Publishing & Business Media.

Audretsch, D. B., Falck, O., Heblich, S., & Lederer, A. (2011). *Handbook of Research on Innovation and Entrepreneurship* (Revised ed.). Edward Elgar Publishing.

Babbie, E. (2010). Research design. The practice of social research, 85-88.

Batool, K., & Niazi, M. (2014). Towards a Methodology for Validation of Centrality Measures in Complex Networks. *PLoS ONE*, *9*(4), e90283. https://doi.org/10.1371/journal.pone.0090283

Baum, J., & Rowley, T. (2008). *Network Strategy* (1st ed., Vol. 25). Van Haren Publishing.

Baumol, W. J. (2004). Entrepreneurial Enterprises, Large Established Firms and Other Components of the Free-Market Growth Machine. *Small Business Economics*, *23*(1), 9–21. https://doi.org/10.1023/b:sbej.0000026057.47641.a6

Bellavitis, C., Filatotchev, I., & Souitaris, V. (2016). The Impact of Investment Networks on Venture Capital Firm Performance: A Contingency Framework. British Journal of Management, 28(1), 102–119. https://doi.org/10.1111/1467-8551.12162

Bertrand-Cloodt, D., Hagedoorn, J., & Van Kranenburg, H. (2011). The strength of R&D network ties in high-tech sectors – a multi-dimensional analysis of the effects of tie strength on innovation performance. *Technology Analysis & Strategic Management*, *23*(10), 1015–1030. https://doi.org/10.1080/09537325.2011.621294

Borgatti, S. (2005). Centrality and network flow. *Social Networks*, *27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Borgatti, S., & Halgin, D. (2011). On Network Theory. *Organization Science*, *22*(5), 1168–1181. https://doi.org/10.1287/orsc.1100.0641

Bottazzi, L., Da Rin, M., & Hellmann, T. (2008). Who are the active investors? Evidence from venture capital. *Journal of Financial Economics*, *89*(3), 488–512. https://doi.org/10.1016/j.jfineco.2007.09.003

Burt RS. (1992). *Structural Holes: The Social Structure of Competition*. Harvard University Press: Cambridge, MA.

Burt, R. (2000). The Network Structure Of Social Capital. *Research in Organizational Behavior*, *22*, 345–423. https://doi.org/10.1016/s0191-3085(00)22009-1

Caseiro, N., & Coelho, A. (2019). The influence of Business Intelligence capacity, network learning and innovativeness on startups performance. *Journal of Innovation & Knowledge*, *4*(3), 139–145. https://doi.org/10.1016/j.jik.2018.03.009

Chang, S. J. (2004). Venture capital financing, strategic alliances, and the initial public offerings of Internet startups. *Journal of Business Venturing*, *19*(5), 721–741. https://doi.org/10.1016/j.jbusvent.2003.03.002

Chen, Y. (2017). Dynamic ambidexterity: How innovators manage exploration and exploitation. *Business Horizons*, *60*(3), 385–394. https://doi.org/10.1016/j.bushor.2017.01.001

Cohen WM, Levinthal DA. 1990. Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* **35**: 128–152.

Colombo, M., & Grilli, L. (2010). On growth drivers of high-tech start-ups: Exploring the role of founders' human capital and venture capital. *Journal of Business Venturing*, *25*(6), 610–626. https://doi.org/10.1016/j.jbusvent.2009.01.005

Crosby, M. (2000). Patents, Innovation and Growth. *Economic Record*, *76*(234), 255–262. https://doi.org/10.1111/j.1475-4932.2000.tb00021.x

Crunchbase. (2018, January 19). *Patents & Trademarks*. https://about.crunchbase.com/marketplace/ipqwery/

Dalle, J., Den Besten, M., & Menon, C. (2017). Using Crunchbase for economic and managerial research. *OECD Science,* *Technology and Industry Working Papers*, *2017*(8), 1–29. https://doi.org/10.1787/6c418d60-en

De Clercq, D., Fried, V. H., Lehtonen, O., & Sapienza, H. J. (2006). An Entrepreneur's Guide to the Venture Capital Galaxy. *Academy of Management Perspectives*, *20*(3), 90–112. https://doi.org/10.5465/amp.2006.21903483

De Carolis, D.M. (2010) Technological Characteristics of Industries. In: Narayanan, V.K. and Colarelli-O'Connor, G. (eds.), *Encyclopedia of Technology and Innovation Management*. Chichester, West Sussex, UK: Wiley. pp. 77–79.

Delgado-Verde, M., Emilio Navas-López, J., Cruz-González, J., & Amores-Salvadó, J. (2011). Radical innovation from relations-based knowledge: empirical evidence in Spanish technology-intensive firms. *Journal of Knowledge Management*, *15*(5), 722–737. https://doi.org/10.1108/13673271111174294

Devigne, D., Vanacker, T., Manigart, S., & Paeleman, I. (2011). The role of domestic and cross-border venture capital investors in the growth of portfolio companies. *Small Business Economics*, *40*(3), 553–573. https://doi.org/10.1007/s11187-011-9383-y

DiMaggio, P., & Powell, W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, *48*(2), 147–160. https://doi.org/10.2307/2095101

Elfring, T., & Hulsink, W. (2007). Networking by Entrepreneurs: Patterns of Tie—Formation in Emerging Organizations. *Organization Studies*, *28*(12), 1849–1872. https://doi.org/10.1177/0170840607078719

Engel, D. (2004). The Performance of Venture-Backed Firms: The Effect of Venture Capital Company Characteristics. *Industry & Innovation*, *11*(3), 249–263. https://doi.org/10.1080/1366271042000265401

Esser, F., & Vliegenthart, R. (2017). Comparative Research Methods. *The International Encyclopedia of Communication Research Methods*, 1–22. https://doi.org/10.1002/9781118901731.iecrm0035

Ferrary, M., & Granovetter, M. (2009). The role of venture capital firms in Silicon Valley's complex innovation network. *Economy and Society*, *38*(2), 326–359. https://doi.org/10.1080/03085140902786827

Ferrati, F., & Muffatto, M. (2020). Using Crunchbase for Research in Francesco Ferrati and Moreno Muffatti. In M. Au-Yong-Oliveira & P. Costa (Eds.), *20th European Conference on Research Methodology for Business and Management Studies* (Revised ed., pp. 342–352). Academic Conferences and publishing limited. https://doi.org/10.34190/ERM.20.120

Field, A. (2017). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications.

Florida, R. L., & Kenney, M. (1988). Venture Capital, High Technology and Regional Development. *Regional Studies*, *22*(1), 33–48. https://doi.org/10.1080/00343408812331344750

Freeman, L. (1978). Centrality in social networks: I. Conceptual clarification. *Social Networks*, *1*(3), 215–239. https://www.bebr.ufl.edu/sites/default/files/Centrality%20in%20Social%20Networks.pdf

Garud, R., Tuertscher, P., & Van de Ven, A. H. (2013). Perspectives on Innovation Processes. 775–819. https://doi.org/10.1080/19416520.2013.791066

Goldschlag, N., & Miranda, J. (2019). Business dynamics statistics of High Tech industries. *Journal of Economics & Management Strategy*, *29*(1), 3–30. https://doi.org/10.1111/jems.12334

Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, *135*(1), 169–190. https://doi.org/10.1016/j.jfineco.2019.06.011

Grabowski, H. (2002). Patents, Innovation and Access to New Pharmaceuticals. *Journal of International Economic Law*, *5*(4), 849–860. https://doi.org/10.1093/jiel/5.4.849

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, *78*(6), 1360–1380. https://doi.org/10.1086/225469

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic Management Journal*, *21*(3), 203–215. https://www.jstor.org/stable/3094185

Hagedoorn, J., & Frankort, H. (2008). The gloomy side of embeddedness: The effects of overembeddedness on inter-firm partnership formation. *Network Strategy*, *25*, 503–530. https://doi.org/10.1016/s0742-3322(08)25014-x

Hair, J., Black, W., Babin, B., & Anderson, R. (2018). *Multivariate Data Analysis* (8th ed.). Cengage Learning Emea.

Haltiwanger, J., Hathaway, I., & Miranda, J. (2014). Declining Business Dynamism in the U.S. High-Technology Sector. *SSRN Electronic Journal*. Published. https://doi.org/10.2139/ssrn.2397310

Hannan, M., Freeman, J. (1984). Structural inertia and organizational change. Am. J. Sociol. 49, 149–164.

Hansen, M. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, *44*(1), 82. https://doi.org/10.2307/2667032

Hansen, T., & Winther, L. (2011). Innovation, regional development and relations between high- and low-tech industries. *European Urban and Regional Studies*, *18*(3), 321–339. https://doi.org/10.1177/0969776411403990

Hecker, D. (2005). High-technology employment: A NAICS-based update. *Monthly Labor Review*, *128*(57), 57–72. https://www.bls.gov/opub/mlr/2005/07/art6full.pdf

Hirsch-Kreinsen, H. (2008). “Low-Tech” Innovations. *Industry and Innovation*, *15*(1), 19–43. https://doi.org/10.1080/13662710701850691

Hirsch-Kreinsen, H. (2015). Patterns of knowledge use in ‘low-tech’ industries. *Prometheus*, *33*(1), 67–82. https://doi.org/10.1080/08109028.2015.1062237

Hirsch-Kreinsen, H., Jacobson, D., & Robertson, P. (2006). ‘Low-tech’ Industries: Innovativeness and Development Perspectives—A Summary of a European Research Project. *Prometheus*, *24*(1), 3–21. https://doi.org/10.1080/08109020600563762

Hochberg, Y., Ljungqvist, A., & Lu, Y. (2007). Whom You Know Matters: Venture Capital Networks and Investment Performance. *The Journal of Finance*, *62*(1), 251–301. https://doi.org/10.1111/j.1540-6261.2007.01207.x

Hyytinen, A., & Toivanen, O. (2005). Do financial constraints hold back innovation and growth? *Research Policy*, *34*(9), 1385–1403. https://doi.org/10.1016/j.respol.2005.06.004

Jeng, L. A., & Wells, P. C. (2000). The determinants of venture capital funding: evidence across countries. *Journal of Corporate Finance*, *6*(3), 241–289. https://doi.org/10.1016/s0929-1199(00)00003-1

Kadushin, C. (2012). *Understanding Social Networks: Theories, Concepts, and Findings* (Illustrated ed.). Oxford University Press.

Kilduff, M. 2010. *Serendipity vs. strategy: A tale of two theories.* Pre- sentation, Intra-Organizational Networks Conference, April 10, University of Kentucky, Lexington.

Kim, J. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

Lerner, J. (1994). The Syndication of Venture Capital Investments. *Financial Management*, *23*(3), 16. https://doi.org/10.2307/3665618

Mendonça, S., Pereira, T., & Godinho, M. (2004). Trademarks as an indicator of innovation and industrial change. *Research Policy*, *33*(9), 1385–1404. https://doi.org/10.1016/j.respol.2004.09.005

Mikle, L. (2020). Startups and reasons for their failure. *SHS Web of Conferences*, *83*, 01046. https://doi.org/10.1051/shsconf/20208301046

Miozzo, M., & DiVito, L. (2016). Growing fast or slow?: Understanding the variety of paths and the speed of early growth of entrepreneurial science-based firms. *Research Policy*, *45*(5), 964–986. https://doi.org/10.1016/j.respol.2016.01.011

Murzacheva, E., & Levie, J. (2020). Entrepreneurial finance journeys: embeddedness and the finance escalator. *Venture Capital*, *22*(2), 185–214. https://doi.org/10.1080/13691066.2020.1767756

NetworkX. (2020). *NetworkX - NetworkX documentation*. https://networkx.org

NetworkX Developers. (2020). *NetworkX Algorithms Centrality Betweenness Centrality*. NetworkX. https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.centrality.betweenness_centrality.html#networkx.algorithms.centrality.betweenness_centrality

Nonaka, I., & Takeuchi, H. (1995). *The Knowledge-Creating Company* (1st ed.). Oxford University Press.

OECD. (2016). OECD Taxonomy of Economic Activities Based on R&D Intensity. *OECD Science, Technology and Industry Working Papers*, *2016/04*, 1–24. https://doi.org/10.1787/5jlv73sqqp8r-en

Office for National Statistics. (2021, February 19). *Internet sales as a percentage of total retail sales (ratio) (%)*. https://www.ons.gov.uk/businessindustryandtrade/retailindustry/timeseries/j4mc/drsi

Olabode, S., Bakare, A., & Olateju, O. (2018). An Assessment of The Reliability of Secondary Data in Management Science Research. *LASU Journal of Employment Relations & Human Resource Management*, *1*(1), 182–194. https://doi.org/10.36108/ljerhrm/8102.01.0102

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research, and Evaluation*, *8*(2), 1–5. https://doi.org/10.7275/r222-hv23

Owen-Smith, J., & Powell, W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*, *15*(1), 5–21. https://doi.org/10.1287/orsc.1030.0054

Parida, V., Pesämaa, O., Wincent, J., & Westerberg, M. (2016). Network capability, innovativeness, and performance: a multidimensional extension for entrepreneurship. *Entrepreneurship & Regional Development*, *29*(1–2), 94–115. https://doi.org/10.1080/08985626.2016.1255434

Park, J., & Bae, Z. (2017). When are 'sharks' beneficial? Corporate venture capital investment and startup innovation performance. *Technology Analysis & Strategic Management*, *30*(3), 324–336. https://doi.org/10.1080/09537325.2017.1310376

Parmentola, A., Ferretti, M., & Panetti, E. (2020). Exploring the university-industry cooperation in a low innovative region. What differences between low tech and high tech industries? *International Entrepreneurship and Management Journal*. Published. https://doi.org/10.1007/s11365-020-00671-0

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, *13*(6), 343–373. https://doi.org/10.1016/0048-7333(84)90018-0

Pimple, K. (2002). Six domains of research ethics. *Science and Engineering Ethics*, *8*(2), 191–205. https://doi.org/10.1007/s11948-002-0018-1

Pratch, L. (2005). Value-Added Investing: A framework for early stage venture capital firms. *The Journal of Private Equity*, *8*(3), 13–29. https://doi.org/10.3905/jpe.2005.516844

Reguera-Alvarado, N., & Bravo, F. (2017). The impact of directors' high-tech experience on innovation in low-tech firms. *Innovation*, *20*(3), 223–239. https://doi.org/10.1080/14479338.2017.1407220

Rogers, E. (1962). *Diffusion of Innovations* (1st ed.). Free Press.

Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rosenbusch, N., Brinckmann, J., & Müller, V. (2013). Does acquiring venture capital pay off for the funded firms? A meta-analysis on the relationship between venture capital investment and funded firm financial performance. *Journal of Business Venturing*, *28*(3), 335–353. https://doi.org/10.1016/j.jbusvent.2012.04.002

Rothaermel, F., & Hess, A. (2007). Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects. *Organization Science*, *18*(6), 898–921. https://doi.org/10.1287/orsc.1070.0291

Sapienza, H. J. (1992). When do venture capitalists add value? *Journal of Business Venturing*, *7*(1), 9–27. https://doi.org/10.1016/0883-9026(92)90032-m

Schilling, M. A. (2011). *Strategic Management of Technological Innovation* (3rd ed.). McGraw-Hill Education.

Shane, S., & Cable, D. (2002). Network Ties, Reputation, and the Financing of New Ventures. *Management Science*, *48*(3), 364–381. https://doi.org/10.1287/mnsc.48.3.364.7731

SICCODE. (2019, September 30). *ISIC Code 2100 Manufacture Of Pharmaceuticals, Medicinal...* SIC & NAICS Codes, Company Search, Business Lists - SICCODE.Com. https://siccode.com/isic-code/2100/manufacture-pharmaceuticals-medicinal-chemical-botanical-products

Smith, D. (2003, January). Five principles for research ethics. *Monitor on Psychology*, *34*(1). http://www.apa.org/monitor/jan03/principles

Som, O. (2012). *Innovation without R&D: Heterogeneous innovation patterns of non-R&D-performing firms in the German manufacturing industry.* Springer Gabler.

Sorenson, O., & Stuart, T. (2001). Syndication Networks and the Spatial Distribution of Venture Capital Investments. *American Journal of Sociology*, *106*(6), 1546–1588. https://doi.org/10.1086/321301

Statista. (2020a, November 5). *Global pharmaceutical industry - statistics & facts*. https://www.statista.com/topics/1764/global-pharmaceutical-industry/#dossierSummary__chapter4

Statista. (2020b, November 27). *E-commerce as share of total U.S. retail sales 2000-2018*. https://www.statista.com/statistics/185351/share-of-e-commerce-in-total-value-of-us-retail-wholesale-trade-sales/

Statista. (2021, January 6). *R&D expenditure per employee in U.S. sectors 2003-2017*. https://www.statista.com/statistics/240032/randd-expenditure-per-employee-in-us-sectors/

Stinchcombe, A. (1965). Social structure and organizations. In: March, J.G. (Ed.), Handbook of Organization. Rand McNally, Chicago, IL, pp. 153–193.

Sukamolson, S. (2007). Fundamentals of quantitative research. *Language Institute Chulalongkorn University*, *1*, 2-3.

Sun, W., Zhao, Y., & Sun, L. (2020). Big Data Analytics for Venture Capital Application： Towards Innovation Performance Improvement. *International Journal of Information Management*, *50*, 557–565. https://doi.org/10.1016/j.ijinfomgt.2018.11.017

Swanson, R., & Holton, E. (2005). *Research in Organizations; Foundations and Methods of Inquiry* (illustrated edition). Berrett-Koehler.

Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *International Journal of Academic Research in Management*, *5*(2). https://doi.org/10.2139/ssrn.3205035

The Economist. (2017, November 16). *A new class of startup is upending America's consumer-goods industry*. https://www.economist.com/business/2017/11/16/a-new-class-of-startup-is-upending-americas-consumer-goods-industry

Tolstoy, D., Nordman, E., Hånell, S., & Özbek, N. (2021). The development of international e-commerce in retail SMEs: An effectuation perspective. *Journal of World Business*, *56*(3), 101165. https://doi.org/10.1016/j.jwb.2020.101165

Tripsas, M. (1997). Unraveling the process of creative destruction: complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*, *18*(S1), 119–142. https://www.jstor.org/stable/3088213

United Nations. (2008). *International Standard Industrial Classification of All Economic Activities, Rev. 4* (No. 4). Department of Economic and Social Affairs. https://unstats.un.org/unsd/classifications/Eco n/Download/In%20Text/ISIC_Rev_4_publication_English.pdf

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

Uzzi B. (1997). Social structure and competition in interfirm networks. *Administrative Science Quarterly,* **42**: 35–67.

Valente, T., Coronges, K., Lakon, C., & Costenbader, E. (2008). How Correlated Are Network Centrality Measures? *Connections*, *28*(1), 16–26. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2875682/

Van Osnabrugge, M. (2000). A comparison of business angel and venture capitalist investment procedures: An agency theory-based analysis. Venture Capital, 2(2), 91–109. https://doi.org/10.1080/136910600295729

Vereniging van Universiteiten. (2014). *De Nederlandse Gedragscode Wetenschapsbeoefening: Principes van goed wetenschappelijk onderwijs en onderzoek.* VSNU. http://vsnu.nl/files/documenten/Domeinen/Onderzoek/Code_wetenschapsbeoefening_2004_%282014%29.pdf

Wang, H., Zhao, J., Li, Y., & Li, C. (2015). Network centrality, organizational innovation, and performance: A meta-analysis. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences de l'Administration*, *32*(3), 146–159. https://doi.org/10.1002/cjas.1316

Wang, L., & Wang, S. (2012). Endogenous networks in investment syndication. *Journal of Corporate Finance*, *18*(3), 640–663. https://doi.org/10.1016/j.jcorpfin.2012.03.004

Winborg, J., & Landström, H. (2001). Financial bootstrapping in small businesses. *Journal of Business Venturing*, *16*(3), 235–254. https://doi.org/10.1016/s0883-9026(99)00055-5

Wolf, M., & Terrell, D. (2016). The high-tech industry, what is it and why it matters to our economic future. *Beyond the Numbers: Employment and Unemployment*, *5*(8), 1–7. https://www.bls.gov/opub/btn/volume-5/the-high-tech-industry-what-is-it-and-why-it-matters-to-our-economic-future.htm

Xue, C., Dang, X., Shi, B., & Gu, J. (2019). Information Sharing and Investment Performance in the Venture Capital Network Community: An Empirical Study of Environmental-Social-Governance Start-Ups. *International Journal of Environmental*

Research and Public Health, *16*(6), 1023. https://doi.org/10.3390/ijerph16061023

Zawislak, P., Fracasso, E., & Tello-Gamarra, J. (2018). Technological intensity and innovation capability in industrial firms. *Innovation & Management Review*, *15*(2), 189–207. https://doi.org/10.1108/inmr-04-2018-012

Zhang, J., & Luo, Y. (2017). Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network. *Advances in Intelligent Systems Research*, *132*, 300–303. https://doi.org/10.2991/msam-17.2017.68

Zhang, M., & Wang, B. (2013). Industrial Technological Features, Strong and Weak Ties and Innovation Performance in Innovative Backward Areas: Evidence from Henan Enterprises Technology Centers [J]. *Science of Science and Management of S. & T*, *4*.

Appendices

Appendix A. Taxonomy of economic activities based on R&D Intensity and STEM-framework

	Manufacturing	R&D as % of GVA[2]	Non-manufacturing	R&D as % of GVA[2]
High R&D intensity industries	303[1]: Air and spacecraft and related machinery	31.69	72: Scientific research and development	30.39
	21: Pharmaceuticals	27.98	582[1]: Software publishing	28.94
	26: Computer, electronic and optical products	24.05		
Medium-high R&D intensity industries	252[1]: Weapons and ammunition	18.87	62-63: IT and other information services	5.92
	29: Motor vehicles, trailers and semi-trailers	15.36		
	325[1]: Medical and dental instruments	9.29		
	28: Machinery and equipment n.e.c.	7.89		
	20: Chemicals and chemical products	6.52		
	27: Electrical equipment	6.22		
	30X[1]: Railroad, military vehicles and transport n.e.c. (ISIC 302, 304 and 309)	5.72		
Medium R&D intensity industries	22: Rubber and plastic products	3.58		
	301[1]: Building of ships and boats	2.99		
	32X[1]: Other manufacturing except medical and dental instruments (ISIC 32 less 325)	2.85		
	23: Other non-metallic mineral products	2.24		
	24: Basic metals	2.07		
	33: Repair and installation of machinery and equipment	1.93		
Medium-low R&D intensity industries	13: Textiles	1.73	69-75X: Professional, scientific and technical activities except scientific R&D (ISIC 69 to 75 less 72)	1.76
	15: Leather and related products	1.65	61: Telecommunications	1.45
	17: Paper and paper products	1.58	05-09: Mining and quarrying	0.80
	10-12: Food products, beverages and tobacco	1.44	581[1]: Publishing of books and periodicals	0.57
	14: Wearing apparel	1.40		
	25X[1]: Fabricated metal products except weapons and ammunition (ISIC 25 less 252)	1.19		
	19: Coke and refined petroleum products	1.17		
	31: Furniture	1.17		
	16: Wood and products of wood and cork	0.70		
	18: Printing and reproduction of recorded media	0.67		
Low R&D intensity industries			64-66: Financial and insurance activities	0.38
			35-39: Electricity, gas and water supply, waste management and remediation	0.35
			59-60: Audiovisual and broadcasting activities	0.32
			45-47: Wholesale and retail trade	0.28
			01-03: Agriculture, forestry and fishing	0.27
			41-43: Construction	0.21
			77-82: Administrative and support service activities	0.18
			90-99: Arts, entertainment, repair of household goods and other services	0.11
			49-53: Transportation and storage	0.08
			55-56: Accommodation and food service activities	0.02
			68: Real estate activities	0.01

Note. Reprinted from “OECD Taxonomy of Economic Activities Based on R&D Intensity” by F. Galindo-Rueda and F. Verger, 2016, *OECD*, p.10. Copyright 2014 by OECD.

[Levels in thousands]

NAICS code	Industry	Employment 1992	Employment 2002	Employment 2012	Employment change: Change in level, 1992–2002	Employment change: Change in level, 2002–12	Employment change: Percent change, 1992–2002	Employment change: Percent change, 2002–12	Median annual wage, May 2004[1]
…	Total nonfarm wage and salary, all industries[2]	109,526	131,063	152,690	21,537	21,627	19.7	16.5	$28,770
…	Total, three levels of high-technology industries	13,415	14,422	16,067	1,006	1,646	7.5	11.4	([3])
…	Level-I industries	4,783	5,883	6,804	1,100	921	23.0	15.6	
3254	Pharmaceutical and medicine manufacturing	225	293	361	68	68	30.2	23.2	43,930
3341	Computer and peripheral equipment manufacturing	329	250	182	–79	–68	–24.0	–27.1	61,830
3342	Communications equipment manufacturing	210	191	201	–19	10	–9.0	5.4	45,520
3344	Semiconductor and other electronic component manufacturing	519	531	452	12	–79	2.3	–14.9	39,210
3345	Navigational, measuring, electromedical, and control instruments manufacturing	549	451	396	–98	–55	–17.8	–12.2	47,960
3364	Aerospace product and parts manufacturing	711	468	386	–242	–83	–34.1	–17.6	51,990
5112	Software publishers	114	256	430	142	174	125.0	67.9	69,880
5161	Internet publishing and broadcasting	16	35	49	19	14	116.1	41.1	53,470
5179	Other telecommunications	16	10	8	–6	–2	–39.2	–21.9	45,470
5181	Internet service providers and Web search portals	39	142	233	103	91	265.3	64.2	52,780
5182	Data processing, hosting, and related services	220	305	430	86	125	39.0	40.8	45,570
5413	Architectural, engineering, and related services	902	1,251	1,306	349	54	38.7	4.3	48,570
5415	Computer systems design and related services	445	1,163	1,798	718	635	161.3	54.6	63,350
5417	Scientific research-and-development services	490	537	573	47	36	9.7	6.7	57,890
…	Level-II industries	4,760	4,528	4,998	–231	470	–4.9	10.7	([3])
1131, 32	Forestry	10	10	10	0	0	.0	4.0	–
2111	Oil and gas extraction	182	123	88	–60	–34	–32.8	–27.8	49,290
2211	Electric power generation, transmission, and distribution	537	436	405	–101	–31	–18.8	–7.1	53,330
3251	Basic chemical manufacturing	246	171	140	–76	–31	–30.8	–18.0	45,970
3252	Resin, synthetic rubber, and artificial synthetic fibers and filaments manufacturing	151	114	89	–37	–26	–24.5	–22.6	42,730
3332	Industrial machinery manufacturing	142	132	125	–10	–6	–7.1	–4.7	39,480
3333	Commercial and service industry machinery manufacturing	138	132	141	–6	9	-4.6	6.6	35,940
3343	Audio and video equipment manufacturing	58	42	38	–16	–3	–27.7	–7.7	32,460
3346	Manufacturing and reproducing, magnetic and optical media	44	57	63	13	6	30.5	11.1	35,720
4234	Professional and commercial equipment and supplies, merchant wholesalers	584	659	790	76	130	13.0	19.8	41,770
5416	Management, scientific, and technical consulting services	358	732	1,137	374	406	104.4	55.4	45,610
…	Federal Government, excluding Postal Service	2,311	1,922	1,972	–389	50	–16.8	2.6	([4])
…	Level-III industries	3,8723	4,010	4,265	137	255	3.5	6.3	([3])
3241	Petroleum and coal products manufacturing	152	119	102	–33	–18	–21.8	–14.8	48,340
3253	Pesticide, fertilizer, and other agricultural chemical manufacturing	54	45	35	–10	–10	–17.7	–21.3	39,680
3255	Paint, coating, and adhesive manufacturing	81	72	62	–8	–11	–10.3	–14.7	35,110
3259	Other chemical product and preparation manufacturing	144	112	79	–32	–33	–21.9	–29.4	35,390
3336	Engine, turbine, and power transmission equipment manufacturing	111	100	100	–11	0	–9.6	.2	37,310
3339	Other general-purpose machinery manufacturing	317	288	339	–29	51	–9.0	17.7	35,320
3353	Electrical equipment manufacturing	219	176	180	–43	4	–19.4	2.2	32,520
3369	Other transportation equipment manufacturing	36	40	40	4	0	10.3	.5	34,230
4861	Pipeline transportation of crude oil	10	7	7	–3	0	–27.0	–2.7	52,020
4862	Pipeline transportation of natural gas	42	29	30	–13	1	–31.0	2.1	49,650
4869	Other pipeline transportation	7	5	5	-2	0	–25.7	–7.7	50,570
5171	Wired telecommunications carriers	637	662	600	25	–62	4.0	–9.4	50,940
5172	Wireless telecommunications carriers (except satellite)	48	196	295	148	99	309.8	50.5	38,480
5173	Telecommunications resellers	173	1856	188	13	2	7.6	1.3	49,400
5174	Satellite telecommunications	19	19	17	0	–2	1.6	–10.4	50,780
5211	Monetary authorities, central bank	24	23	23	–1	0	–2.5	.9	40,840

Note. Reprinted from "High-technology employment: a NAICS-based update" by D. Hecker, 2005, *Monthly Labour Review*, p.60.

Appendix B. Analysis of normality for transformed variables

Table B1.

Analysis of normality of transformed Granted patents variable in high-tech sample.

		Patens Granted	Patents Granted_Sqrt	Patents Granted_Sq	Patents Granted_Ln	Patents Granted_Inv
N	Valid	244	244	244	221	221
	Missing	23	23	23	46	46
Mean		56.50808	5.7436	14,570.91	3.1647	.136
Std. Deviation		106.88578	4.85962	78,973.50	1.52687	.247
Skewness		5.146	1.713	9.753	-.312	2.654
Std. Error of Skewness		.156	.156	.156	.164	.164
Kurtosis		34.754	5.021	107.213	-.395	6.234
Std. Error of Kurtosis		.310	.310	.310	.326	.326

Table B2.

Analysis of normality of transformed Trademarks registered variable in high-tech sample.

		Trademarks Registered	Trademarks Registered_Sqrt	Trademarks Registered_Sq	Trademarks Registered_Ln	Trademarks Registered_Inv
N	Valid	244	244	244	193	193
	Missing	23	23	23	74	74
Mean		11.55562	2.5170	627.3553	1.9406	.2754
Std. Deviation		22.67804	2.28947	3596.66623	1.22001	.30817
Skewness		5.326	1.581	11.697	.138	1.455
Std. Error of Skewness		.156	.156	.156	.175	.175
Kurtosis		39.569	4.461	156.335	-.571	.926
Std. Error of Kurtosis		.310	.310	.310	.348	.348

Table B3.

Analysis of normality of transformed Degree centrality variable in high-tech sample.

		Average Degree centrality	Degree centrality Average SQRT	Degree centrality Average SQ	Degree centrality Average LN	Degree centrality Average INV
N	Valid	244	244	244	242	242
	Missing	23	23	23	25	25
Mean		.0056149	.0650	.0001	-5.9057	2220.725
Std. Deviation		.0058690	.03734	.00014	1.56662	7098.52
Skewness		1.923	.505	4.094	-1.233	5.106
Std. Error of Skewness		.156	.156	.156	.156	.156
Kurtosis		4.357	.143	20.331	1.189	28.389
Std. Error of Kurtosis		.310	.310	.310	.312	.312

Table B4.

Analysis of normality of transformed Betweenness centrality variable in high-tech sample.

		Average Betweenness centrality	Betweenness centrality Average SQRT	Betweenness centrality Average SQ	Betweenness centrality Average LN	Betweenness centrality Average INV
N	Valid	244	244	244	234	234
	Missing	23	23	23	33	33
Mean		.0009745	.0252	0	-7.8623	84617.065
Std. Deviation		.0013795	.01849	.00001	176.796	801057.100
Skewness		2.718	1.068	5.305	-1.318	12.281
Std. Error of Skewness		.156	.156	.156	.159	.159
Kurtosis		8.908	1.159	31.618	3.191	158.964
Std. Error of Kurtosis		.310	.310	.310	.317	.317

Table B5.

Analysis of normality of transformed Granted patents variable in low-tech sample.

		Patens Granted	Patents Granted_Sqrt	Patents Granted_Sq	Patents Granted_Ln	Patents Granted_Inv
N	Valid	265	265	265	77	77
	Missing	7	7	7	195	195
Mean		4.11	.8289	371.1547	1.5588	.4348
Std. Deviation		18.857	1.85447	4326.1501	1.42010	.41259
Skewness		10.738	3.789	15.740	.494	.523
Std. Error of Skewness		.150	.150	.150	.274	.274
Kurtosis		139.705	21.045	252.550	-.642	-1.538
Std. Error of Kurtosis		.298	.298	.298	.541	.541

Table B6.

Analysis of normality of transformed Trademarks registered variable in low-tech sample.

		Trademarks Registered	Trademarks Registered_Sqrt	Trademarks Registered_Sq	Trademarks Registered_Ln	Trademarks Registered_Inv
N	Valid	265	265	265	249	249
	Missing	7	7	7	23	23
Mean		16.61	3.001	2053.377	1.825	.3197
Std. Deviation		42.241	2.768	17380.423	1.344	.33288
Skewness		7.801	2.768	14.687	.547	1.182
Std. Error of Skewness		.150	.150	.150	.154	.154
Kurtosis		81.960	12.289	227.973	-.193	.043
Std. Error of Kurtosis		.298	.298	.298	.307	.307

Table B7.

Analysis of normality of transformed Degree centrality variable in low-tech sample.

		Average Degree centrality	Degree centrality Average SQRT	Degree centrality Average SQ	Degree centrality Average LN	Degree centrality Average INV
N	Valid	263	263	263	256	256
	Missing	9	9	9	16	16
Mean		.0070954	.0681	.0001	-5.9541	2496.38
Std. Deviation		.0090819	.04971	.00032	1.78421	7352.13
Skewness		1.968	.758	4.459	-.724	4.975
Std. Error of Skewness		.150	.150	.150	.152	.152
Kurtosis		4.371	-.086	24.683	-.083	27.346
Std. Error of Kurtosis		.299	.299	.299	.303	.303

Table B8.

Analysis of normality of transformed Betweenness centrality variable in low-tech sample.

		Average Betweenness centrality	Betweenness centrality Average SQRT	Betweenness centrality Average SQ	Betweenness centrality Average LN	Betweenness centrality Average INV
N	Valid	262	262	262	249	249
	Missing	10	10	10	23	23
Mean		.0013194	.0281	0	-7.8133	343985.742
Std. Deviation		.0020777	.02312	.00002	2.07538	4700465.59
Skewness		3.127	1.189	6.180	-1.276	15.689
Std. Error of Skewness		.150	.150	.150	.154	.154
Kurtosis		12.274	1.560	42.667	2.743	247.040
Std. Error of Kurtosis		.300	.300	.300	.307	.307

Appendix C. Assumptions of regression analysis

Low-tech sample assumptions output

Table C1. *Assumption 2: Residuals Statistics output*

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.9077	7.3806	4.0933	1.00256	69
Residual	-4.54799	5.54105	.00000	2.26494	69
Std. Predicted Value	-2.180	3.279	.000	1.000	69
Std. Residual	-1.978	2.410	.000	.985	69

a. Dependent Variable: Innovative performance Startups

Figure C1. *Assumption 4: P-P-plots for normality of residuals Innovative Performance Start-ups*

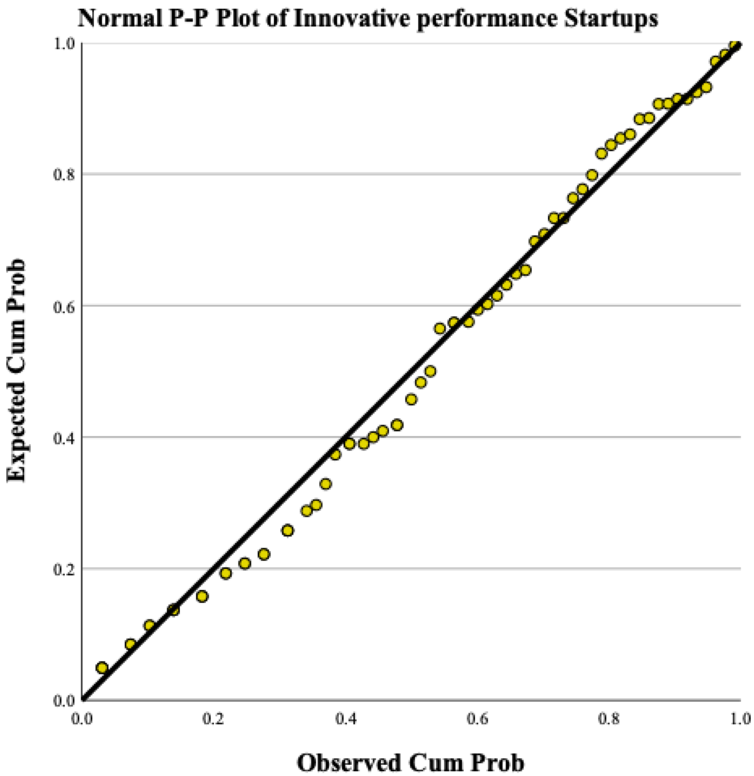


Figure C2. *Assumption 4: P-P-plots for normality of residuals Average Degree centrality*

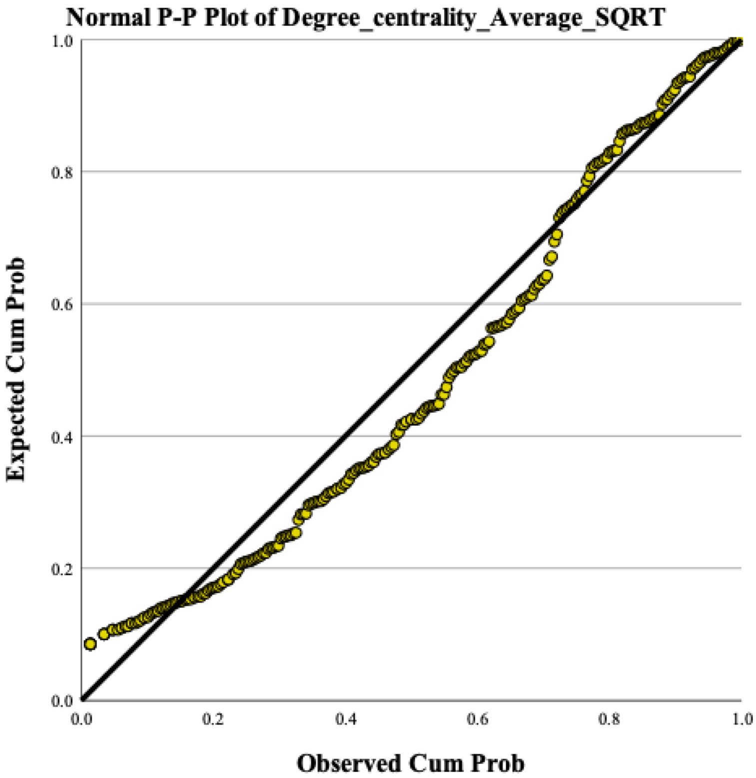
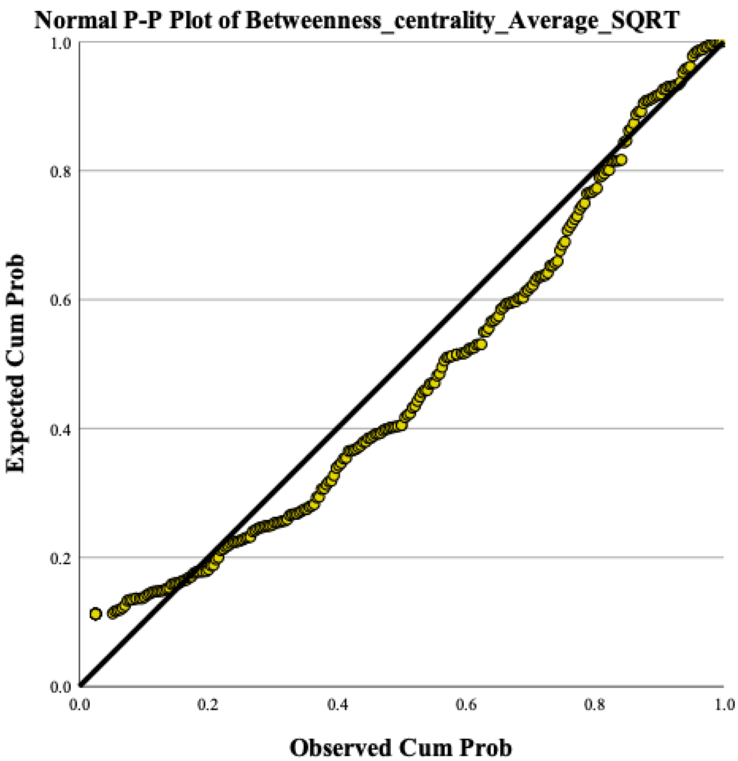


Figure C3. *Assumption 4: P-P-plots for normality of residuals Average Betweenness centrality*



High-tech sample assumptions output

Table C2. *Assumption 2: Residuals Statistics output*

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	5.0645	5.7503	5.3672	.15685	173
Residual	-5.58851	7.19856	.00000	2.17098	173
Std. Predicted Value	-1.930	2.442	.000	1.000	173
Std. Residual	-2.559	3.296	.000	.994	173

a. Dependent Variable: Innovative_Performance_Startup

Figure C4. *Assumption 4: P-P-plot for normality of residuals Innovative Performance Start-ups*

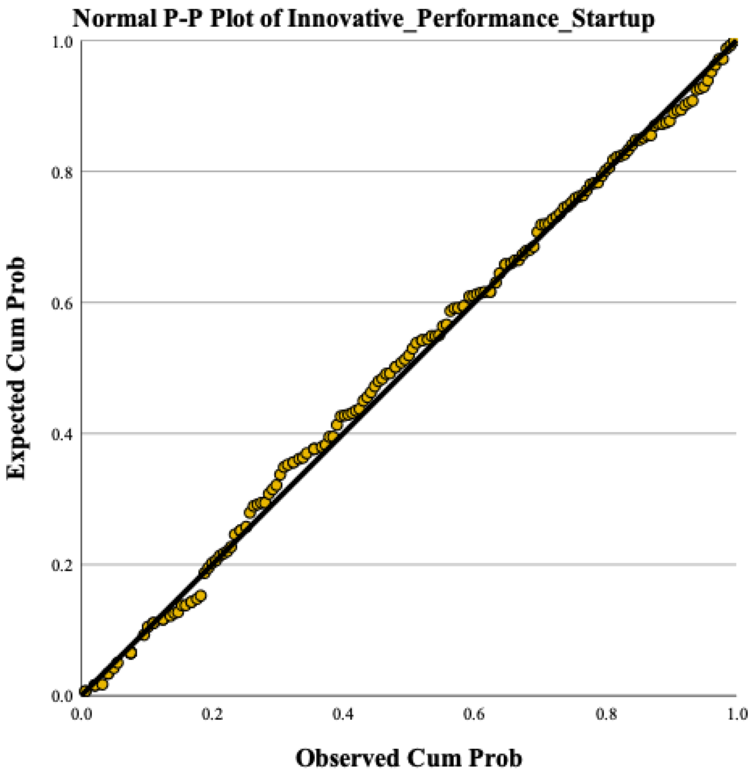


Figure C5. *Assumption 4: P-P-plots for normality of residuals Average Degree centrality*

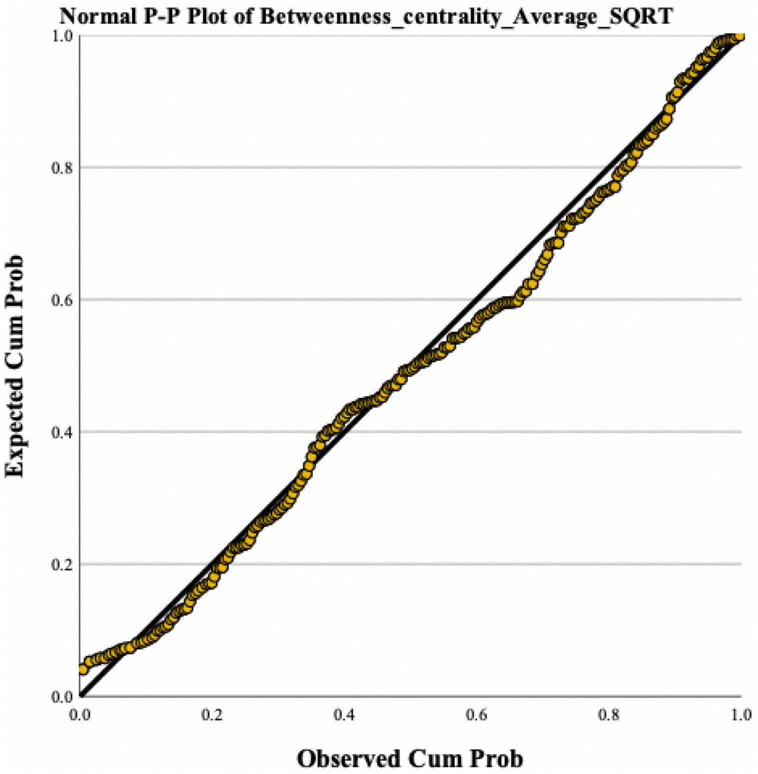
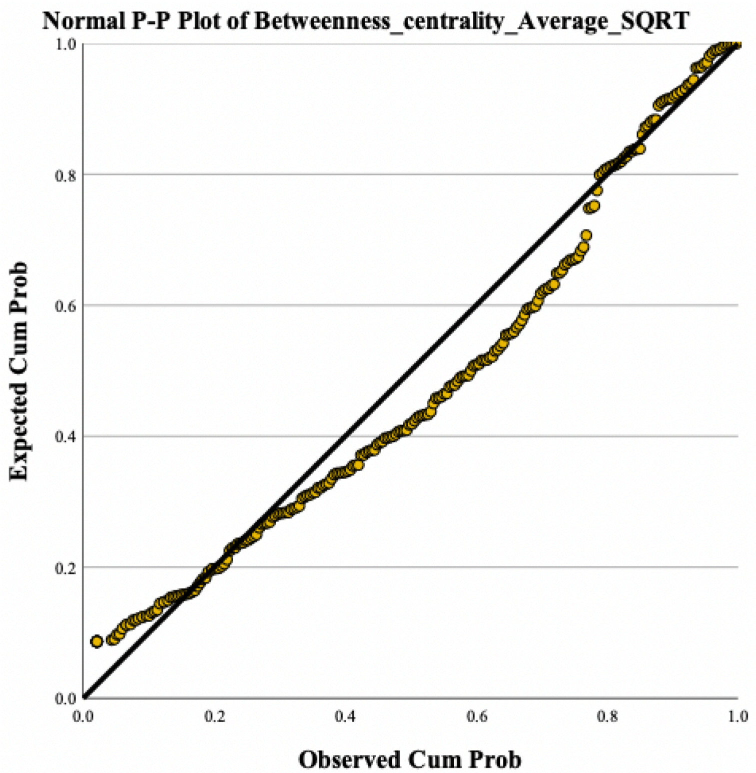


Figure C6. *Assumption 4: P-P-plots for normality of residuals Average Betweenness centrality*



Appendix D. High-tech sample regression analysis relevant SPSS output

Table D1. *Model summary high-tech sample*

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.258[a]	.067	.036	2.12497	.067	2.154	5	151	.062
2	.258[b]	.067	.029	2.13192	.000	.016	1	150	.899
3	.258[c]	.067	.023	2.13903	.000	.005	1	149	.943

a. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees

b. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees, Degree_centrality_Average_SQRT

c. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees, Degree_centrality_Average_SQRT, Betweenness_centrality_Average_SQRT

Table D2. *ANOVA-analysis output high-tech sample*

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	48.625	5	9.725	2.154	.062[b]
	Residual	681.838	151	4.515		
	Total	730.462	156			
2	Regression	48.698	6	8.116	1.786	.106[c]
	Residual	681.764	150	4.545		
	Total	730.462	156			
3	Regression	48.722	7	6.960	1.521	.164[d]
	Residual	681.741	149	4.575		
	Total	730.462	156			

a. Dependent Variable: Innovative performance Startups

b. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees

c. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees, Degree_centrality_Average_SQRT

d. Predictors: (Constant), Number of Lead Investors, Medium number of employees, Age of Company, Number of Investors, Low number of employees, Degree_centrality_Average_SQRT, Betweenness_centrality_Average_SQRT

Table D3. *Regression coefficients of the high-tech sample models*

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations		
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part
1	(Constant)	5.689	.832		6.836	<.001			
	Age of Company	-.003	.012	-.028	-.299	.765	.010	-.024	-.024
	Low number of employees	-1.034	.680	-.221	-1.520	.131	-.222	-.123	-.119
	Medium number of employees	.230	.727	.043	.316	.753	.203	.026	.025
	Number of Lead Investors	.079	.131	.060	.606	.545	.081	.049	.048
	Number of Investors	.030	.043	.072	.703	.483	.058	.057	.055
2	(Constant)	5.656	.874		6.475	<.001			
	Age of Company	-.003	.012	-.026	-.281	.779	.010	-.023	-.022
	Low number of employees	-1.034	.682	-.221	-1.515	.132	-.222	-.123	-.120
	Medium number of employees	.223	.732	.042	.305	.761	.203	.025	.024
	Number of Lead Investors	.080	.131	.061	.610	.543	.081	.050	.048
	Number of Investors	.028	.048	.066	.585	.560	.058	.048	.046
	Degree_centrality_Average_SQRT	.668	5.250	.012	.127	.899	.053	.010	.010
3	(Constant)	5.646	.888		6.355	<.001			
	Age of Company	-.003	.012	-.025	-.265	.791	.010	-.022	-.021
	Low number of employees	-1.035	.685	-.221	-1.511	.133	-.222	-.123	-.120
	Medium number of employees	.221	.735	.042	.300	.765	.203	.025	.024
	Number of Lead Investors	.080	.132	.061	.609	.543	.081	.050	.048
	Number of Investors	.027	.048	.065	.565	.573	.058	.046	.045
	Degree_centrality_Average_SQRT	1.515	12.908	.027	.117	.907	.053	.010	.009
	Betweenness_centrality_Average_SQRT	-1.707	23.729	-.015	-.072	.943	.044	-.006	-.006

a. Dependent Variable: Innovative performance Startups

Table D4. *Correlation matrix of high-tech sample*

Correlations

		Innovative performance Startups	Age of Company	Low number of employees	Medium number of employees	Number of Investors	Number of Lead Investors	Degree_centrality_Average_SQRT	Betweenness_centrality_Average_SQRT
Pearson Correlation	Innovative performance Startups	1.000							
	Age of Company	.010	1.000						
	Low number of employees	-.222	-.309	1.000					
	Medium number of employees	.203	-.008	-.777	1.000				
	Number of Investors	.058	-.265	.235	-.127	1.000			
	Number of Lead Investors	.081	-.189	.118	-.056	.608	1.000		
	Degree_centrality_Average_SQRT	.053	-.247	.073	.041	.493	.271	1.000	
	Betweenness_centrality_Average_SQRT	.044	-.144	.036	.032	.386	.211	.921	1.000
Sig. (1-tailed)	Innovative performance Startups	.							
	Age of Company	.452	.						
	Low number of employees	.003	.000	.					
	Medium number of employees	.005	.460	.000	.				
	Number of Investors	.234	.000	.002	.056	.			
	Number of Lead Investors	.158	.009	.070	.243	.000	.		
	Degree_centrality_Average_SQRT	.256	.001	.182	.307	.000	.000	.	
	Betweenness_centrality_Average_SQRT	.291	.036	.328	.346	.000	.004	.000	.
N	Innovative performance Startups	157							
	Age of Company	157	157						
	Low number of employees	157	157	157					
	Medium number of employees	157	157	157	157				
	Number of Investors	157	157	157	157	157			
	Number of Lead Investors	157	157	157	157	157	157		
	Degree_centrality_Average_SQRT	157	157	157	157	157	157	157	
	Betweenness_centrality_Average_SQRT	157	157	157	157	157	157	157	157

Appendix E. Low-tech sample regression analysis relevant SPSS output

Table E1. *Model summary low-tech sample*

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.526[a]	.277	.211	2.24541	.277	4.205	5	55	.003
2	.566[b]	.320	.244	2.19695	.043	3.453	1	54	.069
3	.604[c]	.365	.281	2.14248	.045	3.781	1	53	.057

a. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees

b. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees, Degree_centrality_Average_SQRT

c. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees, Degree_centrality_Average_SQRT, Betweenness_centrality_Average_SQRT

Table E2. *ANOVA-analysis output low-tech sample*

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	106.011	5	21.202	4.205	.003[b]
	Residual	277.303	55	5.042		
	Total	383.313	60			
2	Regression	122.678	6	20.446	4.236	.001[c]
	Residual	260.635	54	4.827		
	Total	383.313	60			
3	Regression	140.032	7	20.005	4.358	<.001[d]
	Residual	243.281	53	4.590		
	Total	383.313	60			

a. Dependent Variable: Innovative performance Startups

b. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees

c. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees, Degree_centrality_Average_SQRT

d. Predictors: (Constant), Number of Lead Investors, Low number of employees, Age of Company, Number of Investors, Medium number of employees, Degree_centrality_Average_SQRT, Betweenness_centrality_Average_SQRT

Table E3. *Regression coefficients of the low-tech sample models*

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part
1	(Constant)	6.348	1.000		6.350	<.001			
	Age of Company	-.008	.012	-.090	-.646	.521	.087	-.087	-.074
	Low number of employees	-2.676	.786	-.517	-3.405	.001	-.285	-.417	-.391
	Medium number of employees	-1.954	.835	-.355	-2.338	.023	-.013	-.301	-.268
	Number of Investors	.132	.045	.430	2.939	.005	.180	.368	.337
	Number of Lead Investors	-.565	.177	-.488	-3.181	.002	-.112	-.394	-.365
2	(Constant)	5.197	1.158		4.489	<.001			
	Age of Company	-.002	.012	-.021	-.149	.882	.087	-.020	-.017
	Low number of employees	-2.508	.774	-.485	-3.240	.002	-.285	-.403	-.364
	Medium number of employees	-1.766	.824	-.321	-2.144	.037	-.013	-.280	-.241
	Number of Investors	.050	.062	.162	.796	.429	.180	.108	.089
	Number of Lead Investors	-.514	.176	-.445	-2.924	.005	-.112	-.370	-.328
	Degree_centrality_Average_SQRT	17.549	9.444	.342	1.858	.069	.294	.245	.209
3	(Constant)	4.644	1.164		3.989	<.001			
	Age of Company	.002	.012	.026	.187	.852	.087	.026	.020
	Low number of employees	-1.923	.813	-.372	-2.367	.022	-.285	-.309	-.259
	Medium number of employees	-1.298	.838	-.236	-1.548	.128	-.013	-.208	-.169
	Number of Investors	.038	.061	.124	.620	.538	.180	.085	.068
	Number of Lead Investors	-.462	.173	-.400	-2.666	.010	-.112	-.344	-.292
	Degree_centrality_Average_SQRT	48.219	18.265	.940	2.640	.011	.294	.341	.289
	Betweenness_centrality_Average_SQRT	-77.643	39.932	-.613	-1.944	.057	.145	-.258	-.213

a. Dependent Variable: Innovative performance Startups

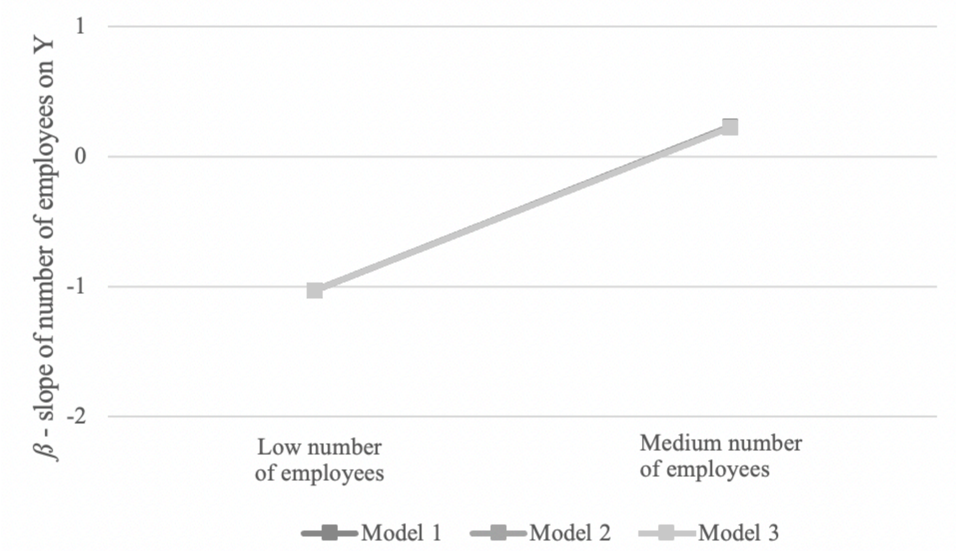
Table E4. *Correlation matrix low-tech sample*

Correlations

		Innovative performance Startups	Age of Company	Low number of employees	Medium number of employees	Number of Investors	Number of Lead Investors	Degree_centrality_Average_SQRT	Betweenness_centrality_Average_SQRT
Pearson Correlation	Innovative performance Startups	1.000							
	Age of Company	.087	1.000						
	Low number of employees	-.285	-.222	1.000					
	Medium number of employees	-.013	-.188	-.503	1.000				
	Number of Investors	.180	-.312	-.027	-.001	1.000			
	Number of Lead Investors	-.112	-.267	-.092	-.135	.599	1.000		
	Degree_centrality_Average_SQRT	.294	-.372	-.010	.002	.773	.417	1.000	
	Betweenness_centrality_Average_SQRT	.145	-.353	.083	.024	.706	.386	.927	1.000
Sig. (1-tailed)	Innovative performance Startups	.							
	Age of Company	.251	.						
	Low number of employees	.013	.043	.					
	Medium number of employees	.461	.074	.000	.				
	Number of Investors	.082	.007	.418	.496	.			
	Number of Lead Investors	.196	.019	.241	.151	.000	.		
	Degree_centrality_Average_SQRT	.011	.002	.471	.493	.000	.000	.	
	Betweenness_centrality_Average_SQRT	.132	.003	.262	.427	.000	.001	.000	.
N	Innovative performance Startups	61							
	Age of Company	61	61						
	Low number of employees	61	61	61					
	Medium number of employees	61	61	61	61				
	Number of Investors	61	61	61	61	61			
	Number of Lead Investors	61	61	61	61	61	61		
	Degree_centrality_Average_SQRT	61	61	61	61	61	61	61	
	Betweenness_centrality_Average_SQRT	61	61	61	61	61	61	61	61

Appendix F. Post-hoc effects of dummified control variable Number of Employee

Figure F1. *Effects of number of employees on Innovative Performance of Start-ups in the high-tech sample*



Note. The scores lie close to each other in such a degree that it is visually presented as a single line (e.g., .230; .223; .221 respectively for model 1, 2 and 3). In reality, there are three lines in this figure.

Figure F2. *Effects of number of employees on Innovative Performance of Start-ups in the low-tech sample*

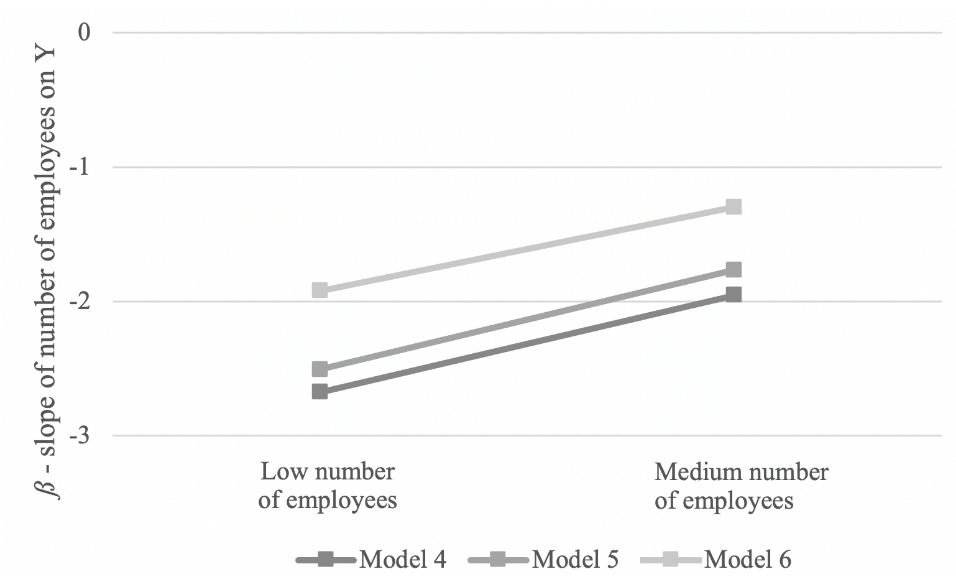


Figure F3. *Total post-hoc effects of number of employees on Innovative Performance of Start-ups in the high-tech sample*

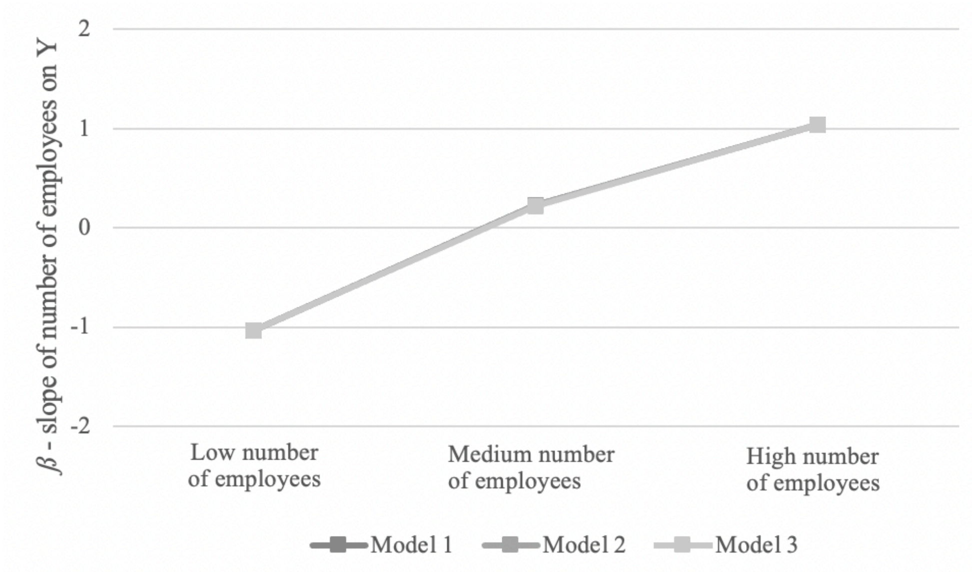


Figure F4. *Total post-hoc effects of number of employees on Innovative Performance of Start-ups in the low-tech sample*