Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

Vegard Jensen Løwe
Fredrik Moger
Harald Tryti Rieber

Measuring and Understanding Venture Capital Fund Performance in the Nordics

Master's thesis in NTNU School of Entrepreneurship
Supervisor: Roger Sørheim
June 2022

NTNU
Kunnskap for en bedre verden

Vegard Jensen Løwe
Fredrik Moger
Harald Tryti Rieber

Measuring and Understanding Venture Capital Fund Performance in the Nordics

Master's thesis in NTNU School of Entrepreneurship
Supervisor: Roger Sørheim
June 2022

Norwegian University of Science and Technology
Faculty of Economics and Management
Dept. of Industrial Economics and Technology Management

NTNU
Norwegian University of
Science and Technology

Preface

This thesis has been written by three students as the final part of their master's degree at the School of Entrepreneurship (NSE) at the Norwegian University of Science and Technology (NTNU). The study has been conducted from January to June 2022.

The authors would like to express their sincere gratitude to their supervisor Roger Sørheim for his encouragement and insightful discussions during the writing of the thesis. Additionally, the authors would like to thank Argentum and Menon for sharing their data. The data has been invaluable for pushing the field of research on Nordic venture capital further.

Trondheim, June 11th, 2022

Vegard Løwe	Fredrik Moger	Harald T. Rieber
Vegard Jensen Løwe	Fredrik Moger	Harald Tryti Rieber

Abstract

This thesis investigates how to measure and understand venture capital (VC) fund performance in the Nordics. Data concerning direct performance such as internal rates of return (IRR) is rarely public, and thus it becomes of interest to find an accurate proxy for performance.

Most existing research on VC performance proxies has used low-quality databases with primarily data on US VCs, like VentureXpert or Venture Economics. The number of new papers conducting research on VC performance proxies is declining; only 18 articles found through our literature study are more recent than 2010. Research concerning Europe and in particular the Nordics is very sparse.

To further the VC research in the Nordics, we create a unique dataset consisting of 141 VC firms and 421 VC funds gathered from Argentum, Menon (NVCA), and significantly expanded by the authors. Using this dataset, we perform a k-means clustering analysis to group VC firms with similar performance using our own proxy. We then inspect the structural characteristics of the clusters to find what separates the top-, middle- and bottom performers.

Our findings expand, contradict and confirm previous research and in certain places contradict the common perceptions of the industry. For example, we find Denmark to be lagging, and find Norway and Sweden to be the most mature investment ecosystems. We also find the maturity and performance of the ecosystem to have significantly improved in recent years and find that funds are becoming more differentiated both in terms of strategy and performance. Later stage firms seem to be both the highest performing and the rarest. Funds specialized in industry have the highest risk and highest return potential. We find that the primary difference between the average funds and the outperforming funds are the experience and amount of funds raised by the firm.

Because a proxy might be valid only for a given ecosystem and time-period, we recommend researchers to inspect the applicability of previous research in a Nordic context. Furthermore, we conclude that a much more solid database is needed to further the field of study, as well as specific research laying the groundwork for unified performance measurements.

Sammendrag

Denne masteroppgaven undersøker hvordan vi kan måle og forstå prestasjonen til risikokapitalfond (VC) i Norden. Data for direkteavkastning som internrente blir sjeldent offentliggjort, og derfor er det ønskelig å finne en god proxy for prestasjon.

Mesteparten av eksisterende forskning på proxier for prestasjon i VC-sektoren har brukt amerikanske databaser med lav datakvalitet som VentureXpert og Venture Economics. Antallet forskningsartikler som omhandler proxier for risikokapital er nedadgående, og vi finner kun 18 artikler som i litteratursøket som er nyere enn 2010. Spesielt forskning i Europa og Norden er veldig begrenset.

For å videreføre forskningen i Norden lager vi et unikt datasett med 141 VC fondsforvaltere og 421 VC fond, mottatt av Argentum, Menon (NVCA). Datasettet er betydelig utvidet av forfatterne. Vi bruker k-means clustering til å gruppere fondene ut ifra deres prestasjon med vår egen proxy. Etter dette bruker vi strukturelle karakteristikker til å kartlegge hva som skiller topp-, midt- og bunnfondene.

Funnene våre utvider, motstrider og bekrefter tidligere forskning, og går mot de vanlige oppfatningene av industrien flere steder. For eksempel finner vi at Danmark henger etter, og at Norge og Sverige har de mest modne investeringsøkosystemene. Videre finner vi at modenheten og prestasjonen til det Nordiske økosystemet har forbedret seg betydelig i senere år, og at fond i senere år har blitt mer og mer ulike i både strategi og prestasjon. Det er færrest fond i senere fase, men prestasjonen deres er bedre enn fond i andre faser. Fond som er spesialisert i industri har både høyest risiko og høyest prestasjonspotensiale. Vi ser at de største forskjellene mellom et gjennomsnittlig fond og et ekstraordinært fond er antallet fond opprettet av fondsforvalteren, og dermed fondets erfaring.

Fordi en proxy muligens kun gjelder for et gitt økosystem i en gitt tidsperiode anbefaler vi forskere å undersøke om tidligere forskning er brukbar i en nordisk kontekst. I tillegg konkluderer vi med at en mye bedre database trengs for å kunne videreføre forskningsfeltet, og at det må lages en felles standard for direkte rapportering av avkastning hvis en god proxy for prestasjon skal kunne lages.

Table of Contents

Figures

Tables

Equations

1 Introduction

This thesis investigates how to measure and understand venture capital (VC) fund performance in the Nordics. The research is conducted by studying a unique dataset consisting of 141 VC firms and 421 VC funds gathered by Argentum, Menon, and the authors. To measure and understand VC fund performance in the Nordics, the thesis outlines the current state of the Nordic VC industry. Following, we propose a new proxy for measuring VC performance as the two most commonly used direct measures - IRR and TVPI - is unobservable on a fund-level due to secrecy. The proxy is constructed by considering the ability to raise the following fund swiftly and increase the fund size. Subsequently, we use the proxy to investigate the structural characteristics of low-, mid- and top-performers to increase the understanding of VC fund performance in the Nordics.

Considering that seven out of the eight most valuable companies in the world today were initially backed by VC firms (Wittenstein, 2022) it is fair to state that VC stimulates the growth and renewal of the global economy. Academics and practitioners have effectively articulated the strengths of the VC model. These include its strong emphasis on governance, capital financing and their network. Indeed, Arrow et al. (1995) once opined that "venture capital has done much more I think, to improve efficiency than anything". In many respects, the VC industry appears to be a bright spot in the increasingly troubled global innovation landscape (Bloom et al. 2020). However, not everything concerning VC is shining.

"*Venture capital is an interesting industry in which at least 75 % of the players you talk to are top quartile performers...*" Leleux (2007) repeats after interviewing a leading US VC. This tongue-in-cheek reference points not only to many VCs' tendency for self-promotion, but also to a more fundamental issue the industry has struggled with since its origination, namely measuring and understanding the drivers of performance (Arundale, 2018). The issue has proven a very difficult one to tackle both by the academic and the professional communities alike, due to a unique combination of factors such as:

- The fundamental challenge of valuing VC investments (restricted securities of early-stage, technology-rich companies) (Woodward & Hall, 2004 Leleux, 2007; Arundale, 2018).
- The very private nature of the industry, where most of the reported numbers are aggregations of self-claimed rates of returns (Muzyka et al., 1996; Bruton & Ahlstrom, 2003; Parhankangas, 2007; Arundale, 2018).

Despite the issues, previous research has been able to explain major differences in performance by studying structural, operational, and wider environmental using aggregated data. E.g., Hege et al. (2003) demonstrated that the magnitude of the difference in performance was related to the time periods measured. Later, Kaplan and Schoar (2005) showed that VCs who outperform the industry in one fund are also likely to outperform in their next fund, implying that structural or operational determinants could explain differences in performance. In the wake of Kaplan and Schoar (2005), several studies sought to explain the influence of a single determinant, or a subset of determinants mainly by applying regression (e.g., Dimov & Shepherd, 2005; Walske & Zacharakis, 2009; Lerner et al., 2011). From these studies the field of research became aware of the related challenges of applying statistical methods to study the influence of factors, as the underlying data quality was remarkably poor.

Arundale (2018) was the first scholar to apply a more holistic approach, gathering and restructuring the majority of determinants proposed by previous research, both structural, operational, and wider environmental determinants. By conducting an extensive first-hand data gathering, Arundale (2018) confirmed findings by Lerner et al. (2011) which suggested that medium sized funds, focusing on the growth stage, with entrepreneurial experience in the management team and with a location in a technology hub was consistently outperforming its peers. Additionally, Arundale (2018) extended the field of research by finding that VC firms with more than one partner working on a deal, and firms that were engaged in a theme-based approach for spotting future investments, tended to outperform the competition.

While several previous studies have pointed to the influence of structural, operational, and wider environmental determinants on VC performance, it is not without contrasting views. Most of the previous research has applied statistical methods on highly criticized datasets. From the literature review, the authors found that of the 81 articles reviewed, 29 either used VentureXpert, Venture Economics or Venture Source. Kaplan and Lerner (2017) reviewed the specific datasets and found them to be severely lacking. E.g., VentureXpert report that less than 10 % of investments end up failing, whereas the actual amount is closer to 20 % (Keplan & Lerner, 2017). Further the data is often only available in aggregate format for each firm, rather than on a fund-by-fund basis (Ljungqvist & Richardson, 2003). Lerner et al. (2011) and Arundale (2018) add that the data is provided largely on a voluntary basis and that the data is based on unrealized as well as realized investments with the former involving subjective valuations.

The contrasting views illuminates key gaps in the literature. First, the critique of the datasets applied in previous research demands a need for research investigating what datasets and research methods are used to demonstrate differences in performance. Furthermore, the consistent characterization of the VC industry's tendency to self-promote skewed results points to the need of investigating how the academic and the professional communities in a best possible way can measure performance, given the shortcomings of previous literature.

Additionally, despite being world leaders in terms of innovation capacity, high-technological development and investments, highly skilled labor and competent and efficient public institutions, the role of venture capital in the Nordics is unexplored. Existing literature is mainly oriented around the US venture capital industry and great efforts have been made to understand the role of venture capital in this context. While measuring and understanding VC performance in the Nordics is a research gap itself, it also underlines the gap in understanding the generalizability of previous research. I.e., we currently don't know if findings in previous comparative research is valid for other geographical regions. Could we understand differences in performance amongst the Nordic peers by considering research concerning other regions?

As a response to the proposed research gaps, the purpose of this study is to *measure and understand venture capital fund performance in the Nordics.* To achieve the proposed purpose, the following research questions (RQ) has been outlined:

- RQ1: What is the current state of the Nordic VC industry?
- RQ2: What proxy can be used to measure VC fund performance?
- RQ3: How does structural determinants influence VC fund performance in the Nordics?

The RQs are explored in a Nordic context. Moreover, the level of analysis is at the fund-level perspective (meso) as shown in Figure 1. To obtain an answer to the outlined RQs, the authors conducted an extensive literature review on the specific field of research during the fall of 2021 to gain an understanding of the previous work on measuring and understanding venture capital fund performance.

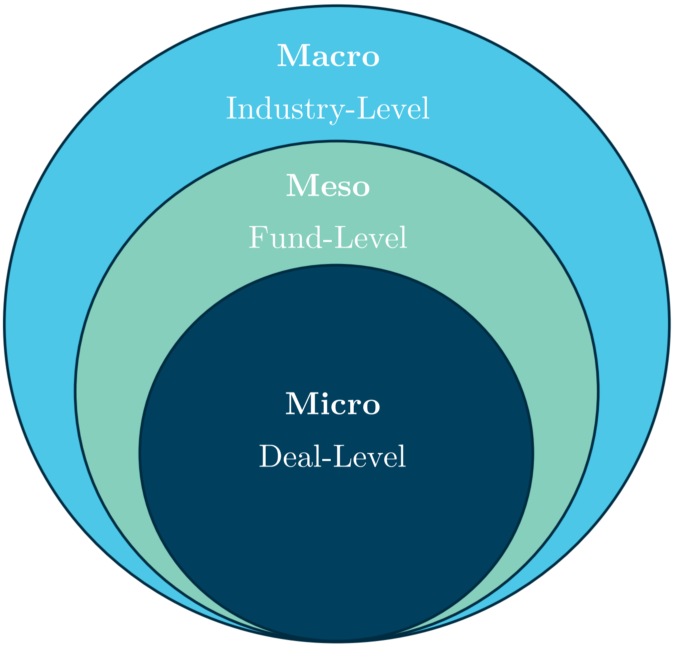


Figure 1: The level of analysis is at the fund-level perspective (meso).

This thesis is to the best of our knowledge, the first to investigate the factors affecting VC fund performance in the Nordics. By studying a unique dataset, this thesis finds that the Nordic VC industry is at the end of a boom-and-bust cycle. A low-interest environment in combination with strong fiscal policies related to the pandemic drove the Nordic market till new heights. Now the economic landscape is looking challenging as interest rates are increasing and the economic activity is declining, indicating a possible stagnation. Further we find that the proposed proxy for performance allows for best practice benchmarking of structural factors. This can yield new insights regarding strategic positioning for players in the Nordic VC scene. To amplify the previous statement, by considering the results of the proxy, we find that later-stage VC investments are in general underrepresented in the Nordics, and that Denmark is lagging behind its Nordic peers. While that later-stage investments are underrepresented extends existing literature, our findings that Denmark is lagging contradicts existing research proposing Denmark together with Sweden are the most mature VC markets in the Nordics.

By answering the RQs, the authors contribute to the literature field of venture capital by first illuminating the current state of the Nordic VC industry. Secondly, this thesis contributes by developing a proxy for measuring VC fund performance based on the time needed for a VC firm to raise the following fund, and the ability to increase the following fund size. Lastly, this thesis contributes by providing a better understanding of how structural determinants influence VC fund performance in the Nordics. The thesis extends previous literature considering VC performance in the rest of the world, in a Nordic context, in addition to present contrasting results on performance amongst the Nordic peers. The findings provide the VC field with quantitative empirical data and evidence for further analysis and research. In addition to answering the RQs, the authors have during their work with this thesis expanded the current state-of-the-art dataset on Nordic VC funds with a twofold, increasing both the coverage and quality of the underlying data. This contribution strengthens the empirical foundation and the room of opportunity for future research.

The study is divided into eight chapters. **Chapter 2** presents the relevant theory on venture capital fund performance. The research method is presented in **chapter 3**; presenting the research design and applied method of the thesis, followed by the limitations of the chosen method. **Chapter 4** includes the qualitative findings. In **chapter 5** the authors answer the RQs by discussing the findings and existing literature, followed by the limitations of the thesis. Further, in **chapter 6**, the authors present their conclusion. Lastly, the authors present their recommendation for further research and evaluate their own work in respectively **chapter 7 and 8**.

2 Frame of Reference

This chapter provides an overview of the relevant literature as the study's frame of reference. The presented literature is derived from a literature review conducted by the authors during the fall of 2021. First, the context of VC is given. We then move forward to present the context of the Nordic VC industry. Further, an elaboration of defining VC fund performance is put forward. Finally, we present the frame of reference on structural determinants' influence on VC fund performance.

2.1 The Context of VC

VC is medium to long term finance that is invested by professional firms in potentially high growth unquoted companies in return for equity stakes in those companies (Arundale, 2007; Lerner et al. 2011). VC is a subsegment within the broader private equity sector, which also includes equity finance for considerably later stage established businesses. The private equity sector is again a subsegment under the investment management umbrella, which is often referred to the handling of financial assets and other assets – not only buying and selling them (Cassis & Minoglou, 2005). VC is characterized by its ability to bring funding to new risky growth companies, often with a potential to disrupt the market, leading growth companies to succeed through strategic and operational support, and by identifying high-growth opportunities at the right time in the right markets (Gompers & Lerner, 2001).

In its simplicity, VC firms has **five** main tasks (Metrick & Yasuda, 2021). **As a first step**, when raising a new fund, a VC firm needs to find investors, often referred to as limited partners (LPs). VC is risky and therefore the reputation and past performance of VC firms are important. **Secondly**, over the next 2-4 years, the VC firms screen a large number of companies to identify investment cases. To source the best deals, VC firms often chase down particular companies or specific investment theses, generating outbound deal flow. Outbound deal flow is often thought of a measure of how hard the VC firm is working. Inbound deal flow, by contrast, indicates how much entrepreneurs value a VC as an investor and refers to the number of deals that come directly to the VC firm. This process results in a portfolio of 10-15 companies. **Thirdly**, VC funds carry out active ownership in the portfolio companies, using their highly specialized knowledge, network, and syndication with other VC firms to increase the chance of success. **The fourth task** starts when the portfolio has matured. The VC fund will start looking for potential buyers in other types of equity markets. This process aims at realizing the value of the investments, often through an initial public offering or a strategic exit. The realized potential and experience are often canalized into a new fund, leading to **the fifth and final task**, distribution of the funds. The five tasks are illustrated in Figure 2 below.

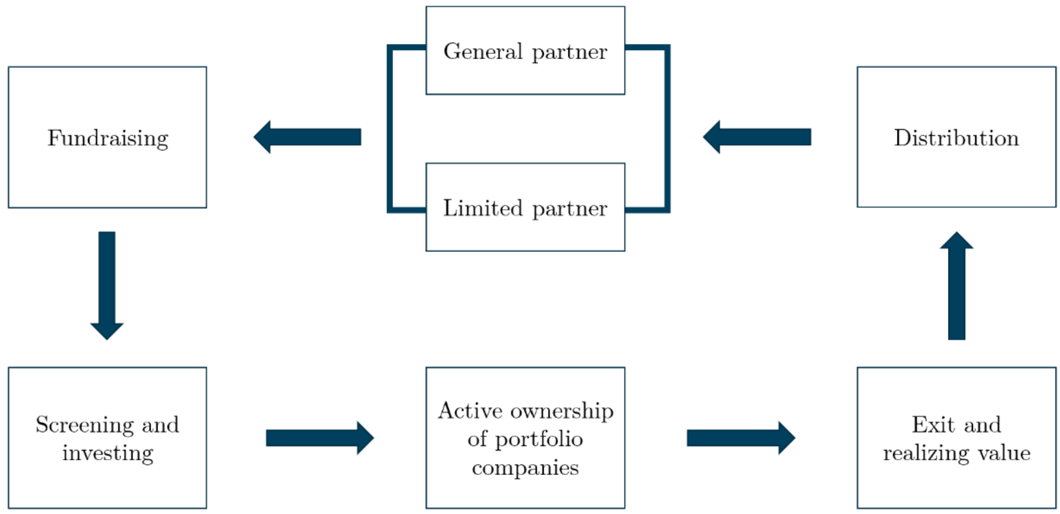


Figure 2: The five tasks of VC firms.

VC plays a crucial role in the capital food chain, bringing up small, innovative startups into proven business concepts (Metrick & Yasuda, 2021). The capital ecosystem can be divided into four financing cycles: (1) seed, (2) venture, (3) growth and (4) buyout (Parhankangas, 2007). Just like it sounds, seed funding describes the first capital sourced at the beginning of a startup's journey. Private investors provide this type of funding, such as angel investors, government venture or seed funds. Often, seed investors take on significant risk without proof of revenue from the business. Seed funding is often used to support the foundational business needs, such as developing prototypes, market research, and research and development costs. Once a business has launched and is beginning to gain traction, it often will require venture capital. Unlike seed capital, venture capital investments often provide entrepreneurs with larger raises (above $ 1 M)

(Schwarzkopf, 2015). Most entrepreneurs seeking venture capital have begun generating revenue and are focused on proving their business model. Growth capital is a form of private equity that in most cases, is a minority investment in established firms seeking funding to expand or restructure operations, enter new markets, or finance major acquisitions. These firms are more likely to be bigger (over ~50 employees) and more established than venture capital-funded businesses, capable of generating a steady recurring revenue but unable to finance significant expansions, acquisitions, or other investments towards the business's growth (Canadian Business Growth Fund, 2022). Access to growth equity is often critical in helping these businesses continue to scale, pursue essential facility expansion, sales and marketing efforts, equipment purchases, and new product development. Buyout funds are pools of capital to be invested in companies that represent an opportunity for a high rate of return. They come with a fixed investment horizon, typically ranging from four to seven years (Corporate Finance Institute, 2022). Exit strategies include IPOs and sale of the business to another private equity firm or strategic buyer. Contrary to seed and venture, buyout funds invest in more mature businesses, usually taking a controlling interest (Corporate Finance Institute, 2022). Buyout funds tend to be significantly larger in size than seed or venture funds.

Looking away from the pure capital itself, an important part of VCs contribution is the societal capital from mentorship and knowledge (Metrick & Yasuda, 2021). There are especially two factors that enable VC firms to give their portfolio companies indispensable guidance in bringing them to success. First, previous experience in the field and in startups. Investment professionals and staff working at VC firms are often previous successful entrepreneurs. Shaw & Sørensen (2019) finds that this is crucial in transforming a good idea into a commercial success, suggesting that previous successful entrepreneurs have 67 % higher sales compared to entrepreneurs without previous experience. Additionally, the majority of VC firms offer specialist knowledge, e.g., in life science with links to academia or computer science with understanding of state-of-the-art machine learning models. This enables VC firms to provide concrete feedback on a product level. According to e.g., Sapienza (1992), Large and Muegge (2008) and Quas et al. (2021) VC firms support their portfolio with finding the right strategy from the start, minimizing product risk and bringing the product to the market, network and bringing the right talent, getting access to other sources of finance, governance and compliance, and choosing the best exit strategy.

2.2 The Context of Nordic VC

Florida and Kenney (1988) indicate that there are differences between VC firms located in different geographic regions. Fried and Hisrich (1995) developed this idea further, arguing that there could be differences between regions based upon regional characteristics (geography, history, and economy), industrial culture (risk-taking versus conservative) and organization and management orientation.

Literature on the Nordic VC industry is limited and few research papers exists. This has induced a theoretical cap in the literature and an inadequate understanding of the Nordic VC industry. However, this theoretical gap can somewhat be filled by analyzing descriptive statistics on the development of the Nordic VC industry and utilizing research conducted on other VC industries in Europe, UK, and US.

2.2.1 Differences within the Nordic

Cetindamar (2003) argues there are three dimensions which are suitable for not only measuring the evolution of the VC industry but also its present maturity. These dimensions are the size of the industry, the diversity of the industry and the competence in the industry. The main measure of size is the total cumulative funds raised for VC investments. A VC industry with enough capital in order to fulfill its role as a supporting industry is perceived as mature. The VC industry should further provide adequate services to firms in a very broad range of industries and technologies. These services should also be available for all stages in the evolution of firms, from seed to buyout. The VC industry appears not to be subjected to international competition in large extent, therefore, the VC industry is largely national in nature and dependent on the national competence (Cetindamar, 2003). This means that it may exist a large and diversified VC industry which is inefficient due to an inadequate competence. Competence, in terms of experience, in specific industries/technologies is required to reduce risks and to mobilize networks of investors. Because of yearly variation in VC flows, our analysis of the maturity is based on the cumulative values. The idea is to measure the cumulative size, diversity, competence, and hence the overall maturity of the industry.

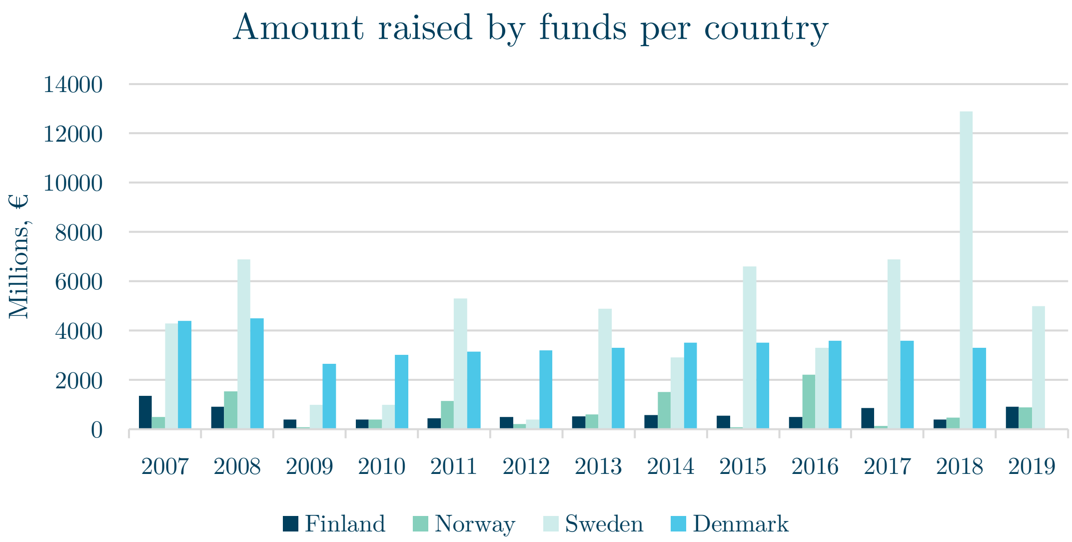


Figure 3: Fundraising split by home country of VC firm. Source: NVCA, SVCA, FVCA, Aktive Ejere

If we consider the amount of capital raised by VC funds over a period from 2007-2019, we observe that Sweden and Denmark have evidently a larger VC market than Finland and Norway. Although it is difficult to say how much VC capital is optimal, it is clear that VC industries with more capital, such as Sweden and Denmark, have greater potential to fulfill its role as a supporting industry (Cetindamar, 2003). Vækstfonden (2019) claims relative to GDP the Danish VCs are the most active in the EU, and by considering absolute numbers, the Danish VCs are only surpassed by the UK, France and Germany. However, Copenhagen Economics (2019) reports Finland as the most active VC market in Europe relative to GDP and Denmark as the ninth most active VC market. Similarly, Argentum (2020) finds that the Swedish VC market has considerably more investments than the Nordic peers, with Finland as the second most active market, and Denmark and Norway as the least active VC markets in the Nordics.

Strong academic and research communities in the Nordics lead to most investments towards ICT and life science in Europe (Copenhagen Economics, 2019). Both Denmark and Norway are registered with a strong investment focus towards respectively life science and energy (Copenhagen Economics, 2019), which makes them less diverse in terms of industry specialization than its peers. According to Copenhagen Economics (2019), both the Swedish and Norwegian VC industry is equally oriented towards early and later stages, whereas the Finnish and Danish VC industry is strongly oriented towards early stage. In terms of diversity in the VC industry, the statistics imply Sweden as the most diverse VC industry in the Nordics.

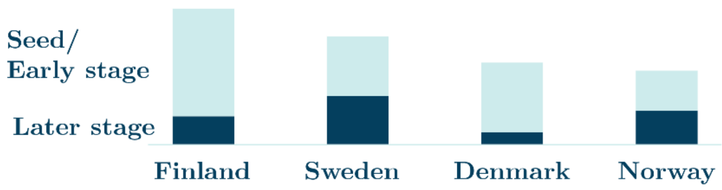


Figure 4: VC investments allocation between stages, 2016. Source: OECD

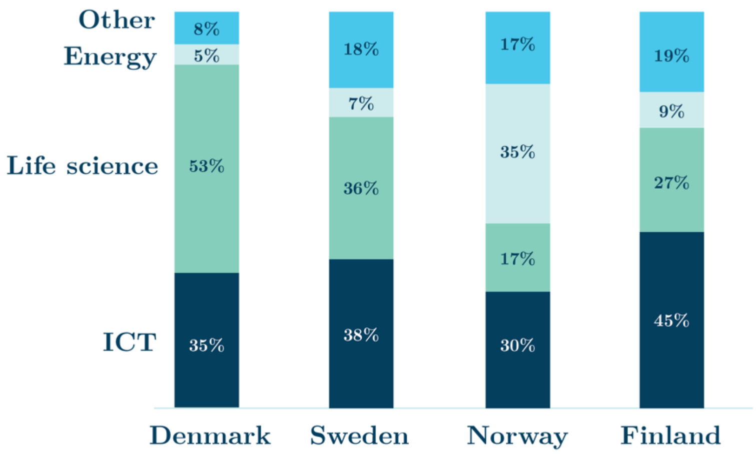


Figure 5: Industry specialization 2007-2017. Source: Invest Europe

The Nordic VC industry was born significantly after that of the US, during the late 1970's and early 1980's. The first VC fund in the Nordics was Företagskapital established in Sweden in 1973. The public sector was also collaborating on the development of the industry, and many of the earliest VCs established in the Nordics were semi-private, as stated by Hyytinen & Pajarinen (2001). In the 1980s the VC industry began to grow as several new private VC firms were founded. By the mid-1980s, there were about 20 VC firms in Denmark, 5-6 in Norway and some 20 private VC firms in Sweden, accompanied by around 30 regional and government run investment companies (Christensen, 2000; Karaömerlioglu & Jacobsson, 2000). In Finland, the growth lagged a bit compared to the other Nordic countries. However, by 1988 there were 48 VC firms in Finland (Seppä, 2000). In 2000, there were registered high activity in the Nordic VC industry: the amount of funds raised was € 852 M in Denmark, € 570 M in Finland, € 497 M in Norway and € 3.6 Bn in Sweden (Hyytinen & Pajarinen, 2001). During the period from 2000 until today, it exists a trend of more VCs operating and make investments in Sweden and Finland than in Norway and Denmark (Argentum, 2020). This development can reasonably draw lines to Sweden and Finland possesses the highest levels of competence and experience in the Nordics.

2.2.2 The Nordic versus EU and US

The Nordic countries are among the most developed economies in the world and a well-educated and skilled population find it natural to focus on new innovative companies seeking to grow on international markets. According to the European Commission's Innovation Union Scoreboard 2021, Denmark, Finland and Sweden are three of the top performing countries. Additionally, the Nordic business sectors score high in different rankings of innovation activity. In the 2018 Global Entrepreneurship Index ranking, the Nordic countries all rank in the top 25 globally. Denmark is ranked as the sixth most entrepreneurial economy globally, Sweden as number 9, Finland as number 12 and Norway as number 21. Figure 6 presents the average ranking among the Nordic countries on selected parameters according to World Economic Forum Global Competitiveness Report 2018.

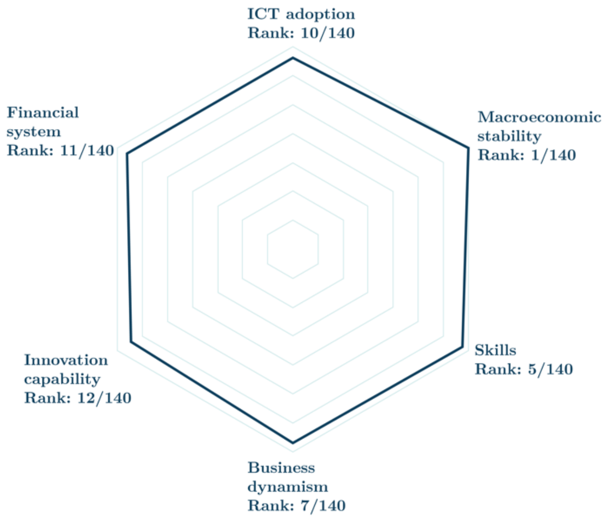


Figure 6: Average ranking Denmark, Norway, Sweden, and Finland on selected parameters, out of 140 countries. Source: World Economic Forum Global Competitiveness Report 2018

Despite the strong base of innovation potential in the Nordics, the Nordic VC market is perceived as less mature than in US and the rest of continental Europe. While the venture capital industry emerged in the US, it is today a global phenomenon that exhibits many regional variations (Wright et al. 2004). However, comparative research on the Nordic VC industry versus the US or Europe is quite limited. Luckily, the financial systems in the Nordic countries are comparable with its European peers. The financial structure is different in the US, where the securities markets are larger and the banking sector is smaller (Gjedrem, 2000). This gives us somewhat understanding of the Nordic VC industry by relating with the European VC industry. A more discussed topic within previous research is comparisons of the European VC industry versus US. Several cross-regional comparisons study the differences between the European and US VC market (Lerner et al., 2011; Kräussl & Krause, 2014; Revest & Sapio, 2012). The results of these studies indicate the main differences include: (1) US funds are larger than European funds which allows them to make twice as many investments, (2) US funds tend to invest with a larger number of co-investors and (3) US funds have a larger focus on internet and communications compared with European funds. Invest Europe (2021) reports in their annual report, a total number of 288 European venture capital funds raised € 15 Bn in 2020, while by comparison fundraising by Nordic venture funds in 2020 was reported to € 906

M (Argentum, 2021). Figure 7 presents the fundraising per region for the three-year period 2018 – 2020.

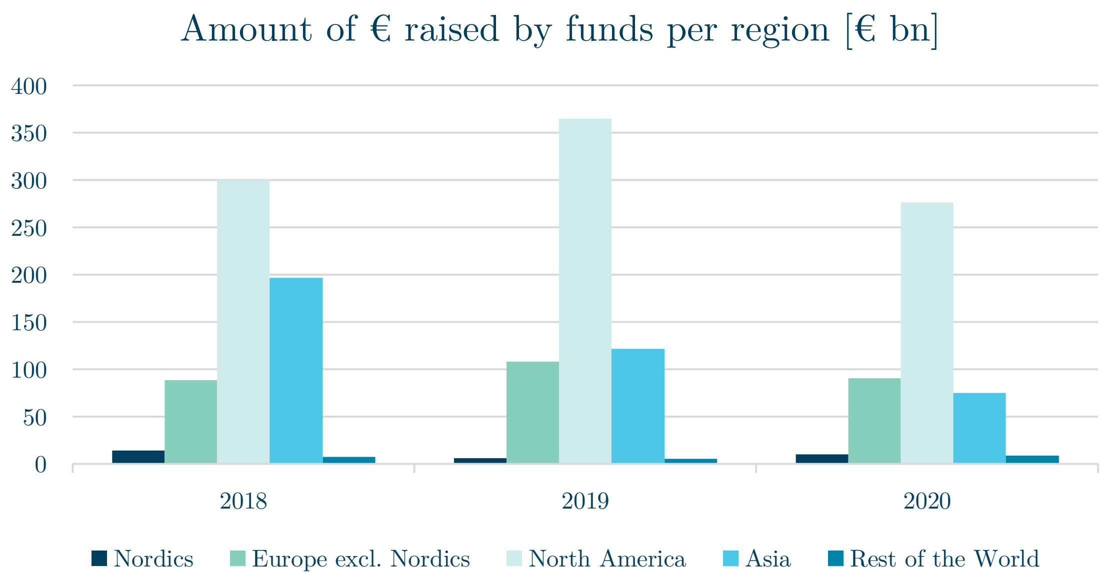


Figure 7: Amount raised per region 2018-2020. Source: Invest Europe, McKinsey (2021)

Copenhagen Economics (2019) reports that ICT and life science industries play a larger role in the Nordics than in the rest of the EU and at level with the US. In general, countries with large ICT and life science sectors have strong VC markets – with US being the important showcase, where VC is dominant in financing these two sectors. Despite strong academic and skilled population, the Nordics lack large enough funds with the financial muscle and experience to commercialize and expand internationally (Copenhagen Economics, 2019). Consequently, later-stage VC investments are in general underrepresented in the Nordics and the average investment size is lower in the Nordics compared to European peers. Data from OECD and Invest Europe is presented in Figure 8 and Figure 9.

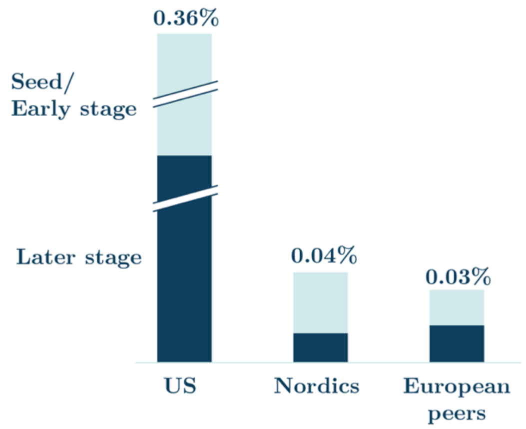


Figure 8: VC investments as a percentage of GDP, 2016. Source: OECD

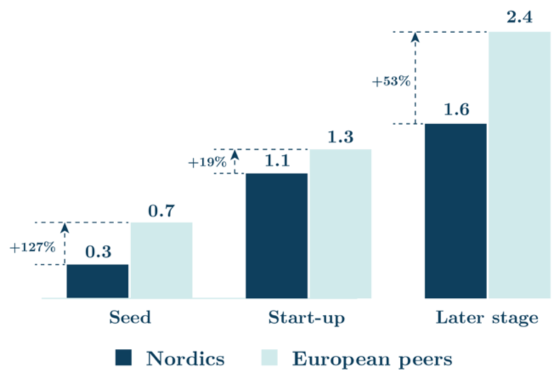


Figure 9: Average investments size in 2015-2017, €M. Source: Invest Europe

2.3 Defining Performance

From a limited partner perspective, the most important measurement is the financial returns from VC fund investments. A long-term lack of competitive returns will force limited partners to avoid VC investments, or only invest in funds with proven track records (McKenzie & Janeway, 2011; Nanda et al. 2020). However, the shortage of reliable industry data has led to a shaky empirical foundation for the field of research (Lerner & Kaplan, 2016). An unappealing consequence is that dubious or misleading studies can linger for many years without rebuttal. In this chapter, we present the frame of reference in regard to defining venture value creation and VC fund performance. Furthermore, we present a comprehensive overview of the datasets and research methods applied in previous research. Lastly, we present proxies proposed to measure VC fund performance.

2.3.1 Entrepreneurs, Venture Value and VCs

The related VC literature considers multiple aspects of performance, mainly dependent on the observer. From an entrepreneurial perspective, performance could be measured in terms of their ability to add value, in addition to capital infusions (e.g., Sapienza, 1992; Rosenstein et al. 1993; Barney et al., 1996). From an inside-out VC firm perspective, performance in terms financial returns is what drive income through management fees, but performance could also refer to the ability to raise the next fund or increase the size of the following fund (Gompers & Lerner, 1998; Nanda et al. 2020).

In this section we present a reworked and expanded model with its origin from the model proposed by Pandher (2021). In order to establish a frame of reference we consider an entrepreneur (E) with a venture of initial value V_0 searching for external capital c in the venture finance market with VCs (VC firm) with differing levels of value-adding capabilities β_V in the range $\beta_V \in [0, \bar{\beta}]$ (Bygrave & Timmons, 2009). The entrepreneur does not know the productivity type of VCs but knows their distribution represented by the probability measure $P(\beta) = P(\beta_V \le \beta)$ (Maula & Murray, 2001).

The basic timing of actions in the interaction between E and VC is as follows over the contracting period $[0, T]$ (where T is in years). At the start of the period E approaches a VC of productivity $\beta_V \in [0, \bar{\beta}]$ for funding (De Clercq et al., 2006). The $VC(\beta_V)$ offers E a term sheet. Next the entrepreneur E decides whether to accept or reject the term sheet. If the deal is accepted, both E and $VC(\beta_V)$ choose their effort levels over the development period $[0, T]$ (Maula & Murray, 2001). At time T, the cumulative impact of factors decides whether the venture will be successful (state "S") or whether it will fail (state "F").

The venture by time T generates profits $Y_T(S)$ and leads to valuation $V(S)$ at exit (e.g., IPO, acquisition) if successful (Kaplan & Schoar, 2005). The profits of the venture over the contracting period $[0, T]$ should depend on the size of the total investment $Y_0 + c$ and efforts a_E and a_v provided by the entrepreneur and VC firm, respectively. They are also subject to uncertainty, and we model them as

Equation 1: Venture value if successful.

$$Y_T(S) = (V_0 + c)(1 + \lambda + \beta_E a_E + \beta_V a_V + \varepsilon), \qquad \varepsilon \sim N(0, T\sigma_Y^2)$$

where β_E represents the productivity of the entrepreneur's efforts a_E; β_V is the productivity of the VC firms' effort A_V (Maula & Murray, 2001; Smith, 2001); λ is the contribution to growth from other factors autonomous of the VC firm and entrepreneur; ε is a normally distributed shock to venture profits with (annualized) variance σ_Y^2; and T is the length of the venture development period (in years) (Leleux, 2007).

If the venture is successful, its exit value $V(S)$ will sell at a random multiple of its expected profits (earnings). If we define $E(Y) = E(Y_T)/T$ as the venture firm's annualized expected earnings at time T and let θ be the valuation multiple of earnings reflecting the venture firm's growth prospects. Then the uncertain value of the successful venture at exit may be represented as

Equation 2: Uncertain value of successful venture exit.

$$V(S) = (\theta + \eta)EY, \qquad \eta \sim N(T\sigma_Y^2)$$

where η is the shock to the venture's valuation parameter at exit and reflects the uncertainty in the venture's value (around the expected value $\theta E(Y)$). In financial valuations, ratios of firm value to various performance measures (e.g., earnings, revenues, enterprise value (EV)) are frequently used to estimate their value and θ may be interpreted accordingly as a valuation multiplier. E.g., if Y represents the venture earnings, then, θ may be viewed as the venture's forward price-to-earnings ratio. It reflects how information on the venture's earnings $E(Y)$ is converted into asset value through the relation $V(S) = (\theta + \eta)EY$ where the valuation multiple $\theta + \eta$ is stochastic with expected value $E(\theta + \eta) = \theta$. Similarly, if Y measures firm revenues, then θ may be viewed as the price-to-revenue ratio for the venture.

If the venture fails (F) at the time T, it is reasonable to assume that the efforts of the entrepreneur and the VC were ineffective and added no incremental value to the investment $(V_0 + c)$. Some of these investments may, however, be recovered though the liquidation of the venture's assets and we define ϕ as the recovery rate of this initial investment (Metrick & Yasuda, 2010). Hence, in case of venture failure at time T, the venture is liquidated and produces the cashflow

Equation 3: Venture value if failed.

$$Y_T(F) = \phi(V_0 + c).$$

Since there is no IPO or acquisition in this state, the venture's exit value is zero: $V(F) = 0$.

For VC funds to measure performance, the aggregated value of the portfolio, e.g., $V(P) = \sum_{mn} V(S)_m + Y_{nT}(F)$, needs to be considered. This is often referred to as TVPI, defined as $\frac{V(P)}{c}$, where $V(P)$ is total distributions to LPs ($V(S)_m$, also known as IRR when periodized and divided on c), in addition to an estimated residual value of unrealized returns ($Y_{nT}(F)$). TVPI and IRR is known as the two most common performance metrics when VCs quantify their own financial performance (see next section). When studying the underlying elements, differences in performance is strongly connected to the efforts a_E and a_v provided by the entrepreneur and the VC, respectively (e.g., Maula & Murray, 2001; Smith, 2001). Furthermore, studies considering the influence of structural factors seek to understand the influence of λ - the contribution to growth from other factors autonomous of VC and E - e.g., US in comparison to Europe (Arundale, 2018) or the influence of year of establishment (Hege et al., 2002).

2.3.2 Direct Measures for VC Fund Performance

When VC firms quantify their own financial performance, either IRR, or TVPI is typically used (Diller & Kaserer, 2004). When LPs measure their investments financial performance, a public market equivalent (PME) is often employed to understand the relative performance (Diller & Kaserer, 2004). All these measures have their shortcomings, however, Ljungqvist and Richardson (2003) show that the IRR of the average fund does not turn positive until the eight years of the fund's life, often referred to as the "J-curve effect", which means it is only at the very end of a fund's life that excess returns are realized. Additionally, external valuations of portfolio companies only exist in the event of IPO's, trade sales based on tradable securities or cash, additional financing rounds including third parties, or if the company files for bankruptcy. Therefore, according to Ljungsqvist and Richardson (2003), the calculations of interim IRRs computed before a fund reaches maturity are not very informative.

Because interim IRRs are a suboptimal measure of performance, TVPI is a relatively common measure of performance for unfinished funds. TVPI is defined as total value / paid-in capital, where total value is total distributions to LPs in addition to an estimated residual value of unrealized returns. Accurately estimating the net asset value of unrealized returns ($V(S) = (\theta + \eta)EY$) is hard, thus TVPI is not a perfect measure of VC performance.

Woodward and Hall (2004) argue that reported returns from VC firms are too low in a rising market but too high in a falling market. Cumming and Waltz (2010) show that experienced VC firms tend to report significantly lower valuation than their younger, especially early-stage and high technology-focused, counterparts. When evaluating and comparing IRRs, there has been shown unclear and inconsistent use of the net and gross returns i.e., whether the reported results include or exclude fees to the VC firms (Døskeland & Strömberg, 2018). Cumming and Waltz (2010) show that there are systematic differences in accounting standards, and thus differing reports of both TVPI and IRR. In addition, funds with different investment horizons are shown to realize their investments at quite different points, so comparing long-term and short-term funds is hard. Kaplan et al. (2005) argues that VC performance should be measured against the public market equivalent (PME), to quantify if limited partners benefit from allocating their capital in private investments instead of the public market. PME constitutes the market performance for the same period as the investment's duration.

Most of the existing studies have used regression analysis on existing datasets to obtain a measure for performance for a specific determinant. Because these existing datasets typically don't contain data on IRR, but do contain exit-rates, this has been a common proxy for performance (Kaplan & Lerner, 2017). To isolate a specific determinant, different authors have taken different approaches. Many attempt to heighten the quality of the existing datasets by appending relevant data from other sources. Other exclude different parts of the original datasets to analyze a smaller subset of the data. Another common approach is to create a mathematical model which attempts to isolate only the effect of a specific determinant on performance. It is difficult to verify which approach to isolating determinants is the most accurate. However, most existing research on the same determinant using the same dataset but using different methods for isolating a determinant and measure of performance has come to similar conclusions regarding the determinants impact. This might indicate that previous studies have successfully managed to isolate determinants and obtain a reasonable proxy for performance.

2.3.3 Databases and Research Methods

Having an accurate database that is representable for the entire population of VC firms in a given ecosystem is necessary to be able to make a good proxy for performance. Most of the early research into VC fund performance relied on information available in IPO prospectus. For the subset of venture-backed firms that eventually go public, a vast amount of information is available. Investments in firms that do not go public are more difficult to uncover, since these investments are usually not publicized (Lerner & Kepler, 2016). Unfortunately, because only a relatively modest fraction of venture-backed companies goes public, researchers must dig deeper.

As a response to this shortcoming, databases such as VentureXpert (VX) and Venture Economics (VE) started to evolve. There are however large inconsistencies in these databases and a general problem of incompleteness (Avdeitchikova, 2012; Kraemer-Eis et al., 2012; European Commision, 2010; Mason & Harrison, 2013; Prohorovs, 2014).

Furthermore, qualitatively, both VX and VE show deterioration in data quality over the past decade. In more recent time, European initiatives on jointly owned and operated data cooperatives by private equity and venture capital associations in Europe has served as a new source of data (Invest Europe, 2021). However, this data is often hard to access for both the professional and academic community (NVCA, 2022). The shortcomings highlighted by existing literature has led to a shaky empirical foundation for statistical analysis. In order to overcome the shortcomings, the more recent research has tended to be more explorative towards alternative qualitative research methods. In the following sections we present the coverage, strengths, and shortcomings of VX and VE, in addition to present what databases and research methods the existing literature has applied to understand differences in performance.

VentureXpert (VX)

VX is a venture capital database provided by Thomson Economics. VX began collecting data in 1961. VX provides data on deals, VC funds and LPs. The coverage is approximately 7,000 funds, and 23,000 portfolio companies. The coverage is thought of being more complete than similar databases (Kaplan & Lerner, 2016).

Gornall and Strebulaev (2015) recorded many missing and miscoded VC data in VX. Kaplan and Lerner (2016) commented “there are large inconsistencies in VentureXpert and VentureSource (…) qualitatively, both show deterioration in data quality over the past decade (…) coverage has dropped dramatically in recent years. Further, Kaplan and Lerner (2016) reports that VX understates the fraction of companies that are defunct. From their study we learn that VX reports that less than 10 % is defunct when, in fact, more than 20 % are defunct. Röhm et al. (2020) pointed out that the data problems in VX caused by the different definitions of “corporate venture capital” were prevalent.

Venture Economics (VE)

Similar to VX, VE has traditionally sourced its data from both LPs and VC firms. As of 2008, VE reported approximately 320 VC funds in its database, which implies a 50 % coverage of the capital commitments in the total VC landscape in the US (Harris et al., 2013).

The major issue with VE was that it appeared to stop updating performance on roughly 40 % of the venture capital and private equity funds in the VE sample. Stucke (2011) finds that between 1980 and 2005, 43 % of the funds had constant net asset values and no cashflow activities for at least two years prior to December 2009. Phalippou and Gottschalg (2009) finds that 300 of 852 sample funds are inactive for more than 3 years, with most for 6+ years. Harris et al. (2014) finds that VC fund performance is lower in VE than that in its peers. Lerner and Kepler (2016) suggest it is likely that for reasons related to poor quality data that is describe above, VE decided to discontinue its database (Lerner & Kepler, 2016).

The previous studies of VC performance have by and large been quantitative in nature. Of the 59 articles reviewed in this study that directly discussing the impact of one or several performance determinants, 54 are quantitative and 5 are qualitative. In addition, nearly all use regression analysis to determine the influence on performance. Of the articles not using regression (or similar), the means of analysis differ quite drastically. Arundale (2018) argues each determinant must be viewed through its own theoretical lens and employs a comprehensive multi-theoretical framework to analyze individual determinants. Bygrave (1987) uses graph theory to analyze the strength of each VC firms' network, to determine the correlation between network and performance. Amit et al. (1998) attempt to create a new theoretical framework which focuses on the asymmetry of information between VC's and entrepreneurs. Kim and Park (2021) use fuzzy-set qualitative comparative analysis to ascertain whether single conditions had to be in place to reach the outcome. This analysis seeks to determine whether the VC fund could have reached the same profitability and the same deal flow with a significantly different strategy.

The fact that the underlying data for most of the studies on VC fund performance is severely inaccurate, limits the contribution of these studies. Lerner and Kaplan (2016) argue that this, in part, may reflect the challenges associated with the reliance on commercial data providers, who may decide on an investment in ensuring data quality that while profit-maximizing, is less than an academic financial economist would prefer.

Furthermore, the lack of availability of the datasets used in studies where data was collected first-hand limits the possibility of reproducing the results of these studies. Appendix A contains an overview of the datasets and research methods used by the 59 articles reviewed in this thesis. We learn that 29 use either VX or VE, and that 54 of the articles apply statistical analysis to understand the influence of determinants on VC fund performance.

2.3.4 Proxies for VC Fund Performance

Hochberg et al. (2007) use exit-rates as a proxy for performance. Exit-rates are defined as the percentage of companies exited through either an initial public offering (IPO) or through mergers and acquisitions (M&A). Gompers et al. (2010) use a similar measure, where success is equivalent to performance, and success is measured by a portfolio company undergoing or registering for an IPO. Smith et al. (2010) studied how financial returns measured as internal rate of return (IRR) or total value to paid-in capital (TVPI) correlates with exit-rates and found the correlation to be strong. The paper thus concludes that exit-rates provide a suboptimal measure for performance. The reasoning is that many studies have relied on using exit-rates instead of other measures for performance mostly because of the unavailability of other data that could act as a more fitting performance measure.

Kaplan and Schoar (2005) finds that fund flows are positively related to past performance, indicating that top performing funds have an easier time raising more capital in the following fund. Additionally, Kaplan and Schoar (2005) document substantial persistence in VC fund performance, meaning that firms who outperform the industry in one fund are likely to outperform the industry in the next and vice versa. This was not only true for two consecutive funds, but also true for the current fund and the second previous fund, indicating that the numbers of funds raised by a firm could imply the level of performance (Kaplan & Schoar, 2005). Kaplan and Schoar is supported by Gompers and Lerner (1999) who finds that older and larger VC firms tend to perform well. Although this might be true, Kaplan and Schoar finds that on average, top performing funds grew proportionally slower than the lower performing funds in the sample period, indicating that relative growth should be adjusted for the sample population's mean fund size.

Morris et al. (2020) recently proposed a little discussed proxy based on markups, i.e., when one VC firm invest after another firm at a higher price. More markups mean more funding, and so they are used as a proxy for successful ventures, and at an aggregated portfolio-scale for VC firms. According to Morris et al. (2020) the industry is arranged around this notion. The fact that Cambridge Venture Capital Index separate the "better" VC firms from other by quartile based on their markups in addition to hard numbers such as IRR supports their finding. Another characteristic of markups, strengthening its popularity is VCs desire to showcase how their portfolio gains value, so markups are used as an intermediate measure of success on the road to realized returns. Morris et al. (2020) concludes that markups are not an accurate proxy for the value of VC investments as it is mainly a phenomenon driven by access to capital, not startup business merit.

2.4 Structural Determinants

According to the literature, the determinants that influence VC performance fall into three main groups: (1) Structural determinants, (2) Operational determinants and (3) Wider environmental determinants. The first group of determinants are the most proven and include specific structural aspects of VC funds, including specialization in geographical focus, industry or investment stage, age of VC and VC fund size.

2.4.1 Specialization in Geographical Focus, Industry, and Investment Stage

When establishing a VC fund, the managers must decide on their general investment focus. VCs will select their investment stage (broadly seed, start-up, early stage, and later stage), industry focus (the industry sectors in which they want to invest), and geographic focus (local, regional, country wide, multi-country, global). All these decisions may affect a VC fund's performance.

Most VC firms are specialized in one or more industries. On average, the funds invest close to 40 % of their capital in a single industry (Ljungqvist & Richardson, 2003). Specialization in one industry enables the firms to develop their industry-specific capabilities and experience. This is of utmost importance in e.g., high-tech, and medical industries. The chaotic uncertainty and opacity of an emerging technology market may be too high a barrier for other than the most specialized investors (Lockett et al., 2002). In contrast, Aigner et al. (2008) find no advantage of specialization for a firm. Similarly, Walske and Zacharakis (2009) find, based on interviews with US VCs, that specialization in geography (US states) have no impact on fund performance.

Risk reduction is important for VC firms to obtain a persistent performance. The managers can affect their risk exposure by choosing an appropriate set of industries (which also have their own level of risk) or development stages of ventures (Buchner & Schwienbacher, 2017). For example, expected risk exposure is likely higher when allocating a larger fraction of funds to early-stage ventures. VCs may specialize in certain stages; seed, early and later venture stages, and growth and scale-up (EVCA, 2010). Early-stage investments, such as investments in seed or start-up companies, are associated with higher levels of downside risk and higher upside potential, whereas later-stage investments are associated with lower downside risk and moderate upside potential (Buchner & Schwienbacher, 2017). Nanda et al. (2020) explains how successful VCs have a preferential access to deals, both entrepreneurs and other VC firms want to partner with them. Successful VC firms therefore get to see more deals, particularly in later stages, when it becomes easier to predict which companies might have successful outcomes, while unexperienced VC firms may not have the opportunity to invest in later stages.

Diversification is usually seen as a way to reduce and manage risks. Prior research by Humphery-Jenner (2012) documents a positive relationship between industry and geographical diversification and performance for US VC's. In contrast, Cressy et al. (2014) studied 649 UK VCs to find that industry diversification reduces fund performance, but geographical diversification improves performance. This corresponds with several researchers that encourage VC firms to specialize in industry to reduce risk (Murray & Marriott, 1998; Wang & Ang, 2004).

2.4.2 Age of VC Firm

It's reasonable to assume that the more experience a VC company has in making and managing investments, the better its performance will be. The age of the VC company is measured as the difference between founding date and investment date (Espenlaub et al., 2014). The older the firm, the more contacts, experience, and prominence it has (Jääskeläinen et al., 2006). Jääskeläinen et al. (2006) explain moreover, the younger the firm, the more it tries to establish a reputation by opportunistically striving toward successful exits. This is a phenomenon called "grandstanding" (Gompers, 1996). A higher risk tolerance to strive toward successful exits could lead to higher returns. In contrast with Jääskeläinen et al. (2006), Aigner et al. (2008) interestingly find that experienced firms tend to take more risk in the sense that they have more underperforming or defaulting portfolio companies.

Liu and Zhiqi (2014) argue that a young VC firm has an initial learning period during which it may be more prone to mistakes and consequently its performance may not necessarily increase with age. In the longer term, however, we expect that older VC firms have gained more experience and become better at selecting promising portfolio companies to invest in and providing the latter with better services, both of which should lead to superior investment performance (Liu & Zhiqi, 2014; Gompers et al., 2008).

Buchner et al. (2018) find an effect that cross-border deals are most often carried out by older funds. Perhaps the fund age reflects a tendency to invest initially in domestic deals that are perceived as less risky, and to wait until the fund have matured before 'gambling' on possibly riskier cross-border deals. As a result, after previous domestic deals show evidence of success, VC firms might be confident that they can mitigate the risk of cross-border deals by leveraging their domestic accomplishments. Funds that take a higher risk approach to investing in cross-border deals could lead to increased returns.

2.4.3 Fund Size

Prior research has shown that the fund size influences the fund's performance. Walske and Zacharakis (2009) defines fund size as the dollar amount of capital raised in the fund. Funds with assets under management below $84 million can be categorized as small funds, $84-$364 million as medium-sized and above $364 million as large funds (Lerner et al., 2011).

SVB Capital (2010) presents multiple certain explanations, based on historical returns data from the data provider Preqin, on advantages of small funds over large funds. Smaller funds tend to have more attractively structured partnership terms that create a stronger alignment between management and investors. Managers of smaller funds typically develop a tighter focus on a specific niche or strategy that gives them a competitive edge over the other investors. Focused managers can leverage their sector and geographic-specific networks for high-quality deal flow. However, SVB Capital (2010) describes smaller fund sizes, and hence fewer investments, force managers to focus on capital efficiency. Maula et al. (2006) recommend that fund should have a sufficient fund size to ensure the possibility of follow-on investments to avoid excessive dilution in later rounds and to diversify the portfolio. Larger organizations also have greater influence on their environments, enabling them to engage in "environmental manipulations" (Walske & Zacharakis, 2009).

Large funds allow the management to perform more investments, follow-up their most successful investments and spread the costs more widely. Lerner et al. (2011) argue that the size of the fund is positively related to the experience of the firm itself, which subsequently impact VC fund performance. Some large funds have multiple strategies with partners and operate like multiple small funds combined (SVB Capital, 2010). For example, a $500 million fund that targets three different industry sectors may have specific allocations to those sectors that do not change over the life of the fund.

According to the findings of the research, there may be an ideal, medium-sized fund for improved fund performance. Lerner et al. (2011) finds that medium-sized funds outperform large funds. Kaplan and Schoar (2005) confirm this by suggesting a positive concave relation between fund size and performance. The average IRRs for funds with assets under management of less than $84 million were around 7% lower than for bigger funds, while funds with assets under management of more than $365 million did not outperform small funds with assets under management of less than $84 million.

2.5 Summary of Frame of Reference

The presented literature in this frame of reference brings the relevant literature for this study. Venture capital plays a crucial role in the capital food chain, bringing up small, innovative startups into proven business concepts (Metrick & Yasuda, 2021). When studying a specific region, such as the Nordics, it is important to acknowledge the differences in previous literature's area of interest and this study. The Nordic VC industry is historically limited considered in previous research and the reported data from national venture capital associations are interpreted as contradictory. However, the Nordic VC industry is perceived as smaller and less mature than the EU and US. This will introduce considerations in how comparative studies shall be conducted and how VCs' performance is measured.

Lerner & Kaplan (2016) finds that researchers lack reliable industry data of the VC industry, which has resulted in a shaky empirical foundation. In this study, we present a comprehensive overview of datasets and research methods applied in previous research and discussing the shortages involved in previous research. To overcome the shortages, we propose a proxy for measuring VC performance in a specific geographical region (Nordics). According to previous literature, it exists several determinants that influence VC performance. Structural determinants are most proven to influence VC performance, including specialization in geographical focus, industry and investment stage, age of VC firm, and fund size. The research methodology in Chapter 3 will further apply the findings of the frame of reference to design a relevant methodology for analyzing the data.

3 Research Methodology

In this chapter, the applied research methodology of the study is outlined. First, the method and research design are presented. Secondly, the data acquisition and preparation are covered. The third section outlines structuring of the data, before the last section presents the applied statistical method.

3.1 Research Process and Research Design

Scientific research is a method to produce coherent answers to problem statements. The scientific approach allows the researchers to find general rules, test and discuss findings, and collect objective evidence (Mitchell & Jolley, 2010). Research design is used to provide a systematic presentation of data so that it is possible to discuss and draw conclusion within a hypothesis. A research design starts with deciding which methods to use for data collection and analysis.

3.1.1 Research Process

The authors have divided the research process into four steps: (1) Literature review and research design, (2) data collection, (3) data analysis, and (4) public presentation. This study began with a literature review which led to the outlined RQs and the chosen research design. Next, the authors started the data collection process by cleaning data from Argentum and NVCA, in addition to expanding the dataset through additional data gathering efforts. After the data was collected, the authors used statistical methods to cluster subgroups of funds by their performance. Next, to answer the RQs, the authors discuss descriptive statistics for the clusters in comparison with the findings from previous literature. The process and outcome of the authors' steps during the study is illustrated in Figure 10 and further described in the following sub-chapters.

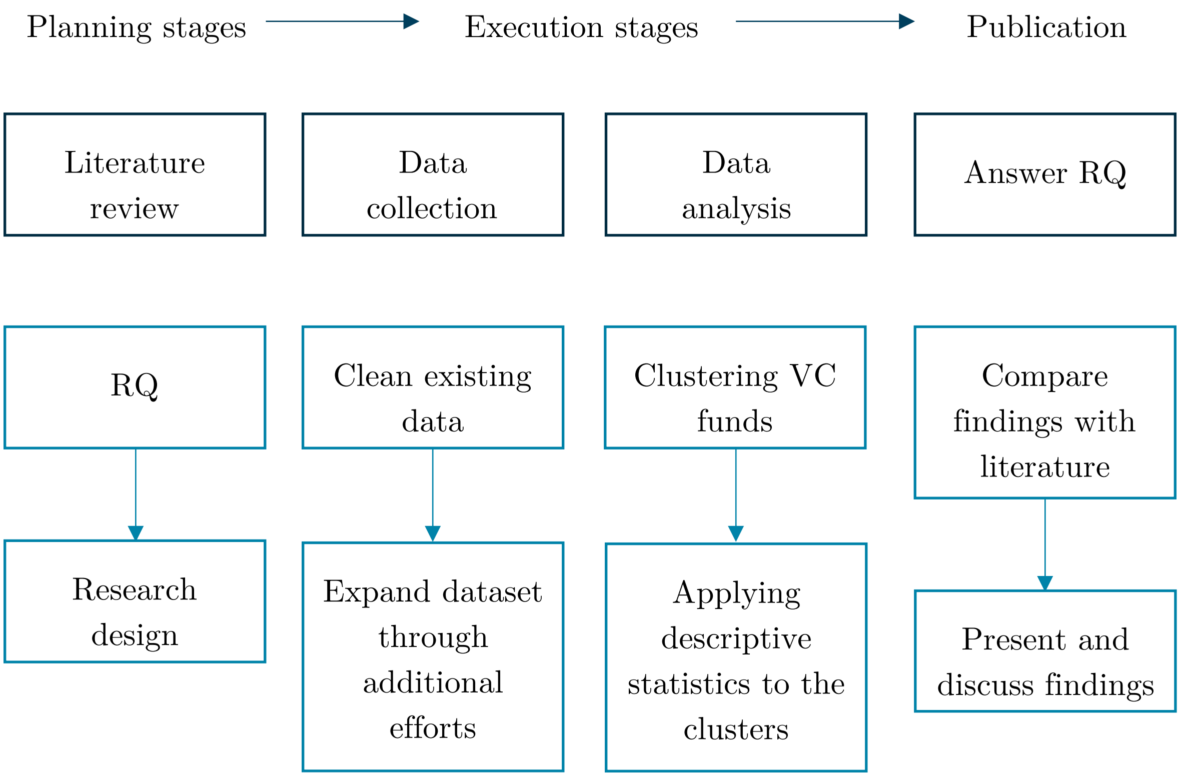


Figure 10: Steps of the study process

3.1.2 Literature Review

This section outlines applied literature and findings from the literature used to conduct this thesis. The majority of the literature emerged from the project thesis the authors conducted, titled "Exploring Venture Capital Fund Performance and Determinants Driving the Performance" in the fall of 2021. In addition to this project thesis, further literature research has been conducted, especially to further expand the literature on data-clustering and descriptive statistics to create a theoretical basis for the data analysis.

A thorough and refined literature review lays the foundation for meaningful and valuable research (Boote & Beile, 2005). Through the literature review, the authors gained insight regarding current VC data quality, accessibility constraints in the venture capital industry and the current research's stance on the influence on performance of structural, operational, and wider-environmental factors. Key findings from the literature were that the structural determinants had the most apparent impact on performance. However, most of the research had used the same research methodologies and the same datasets. Therefore, the fact that most studies found the same performance effects were not surprising.

Of the articles used in the project thesis and the additional literature review conducted as part of this thesis, the majority of the literature is from before 2010. In later years, less has been written on the topic of determinants affecting performance.

Table 1: Overview of the year articles were published.

Year	**# of articles**
Before 2000	15
2000 – 2005	11
2006	3
2007	7
2008	8
2009	4
2010	3
2011	3
2012	1
2013	1
2014	2
2015	0
2016	0
2017	4
2018	4
2019	0
2020	1
2021	1
2022	0

Of the determinants studied in previous literature, structural and operational determinants have been the most widely used.

Table 2: Overview of determinants studied in existing articles.

Determinant category	Number of articles
Structural	28
Operational	43
Wider environmental	6

Most previous research has either used US VC databases like VentureXpert and Venture Economics. Few have used European databases such as national VC databases, CEPRES or PREQUIN. A good amount has either expanded on existing datasets or created their own through questionnaires or through their contact network.

Table 3: Overview of datasets used in existing articles.

Dataset	Amount used
VentureXpert	23
Venture Economics	9
Venture Source	1
Zdatabase	1
Capital IQ	2
CEPRES	3
PREQUIN	2
BVCA (British Venture)	1
CVCA (Canadian Venture)	2
Dealogic	2
Australian Bureau of Statistics	1

Exucomp	2
Diane	1
FAME	1
Custom	17

Of previous research methodologies, almost all aim to isolate a specific determinant and quantify its impact using a form of regression analysis. Only a limited number of articles have attempted other approaches, such as using graph theory, time-series analysis or using qualitative approaches such as a multi-theoretical conceptual framework.

Table 4: Overview of research methods used in existing articles.

Research Methodology	Amount Used
Regression, single variable	33
Multivariate Analysis	6
Network Analysis, graph theory	1
Time-series analysis	1
Descriptive statistics	4
Multi-theoretical conceptual framework	4

From the literature review, a research gap regarding measuring and understanding venture capital fund performance in a European and Nordic context was identified, which this thesis aims to cover. A deductive approach to this thesis was chosen, and the existing literature guided the authors through the study (Wilson, 2014).

After the data analysis, the literature was slightly expanded to include research on descriptive statistics and data clustering to ensure meaningful conclusions could be drawn from the study and the research process would be easy to follow for the reader.

3.2 Data Collection

The data collection was conducted to provide a rich and unique understanding of how structural factors impact venture capital performance. The basis for the data is threefold; (1) Argentum's dataset, (2) NVCA's dataset, and (3) self-collected data through market research. The data collection process is described in the following subsections.

3.2.1 Argentum Data

Argentum is a leading private equity investor in Northern Europe managing investments on behalf of the Norwegian government and institutional investors. Argentum was incepted in 2001 and has invested over € 1.9 Bn in more than 200 funds.

Argentum's dataset consisted of fund-level data on 198 funds managed by 102 unique firms across Norway, Sweden, Finland, and Denmark. The data was structured as follows:

- Investment Year: 2006 - 2021
- Fund Manager (VC firm)
- Fund Name
- Size (M EUR)
- Country: Norway, Sweden, Denmark, Finland
- Stage: Seed, Venture, Growth, Buyout
- Industry: Life Science, ICT, Energy, Cleantech, Food, Manufacturing, Forestry, Oceantech, Edtech, Impact and Generalist.

3.2.2 NVCA Data

NVCA was founded in 2001 by the leading players in the Norwegian venture capital sector. NVCA is an independent, non-profit association supporting the interests of entities acting in the private equity and venture capital markets. NVCA is responsible for reporting Norwegian venture capital activity to the European Venture Capital Association (EVCA).

NVCA's dataset consisted of investment-level data on 3684 investments by 576 unique firms in 2482 unique companies (investment targets). Although the dataset is oriented around the Norwegian market, it covers a vast number of transactions in rest of the Nordics. The data was structured as follows:

- Investment year: 2008 - 2021
- Fund Manager (Firm)
- Investor country: 37 different countries
- Company Name (investment target)
- Headquarters: Norway, Sweden, Denmark, Finland
- Stage: Venture, Seed, Growth, Buyout
- Industry: ICT, Cleantech, Healthcare & Life Science, Industrials, Energy, Consumer, Financials, Utilities, Other

Because the NVCA data does not contain any fund level data, this had to be added manually by the authors. The majority of the firms contained in the NVCA dataset were not Nordic funds and were thus excluded from this study.

3.2.3 Data Collection Through Market Research

We extended the dataset from Argentum and NVCA by conducting market research. The majority of the market research consisted of analyzing official information provided by the firms through their own channels, e.g., press releases, websites, social media, etc.

Because the NVCA dataset did not contain fund level data, we could not use this data directly. To make the NVCA data applicable in this study, the authors found the firms operating in the Nordics not contained in the Argentum dataset. For these firms, the authors manually inspected the firms' websites to find the fund level data. This data was then appended to the Argentum data.

By expanding the NVCA's dataset to contain fund level data in combination with Argentum's fund-level dataset we created a basis consisting of 145 unique firms, e.g., an increase of 43 firms compared to the initial dataset provided by Argentum.

In order to fill the missing metadata on the 43 new firms, we started to review official information published by the firms. Three main sources of information were investigated: (1) press releases, (2) websites, and (3) social media. The metadata we tried to find was:

- Fund Name
- Year
- Size (M EUR)
- Country
- Stage
- Industry

As we were collecting data for the new firms, we noticed that the original dataset from Argentum had discrepancies from our findings. Thus, we also reviewed the metadata listed above for the 102 original firms. This led to both updating metadata, and adding new funds previously not listed. After both expanding the dataset with new firms and their funds, in addition to updating information on existing firms, we were able to increase the fund-level data from 198 to 404 funds.

3.2.4 Summary of Data Used in this Thesis

After collecting new data and updating existing data through market research, the authors have created a unique dataset containing significantly more fund level data than the original datasets. The data-quality has been improved, and we believe the dataset now provides a more correct and unbiased view at the Nordic VC ecosystem.

For the fund level data, both additional funds from the existing Argentum dataset, and additional funds uncovered through new firms in the NVCA dataset has been added. The Argentum dataset mainly contains data from Norwegian firms, but the NVCA database has data evenly distributed for the entire Nordic region. For this reason, the authors believe the resulting dataset to be less biased towards Norway, and a better representation of the Nordic region.

Table 5: Overview of original dataset from Argentum, our dataset, and the % increase in number of entries.

Fund level data	Argentum	Our dataset	% increase
# of Firms	102	145	42 %
# of Funds	198	404	104 %

Through the data collection, fund level data have almost doubled for each country in the Nordic region. Because the firm data has become more representative for the entire Nordic region, one interesting finding is that Norway still has significantly more funds than the rest of the region. One explanation for this might be that Norway has significantly many more funds than the rest of the Nordics, but a smaller average fund size.

Table 6: Breakdown of the original and expanded dataset in regards of geographical coverage.

Geography	Argentum	Our dataset	% increase
Norway	74	161	118 %
Sweden	47	94	100 %
Denmark	22	43	95 %
Finland	55	96	75 %

The original dataset did not contain any data concerning the growth and buyout stages. By collecting additional data, we have been able to create a more even distribution by investment stage, and thus hope to achieve an accurate dataset.

Table 7: Breakdown of the original and expanded dataset in regards of stage coverage.

Investment Stage	Argentum	Our dataset	% increase
Seed	60	97	62 %
Venture	138	193	40 %
Growth	0	33	N / A
Buyout	0	71	N / A

Interestingly, the data collection significantly expanded investments before 2011. Argentum's data included data from 2006, and NVCA's data started in 2008. Therefore, we have been able to significantly increase the timespan in which we can inspect performance. In addition, we believe this expansion in early funds will let us create a better representation of performance, as many of these early funds were in fact the first funds of firms already included in our dataset. Thus, we have a much more complete fund-level history for each firm, and therefore a better data-basis for exploring performance.

Table 8: Breakdown of the original and expanded dataset in regards of year coverage.

Year	Argentum	Our dataset	% increase
<2011	66	195	195 %
2011 – 2016	58	86	48 %
2017 - 2022	74	110	49 %

3.3 Data Analysis using K-Means Clustering & Descriptive Statistics

The process of finding useful research or business insight from a dataset is called data mining or data analysis (Jackson, 2002). Data clustering is a primary methodology in data analysis to ensure high data-quality and subsequent high-quality insights (Kameshwaran & Malarvizhi, 2014). Descriptive statistics are used to summarize and explain a dataset (Kaur et al., 2018). In this paper, the data analysis process is presented sequentially, starting with data-cleaning before we create our proxy for performance and finally run k-means and inspect the results using descriptive statistics. The entire data analysis process was completed using the programming language python, and the code used for this process is included in Appendix B. For ease of understanding the research methodology, we recommend the reader inspects Appendix B while reading this section.

3.3.1 Data Cleaning & Preparation

Before clustering the data, we must ensure our data is uniform and in the correct format (Kasliwal et al., 2012). Datapoints that do not represent our population must be removed (Kasliwal et al., 2012), like funds not in the Nordics or firm with only one fund.

After loading the data in (Appendix B, section 1.1. Hereafter, B1.1), we remove funds not in the Nordics (B1.2). This is done because we do not want to include funds outside of the Nordics. Subsequently we remove new firms with only one fund (B1.3), because we don’t want these to end up together with defaulted funds, and there is too few datapoints to measure performance. Thereafter, we remove discontinued funds (B1.4) and funds with missing data (B1.5). This leaves us with a dataset of only Nordic firms with valid data.

3.3.2 Proxy for Performance

As discussed in section 2.3.4, there has been several attempts at using proxies to determine performance. Hochberg et al. (2007) used exit-rates as a proxy for performance. Smith et al. (2010) conclude exit rates are an inaccurate measure for performance, and that this proxy is only used because data constituting a better proxy was not available.

Morris et al. (2020) suggested making a proxy based on markups, i.e, VC firms investing at a higher price than previous VCs. The paper concludes that markups are an inaccurate performance proxy because markups are mostly a result of higher access to capital and not the underlying performance of the firm.

Kaplan and Schoar (2005) found that high performing funds raise funds that are relatively larger than worse performing funds. This effect lasted across several funds, and outperforming funds were shown to continue the outperformance in subsequent funds. Kaplan and Schoar (2005) also found that firms with more experience and more funds typically outperformed firms with fewer vintage funds. For high performing funds that had grown significantly larger than the investment stage average, the percentage increase per fund decreased without this affecting performance (Kaplan and Schoar, 2005).

The effects suggested by previous research on performance proxies have been considered when selecting a proxy in this thesis. The first part of our performance proxy looks at the relative difference between subsequent fund sizes. This effect aims to capture the fact that high performing funds raise larger funds faster than worse performing funds. The second part of the equation adds the relative difference between a fund and the investment stage average for the fund. As funds that grow significantly larger than the market average experience slower growth while remaining overperformers, this part of the equation aims to capture this effect. As a result, our proxy for performance becomes:

Equation 4: Adding subsequent fund deltas and investment stage deltas to form a performance proxy.

$$\sum\left(\frac{curr.fund}{prev.fund}-1\right)+\left(\frac{curr.fund}{stagemean}-1\right)$$

To implement this proxy for performance in our study, we start by calculating and adding the stage mean (B2.1.1), the average time to raise a new fund (B2.1.2) and the mean fund size for each firm (B2.1.3). Thereafter, we calculate the percentage delta between sequential funds (B2.2.1) and the percentage delta between the current fund and the stage mean (B2.2.3). After calculating the individual deltas, we add them together to form our metric (B2.4).

3.3.3 K-Means Clustering

There are several statistical methodologies used in research to categorize data-points. Primarily, either classification or clustering is utilized. The most important difference in research applications is that classifications are predictive (Kaushik & Mathur, 2014), while clustering is descriptive (Kasliwal et al., 2012). As we intend to describe the characteristics of the Nordic VC ecosystem, clustering is the most appropriate data-categorization methodology in this paper.

The goal of clustering is to group sets of data that have similar characteristics, while remaining dissimilar to the data-points in the other clusters (Kaushik & Mathur, 2014). In order to describe how structural factors impact performance of VC firms, we cluster VC firms with similar performance.

The most commonly used clustering algorithm is the unsupervised machine learning algorithm K-means clustering (Kaushik & Mathur, 2014; Forgy, 1965). K-means clustering works by selecting k clusters and randomly initializing cluster centers, and then iteratively moving the cluster centers to a location where the Euclidean distance to each data-point in the cluster is the smallest (Yuan & Yang, 2019). K-means is widely used as it groups data accurately into the desired number of clusters (Rai & Shubha, 2010).

There are certain limitations to k-means clustering. The data-points must be continuous and can't be discrete (categorical). The primary weakness of k-means is that it is very sensitive to outliers (Kaushik & Mathur, 2014). Thus, we in this paper remove outliers before applying the k-means clustering. Outliers is defined as a sparse quantity of data-points with significantly higher or lower values than the rest of the dataset, creating a bias for the cluster centers (Kaushik & Mathur, 2014). Because outliers significantly impact the cluster centers, outliers can be removed or manually clustered before k-means is applied (Rai & Shubha, 2010). In addition, data that can't be correctly clustered by k-means should be set as its own cluster (Rai & Shubha, 2010).

A silhouette clustering criterion is a commonly used methodology to find the appropriate number of clusters to use in the k-means algorithm (Yuan & Yang, 2019). Where the silhouette criterion yields similar results for several k-values, we choose a small k-value that is close to the maxima and fits the research criterion (Yuan & Yang, 2019). In this study we end up with 4 cluster centers with a silhouette score of 0.41 as this is close to the maxima of 0.43 and allows us to easily categorize the clusters as bottom-performers, mid-performers and high-performers.

To minimize spurious local minima problems (Fränti & Sieranoja, 2019) and obtain robust results, we ran the algorithm 10 times and chose the result with the lowest sum of distances found.

In our study, firms with only one fund can't be clustered properly, as we can't estimate their performance. Thus, these funds are clustered together manually (B3.1.1). In addition, there are three significantly outperforming firms, and we group these together manually (B3.1.2). Before running k-means, we must also normalize and standardize our data (Kaushik & Mathur, 2014). We do this to get the correct Euclidian distances, and thus the correct cluster centers in (B3.2.1 & B3.2.2). Then we calculate the silhouette scores (B.3.2.3). Silhouette scores are very similar for all numbers of k-clusters, and we choose a k value of 4 with a silhouette score of 0.41. After preparing our data and finding k, we run the k-means algorithm to categorize our data (B3.1.2).

3.3.4 Descriptive Statistics

After creating the clusters, the next step was to evaluate the determinants of the groups obtained, in order to understand which structural factors impact performance. Descriptive statistics is a common methodology for drawing conclusions from data (Kaur et al., 2018). In this study we use descriptive statistics and plots to observe differences in the structural factors. There are many different descriptive statistical methodologies that can be utilized, and in this paper, we employ several of them. As a comparison of the clusters, we examine number of funds raised, which utilizes the descriptive methodology of absolute frequency measuring (Kaur, Stoltzfus & Yellapu, 2018). Furthermore, we examine averages for time to raise subsequent fund, growth rate and absolute size of funds. For comparing stage focus we use the descriptive methodology of measures of position, or percentile placement to see differences in stage focus for the clusters.

To conduct a descriptive statistical comparison of the different cluster characteristics, we download the resulting data from running our k-means algorithm (B4.2.2) and plot the data using Microsoft Excel.

3.4 Reflection of the Methodology

To assess the research quality of a study, a common evaluation criterion is trustworthiness. The authors use the framework proposed by Lincoln and Guba (1985) which discuss the internal validity, external validity, reliability, and objectivity of the study.

3.4.1 Internal Validity

A study is internally valid if it is able to determine whether a causal relationship exists between one or more independent variables and one or more dependent variables (Heffner, 2017). Our initial datasets, provided by Argentum and NVCA, were produced by direct interaction between VC partners and research professionals in the mentioned organizations. Since the data is collected yearly over a long-term, we assume the data quality to be acceptable. However, as the data is provided by a third-party, we cannot validate the data quality beyond the methodology details provided in an interview.

The self-collected data is collected through analyzing official information provided by the firms through their own channels, e.g., press releases, websites and social media based on a fixed list of firms, without any interaction with the VC firms, VC partners or related parties. This makes the firms in the dataset unaware of the data collection to avoid biasing the perception and behaviors and thus the outcome of the study.

Moreover, a strict data collection protocol is used that outlines the procedures of the data collection. If VC funds are omitted from the data basis, their characteristics are examined to make sure there is no systematic bias in terms of which funds stays in the data basis.

3.4.2 External Validity

External validity describes the ability to generalize a study, which is particularly threatened if people, places, or times are poorly chosen (Trochim, 2006). The population of interest is clearly defined as Nordic firms in the venture capital segment and the data collection is aimed to include all firms in the particular segment. Thus, the population sample is assumed to be a representative sample of the Nordic VC market.

The self-collected data is collected from publicly websites within a short time-period, hence a content-changing environment based on the source authors' call. This introduces situational factors concerning the data collection timing, consequently data collection performed at an earlier or later occasion may result in different data points collected. Thus, leading to limited generalizability of the findings.

3.4.3 Reliability

Reliability means repeatability and is achieved if a measurement always creates the same result (Trochim, 2006). To obtain the same datapoints for each VC firm in the dataset, we formulated a strict data collection methodology with specific search words and guidelines. This lets us find the required data without being influenced by concentration of the researchers.

Multiple researchers were involved in the data collection and analysis. When data points were collected and assigned subjective ratings or categories, we made sure that all the researchers observed and assessed the same thing. In this way, we aimed to minimize subjectivity as much as possible so that a different researcher could replicate the results.

3.4.4 Objectivity

Objectivity describes the researchers' ability to remain distanced from what they study to generate findings based on what was studied rather than personality, beliefs and values of the researcher (Payne & Payne, 2004). Since the self-collected data is collected without interaction with participants, it is ensured an appropriate distance between the researchers and the VC firms to lessen bias.

3.5 Limitations of the Methodology

The results reported in this study should be considered in the light of certain limitations. Despite efforts in constructing a comprehensive and complete data basis, the data sample still involve several shortages.

Firstly, information concerning discontinued VC funds and VC firms is often missing and not available on the web due to no official communication channel from the VC firm. This is mainly for discontinued or defaulted funds, which performance we can judge to be poor. This results in possible incomplete data registered on discontinued VC funds and a potential survivorship bias in our data.

Secondly, the dataset is perceived as the largest achievable on VC funds in the Nordics in 2022. With a bigger dataset, it would be easier to use more advanced statistical methodologies as well as possibly a clearer answer to the differences between clusters. This might not be possible in the Nordics, because we simply do not have many more funds to analyze. For a slightly larger ecosystem, more conclusive findings on performance could be possible.

The websites of the VC firms are continuously changing to reflect their current state. For many of the discontinued VC firms, their website was no longer available and important information was lost. For existing VC firms, there were several occasions where the best funds were easily visible, while less impressive fund data was harder to find. Because the websites change continuously, the repeatability of the study might be affected if performed a significant amount of time after the publication of this thesis. Therefore, it is advised to carefully examine the VC's self-reported data, as it might be different from the dataset used in this study. In addition, the dataset obtained from both Menon and Argentum were acquired in confidence that it would not be shared. Thus, we are not able to easily attach our aggregated dataset in this study to increase repeatability.

The VC industry is very private and most of the reported numbers are aggregations of self-claimed rates of returns. As a consequence, the availability of quality data on VC performance is limited and a proxy for performance is required. However, it exists limited previous research studies defining VC performance based on factors excluding rate of return. This makes our proxy for VC performance limited empirical tested and verified.

4 Results and Analysis

Venture capital funds operate in a number of sectors, across various geographic regions and target different investment stages. All of this makes the analysis of venture capital fund performance a complicated task with many intertwined variables.

In this study, multidimensional cluster analysis is applied to group venture capital funds according to common performance characteristics, i.e., the percentage delta between subsequent funds and the delta for a fund and the investment stage average. In this chapter we inspect each cluster and their structural characteristics.

4.1 The Current State of the Nordic Venture Capital Industry

In this section, we will present the results of collecting the most up-to-date dataset on the Nordic VC industry. Previous datasets utilized for research-purposes in the Nordic VC industry are sparse and only dated until 2021, whereas the dataset used in this study involves a comprehensive expansion of original datasets and including data from the first half of 2022. This section aims to provide the current state of the Nordic VC industry.

When we observe the absolute amount raised per country per year, as shown in Figure 11, Sweden outperforms its Nordic neighbours. Sweden raises €38 Bn. in the period 1989-2022, almost twice as much as Norway with its €22 Bn. Both Finland and Denmark are relatively far behind with €7 Bn. and €4 Bn. respectively in raised capital during the period. Notably, we observe an incline increase post 2015 for all countries in the Nordics, especially Norway and Sweden.

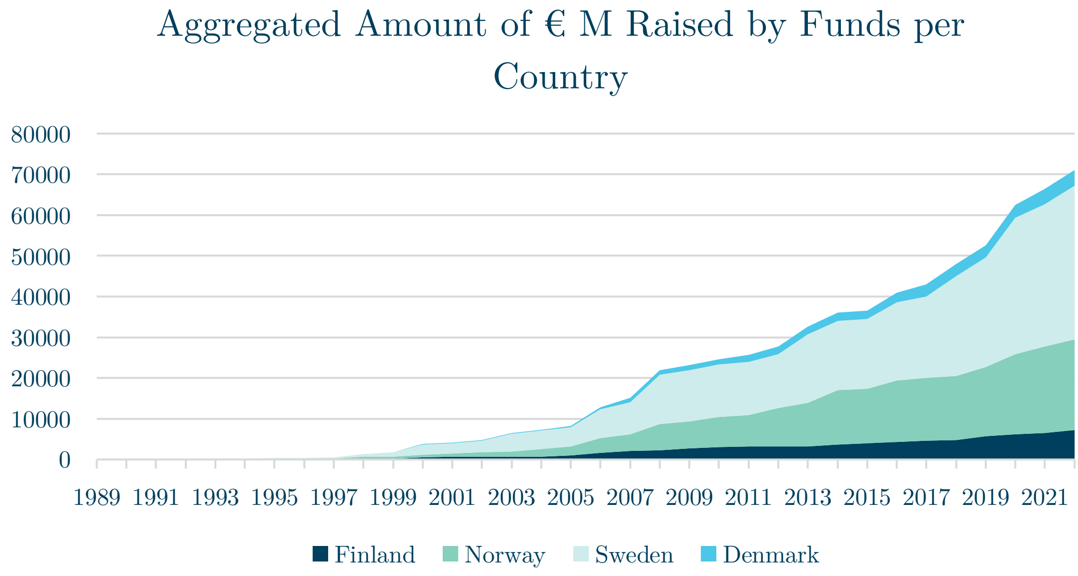


Figure 11: Aggregated amount of euros raised by funds per country.

Later-stage VC funds are underrepresented in the Nordic countries, especially in Denmark, as observed in Figure 12. Sweden has a similar number of seed, growth and buyout funds, while having a significantly larger portion of venture funds. Finland functions as an average of all the other countries with a quite even distribution of funds. Norway has the most funds in every stage and a distribution quite similar to Finland.

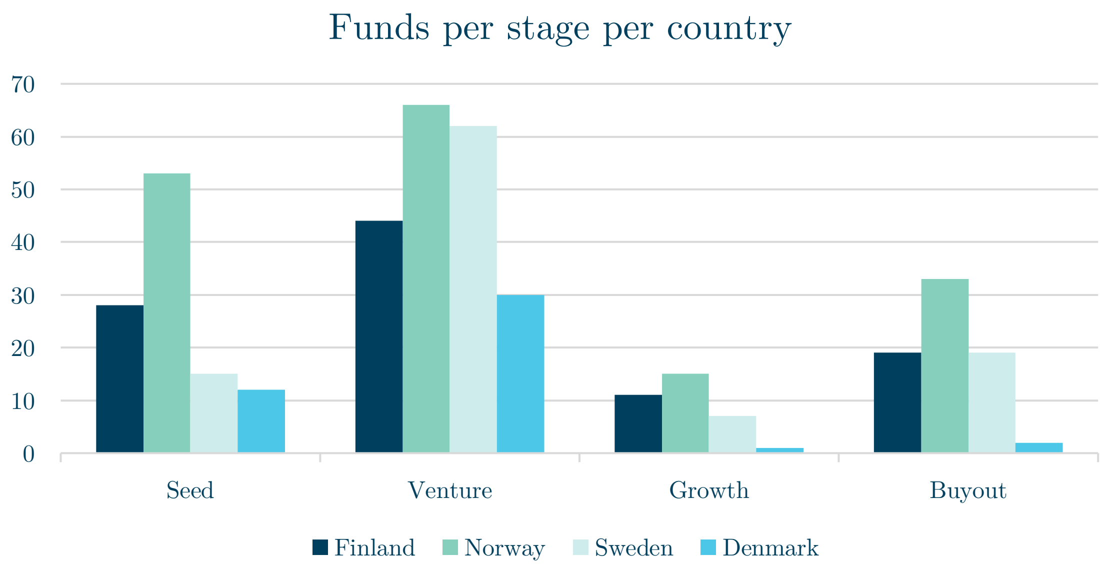


Figure 12: Number of funds per stage per country in Nordics

We observe in Figure 13 that Finland, Norway, and Sweden all have mostly generalist funds in the period 1989-2022. On the other hand, Denmark majorly has ICT focused firms, the second most common industry specialization for the other Nordic peers. Generalist and ICT are by far the most common industry specializations for VC funds in the Nordics, whereas we observe smaller allocations of VC funds with a specialization towards life science, energy or cleantech. Notably, Norway is the only country in the Nordics with VC funds that have an industry specialization towards energy.

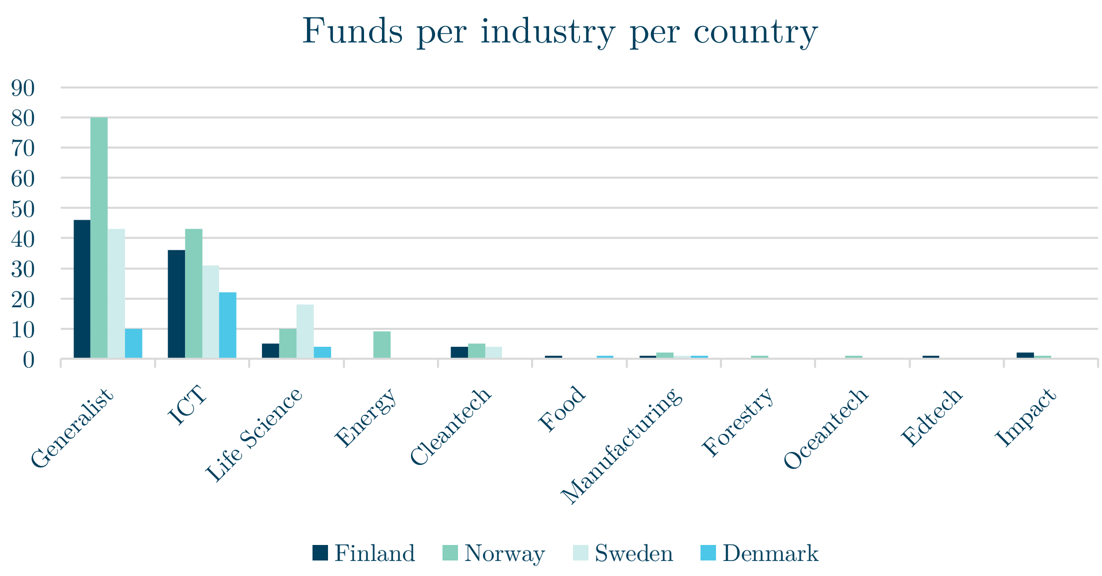


Figure 13: Funds per industry per country.

As seen in Figure 14, the number of newly established funds has in recent years been close to all-time high due to a solid increase in especially Norway and Finland compared to previous years. The results also highlight the two previous boom and bust cycles with the Dot-com boom (and bust) in 1999-2003, and the financial crisis of 2007-2008. Another interesting finding is that Denmark is raising substantially fewer funds than its Nordic peers. When we aggregate the datapoints in the figure below for the 1989-2022 period, Norway outperforms the rest with 166 funds raised. Sweden and Finland follow with 102 and 103 funds.

Denmark is lagging with 45 funds raised in the period from 1989 to 2022.

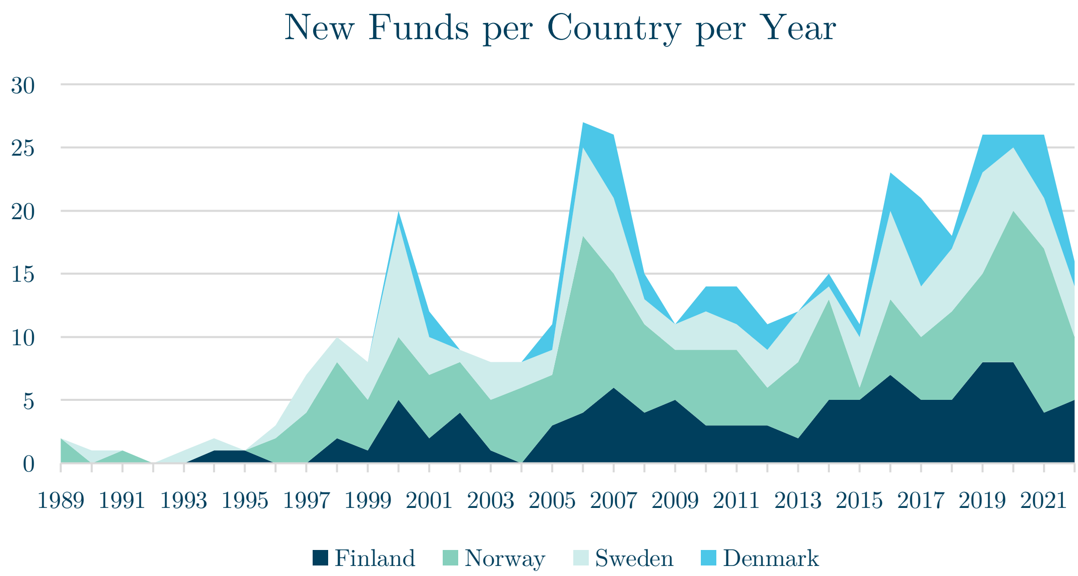


Figure 14: New funds per country per year

4.2 Cluster Analysis

The cluster analysis suggests that there are six different kinds of Nordic VCs, some more prevalent than others. Named after their growth pattern, we find Defaulted, Slow-raisers, Commoners, Fast-raisers, Visionaries, and Superstars. Figure 15 gives an overview of the six clusters. These six clusters are then grouped into three distinctive categories, namely bottom-performers, mid-performers and top-performers.

Figure 15: The six different clusters resulting from the k-means clustering.

Table 9: Descriptive statistics of clusters

Cluster	# of VCs	Avg. time to raise	St.dev of time to raise	Avg. increase in fund size	St.dev of increase in fund size
Superstars	3	1.7	0.701	41.376	5.375
Visionaries	14	2.3	0.586	10.665	4.44
Fast-raisers	34	1.1	0.470	0.545	2.843
Commoners	31	2.5	0.388	0.493	1.981
Slow-raisers	10	4.2	1.086	2.644	1.568
Defaulted	25	-	-	- 0.291	0.785

4.2.1 Bottom Performers; Defaulted & Slow-raisers

The bottom performing defaults and Slow-raisers share some common characteristics. New funds are raised slowly, and new funds are typically not much larger than the previous funds. In addition, bottom performing funds tend to be smaller than the investment stage average. In the following section we present the results for the underperforming firms.

Defaulted

Defaulted firms are either VCs that have announced their resignation or VCs established before 2017 only being able to raise one fund. Defaulted firms make up 21 % of the firms. Defaulted firms are able to raise 1.08 funds on average, implying that having a successful first fund is of great importance to survive in the VC industry. As the majority of the firms in this cluster only raise one fund, we do not have any significant statistics concerning the average time to raise the following fund or the average increase in fund size for the following fund. However, the defaulted firms tend to raise substantially smaller funds compared to peers investing in the same phase (39 % smaller funds than phase-specific average).

Slow-raisers

Slow-raisers are the slowest raising firms and the second smallest cluster making up 8.5 % of all the firms in our dataset. The VC firms raise at average 3.8 funds that are 20 % larger than their phase-specific average. The Slow-raisers are characterized by their slow time to raise new funds, taking between 3 – 7 years on average. On average, the Slow-raisers increase their fund size 58 % for each following fund.

4.2.2 Mid Performers; Commoners & Fast-raisers

The mid-performers are numerous representing 55 % of the firms in the study, whose performance is best described as typical for the ecosystem. For being mid performers, the growth of these funds is much larger than one might expect. Mid-performers record positive growth rates across the performance proxy metric, increasing each subsequent fund by roughly 80 %. Even though these funds grow with an impressive pace, these funds still remain 40 % smaller than the ecosystem average. This speaks to the immaturity of these funds, as the mid performers are typically less experienced.

Commoners

The Commoners are the second largest cluster and make up 26 % of all firms in the study. The Commoners raise 3.38 funds on average and achieves an impressive 81 % fund-to-fund growth in fund size. The Commoners record the lowest fund size compared to peers, averaging 48 % lower than their stage average. The Commoners raise quick follow-up funds, averaging 2 years and 10 months for subsequent funds. Even though this speed might seem impressive, it is slower than its mid-performing peer: the Fast-raisers.

Fast-raisers

Fast-raisers is the largest profile in the study, including 29 % of all firms in this study. Fast-raisers is the profile with the lowest average time to raise a new fund at a staggering 1 year and 5 months. Despite the incredible pace, the average fund size increase in the following fund averages at an impressive 87 %. Compared with their stage-specific peers, Fast-raisers raise funds in average 42 % smaller in fund size, speaking to their immaturity together with the Commoners. The average number of funds raised by firms in this cluster are 3.42, meaning that they barely beat Commoners, but still are lagging behind the Top Performers.

4.2.3 Top Performers; Visionaries & Superstars

When everything goes according to (business) plan, firms continue raising faster and bigger funds. This is what differs the Top Performers from the rest: namely the ability to raise the following fund in a short amount of time in addition to increasing the fund size considerably. The mid performing Commoners and Fast-raisers are easily beat in growth rate by the Top Performers, raising over 160 % larger subsequent funds. This growth rate and fund sizes can’t be described as anything less than extraordinary.

Visionaries

Visionaries performs very well both in raising large and fast, ranking them as the second-best cluster. The Visionaries make up the third largest clusters of the VC firms, and averages 5.5 raised funds. The Visionaries raise fast, averaging 2 years and 6 months between funds. These firms have an incredible growth rate, with subsequent funds increasing in size by 157%. Not only do these funds grow extraordinarily fast, in fact these funds are typically very large at 38% bigger than the investment stage average.

Superstars

Superstars are the best performing profile according to the performance proxy metric. The profile makes up 3 % of all firms, making it the smallest cluster. Superstars raise the following fund in 2 years and two months on average, and the average increase is 192 % fund-on-fund. The Superstars averages 10 funds raised and exceeds its peers with 377 % in terms of fund size raised.

4.2.4 Comparing the Clusters

We have seen that high-performing firms have raised significantly many more funds than their peers. Especially the Superstars have a much higher average number of funds than even their high performing visionary peers. Only raising faster does not seem to be the key to success, as the fastest raising funds are not the best performing. When considering growth rate, high performers significantly outshine their peers. Their growth rate average is more than twice that of the mid-performing firms. The high-performers funds are also significantly larger. Again, we can see that the Superstars outclass even the Visionaries by being more than twice as large on average. As a result, we can see that a common characteristic of a high-performing firm is that they are experienced, grow significantly and raise big, but are not necessarily the fastest.

4.3 Structural Characteristics

In the following sections, we will present the characteristics of the clusters obtained to attempt to uncover which structural factors impact performance. The results are given through plots and descriptive statistics related to specialization in geographical focus, industry or investment stage, age of VC and VC fund size.

4.3.1 Geographical Focus

Figure 16 shows the percentage allocation of the geographical focus of each cluster. We observe allocation of investment focus towards Denmark is primarily within the lower half of performing clusters, such as Defaulted, Slow-raisers and Commoners. On the other side, we clearly see a trend towards both Finland and Norway in the upper half. Consequently, the geographical focus in Top Performers has an allocation of 42 % towards Finland, 36 % towards Sweden, 17 % towards Norway and 4 % towards Denmark. Across all clusters, the pattern of allocation towards Sweden is more ambivalent. Bottom Performers, Mid Performers and Top Performers have an allocation towards Sweden of respectively 52 %, 19 % and 36 %. We notice that Slow-raisers does not include any VC firms with an investment focus towards Norway.

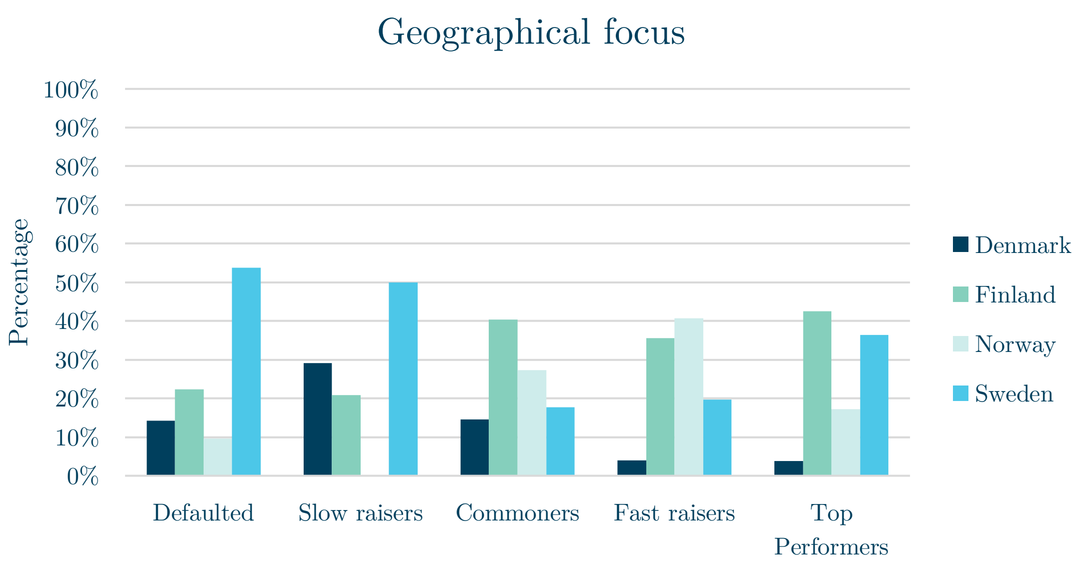


Figure 16: Geographical focus per cluster.

4.3.2 Industry Focus

Any VC that currently has or previous had a majority of funds investing in more than 3 industries are characterized as generalists. Less than 3 industries are proportionately weighted within every industry. Figure 17 reports the allocation of industry focus, showing a majority towards both generalist and ICT across all clusters. We observe the lowest allocation towards generalist among Top Performers compared to peers. Noticeable, we observe that the group Defaulted has the most wide-ranging industry focus, with more VC firms focused on specific industries that are not present in other clusters, e.g., food, manufacturing, edtech or impact. Lastly, we acknowledge that Fast-raisers has the largest allocation towards generalists with an allocation of 56 %.

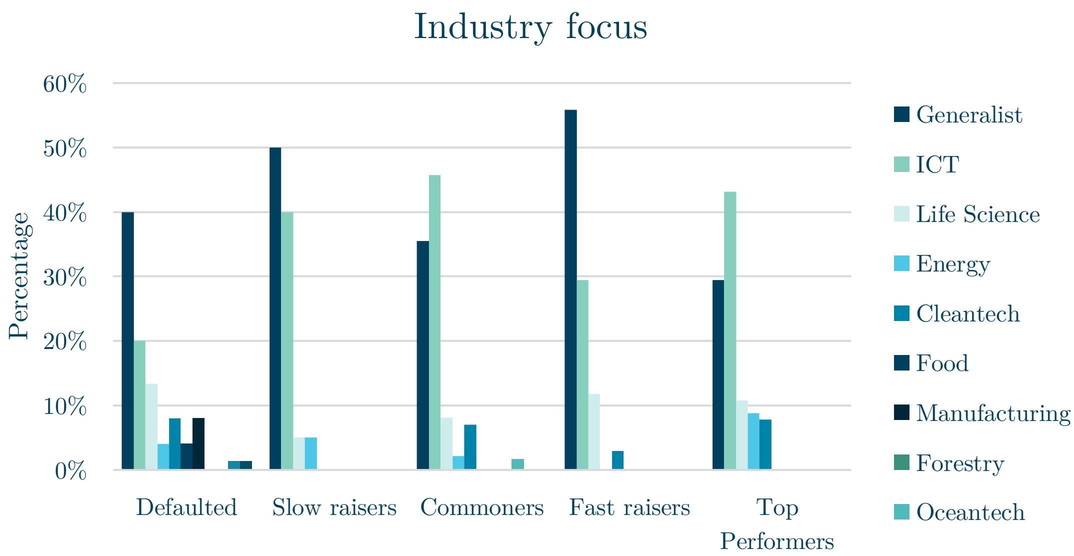


Figure 17: Industry focus per cluster.

4.3.3 Stage focus

Figure 18 presents the output from the data analysis of the VC firms' specialization in stage. We observe a general skewness towards the venture stage across all clusters. Bottom Performers and Mid Performers have corresponding stage focus towards venture, whereas Top Performers has the smallest allocation of stage focus towards the venture stage. Stage focus towards the seed stage is evenly distributed across clusters. On the other side, we observe a trend of stage focus towards the buyout stage at higher performing clusters, such as Fast-raisers and Top Performers, with the highest reported allocation towards the buyout stage of 24 %.

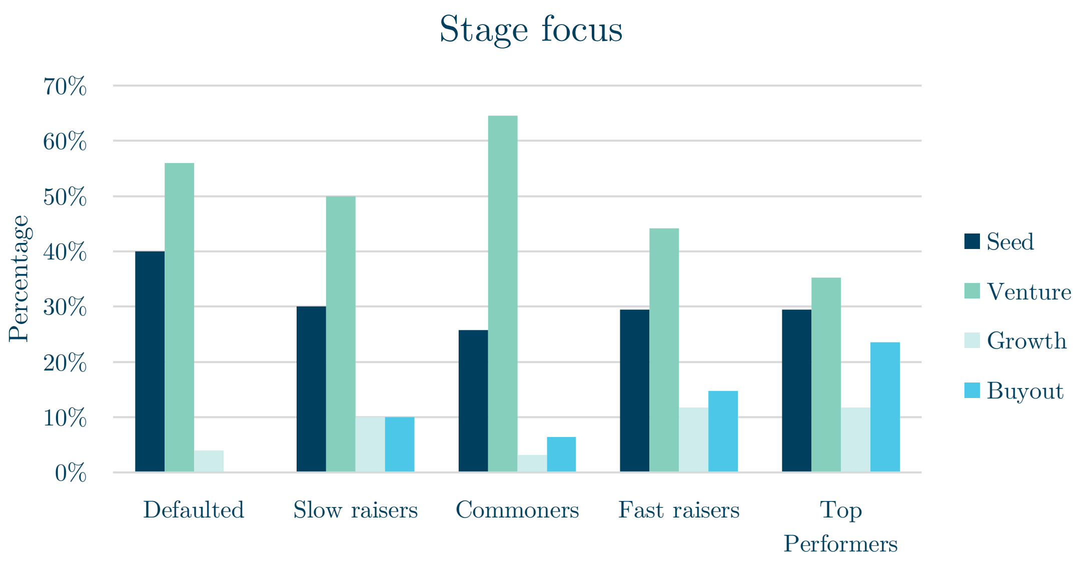


Figure 18: Stage focus per cluster.

4.3.4 Age of VC firm

When looking at the results in Figure 19, we see no clear trend regarding the age of VC firm. We observe that Defaulted and Fast-raisers are established relatively recent, respectively in average 2010 and 2011. On the other side, both Slow-raisers, Commoners and Top Performers is observed as relatively longstanding VC firms, with an average establishment year to respectively 2002, 2004 and 2005. Additionally, we notice that the standard deviation increases in line with performance from Defaulted to Top Performers with the values from 5.5 to 14.

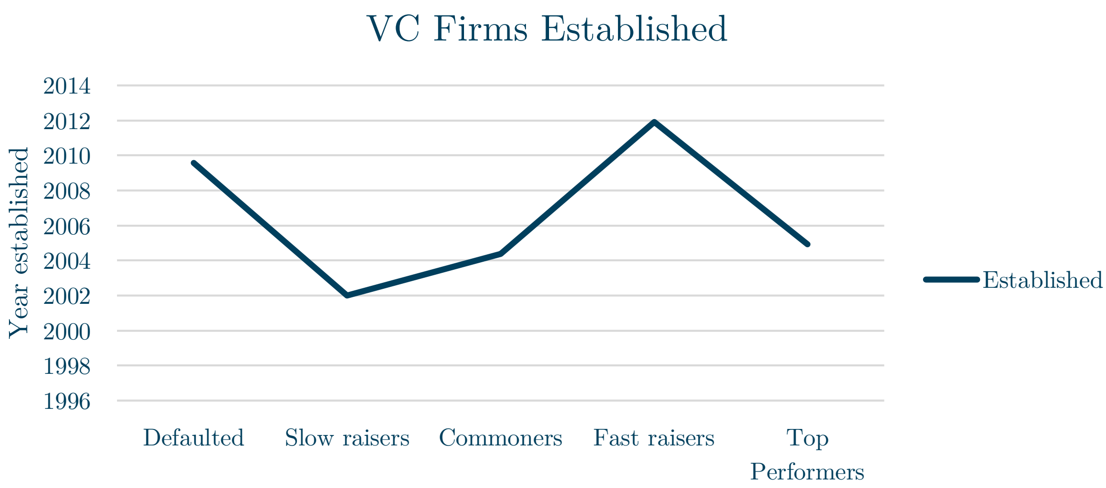


Figure 19: Year established per cluster.

4.3.5 Fund size

The results in Figure 20 of mean fund size across the clusters show that Top Performers has considerable higher mean fund size than peers with a fund size of € 426 M. However, we observe the largest standard deviation of € 573 M for Top Performers. For the clusters with lower performance than Top Performers, we observe a slight diverse distribution. Fast-raisers and Commoners report a mean fund size respectively of € 75 M and € 56 M, with a corresponding standard deviation of 87 and 68. On the same track, we observe the mean fund size of Slow-raisers and Defaulted at € 101 M and € 49 M. Slow-raisers is observed with the lowest proportionately standard deviation of € 59 M, whereas Defaulted has a standard deviation of 74.

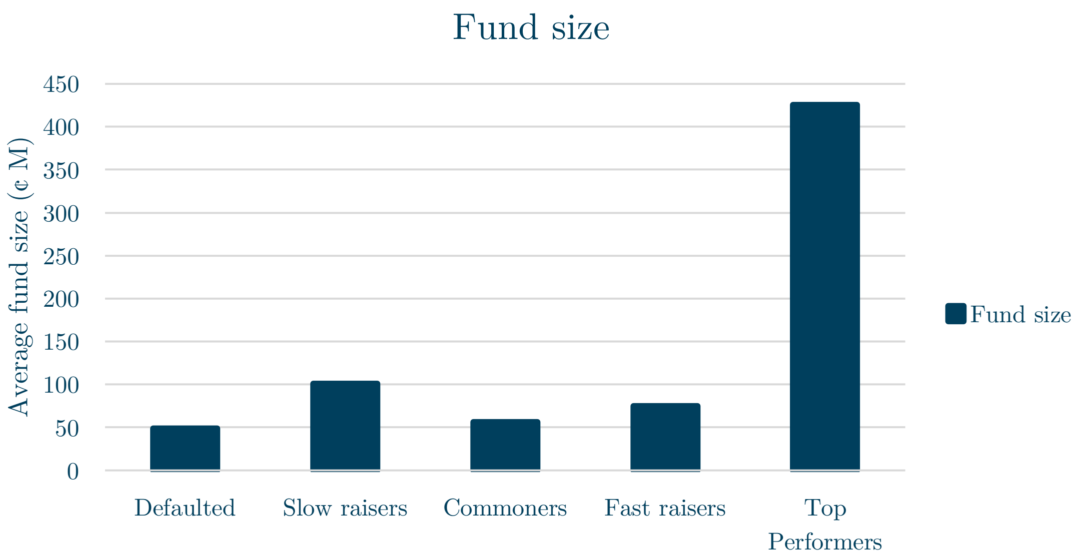


Figure 20: Fund size per cluster.

4.4 Nordic differences

In the Nordic region, we notice distributed allocations of VC firms across clusters. Figure 21 shows that both Defaulted, Slow-raisers and Top Performers has most VC firms located in Sweden with over 35 %. However, Commoners and Fast-raisers have only an allocation of VC firms in Sweden at respectively 13 % and 15 %. We observe that Fast-raisers has an allocation of 47 % of VC firms in Norway, whereas the other clusters have an allocation of approximately 30 %. The allocation of VC firms in Finland is on the other hand quite diverse.

Commoners has the highest allocation in Finland with 39 %, while Slow-raisers is reported with 10 % and the other three clusters in between around 22 %. Denmark is the total lowest allocation across all clusters. We observe that Defaulted and Top Performers have the lowest allocation of VC firms in Denmark with approximately 7 %, whereas Slow-raisers, Commoners and Fast-raisers have an allocation of approximately 15 %.

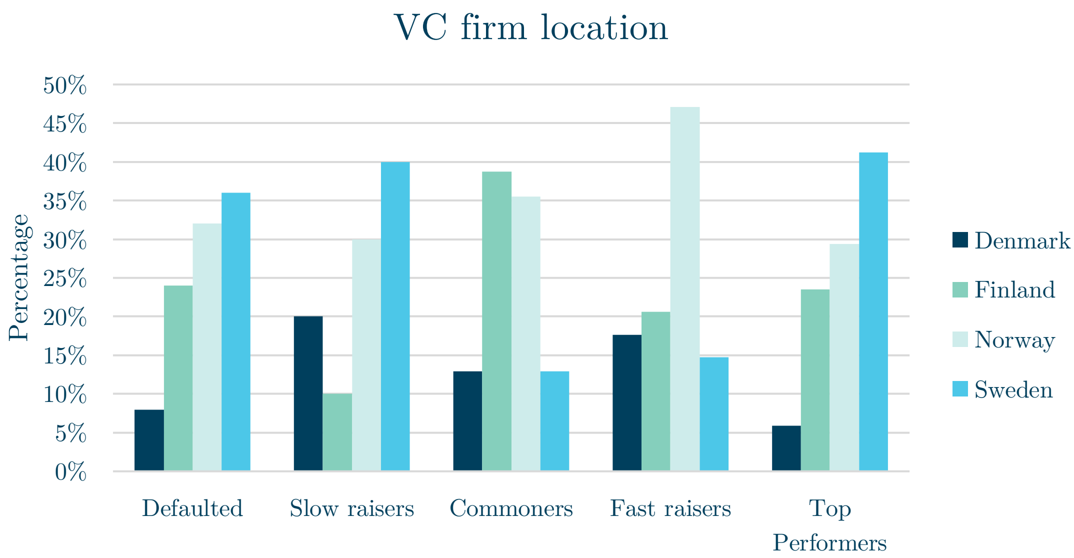


Figure 21: VC firm location per cluster.

5 Discussion

The scope of this study has been to measure and understand venture capital fund performance. The authors' analysis reveals both findings that confirm, contradict, and extend the existing literature regarding the research question(s):

- RQ1: What is the current state of the Nordic venture capital industry?
- RQ2: Is there an alternative proxy for measuring venture capital fund performance?
- RQ3: How can we understand venture capital fund performance in the Nordics by considering structural factors?

Throughout this chapter, themes from the findings will be discussed in view of the theory presented in chapter 2. The chapter presents a discussion of four aggregated themes: (1) Measuring venture capital fund performance, (2) The development of the Nordic venture capital industry, (3) Understanding structural characteristics, and (4) Persistency of performance.

5.1 The Journey of The Nordic Venture Capital Industry

When considering the development of the Nordic venture capital industry, previous research has focused on the aggregated fund raising in the region, in addition to number of investments made within the countries. In order to discuss the findings in context of previous literature, we apply Cetindamar's (2003) three-dimensional method to evaluate VC industry maturity: size, diversity, and competence.

Size

Cetindamar (2003) claims that enough VC capital to fulfil its role as a supporting industry is vital for mature VC industries. Data sources used in previous activity reports and research (NVCA, 2021; SVCA, 2020; FVCA, 2020; Aktive Ejere, 2019) show that Sweden and Denmark have a larger VC market than Finland and Norway during 2007-2019, in terms of funds raised. When studying the investment activity in the Nordic countries, previous activity reports state contradictory numbers on activity levels. Vækstfonden (2019) claims Denmark to be the most active VC market in the Nordics, whereas other research (Copenhagen Economics, 2019; Argentum, 2020) find that the Swedish VC market has considerably more investments than the Nordic peers, with Finland as the second most active market.

Our analysis reports absolute amount raised per country per year, as shown in Figure 11, Sweden outperforms its Nordic neighbours. Sweden raises €38 Bn. in the period 1989-2022, almost twice as much as Norway with its €22 Bn. Both Finland and Denmark are relatively far behind with €7 Bn. and €4 Bn. respectively in raised capital during the period. Notably, we observe an incline increase post 2015 for all countries in the Nordics, especially Norway and Sweden.

As we can see in Figure 11, the incline of raised capital has increased post 2015, which is in line with the general economic development in the Nordic region. This is especially true for Norway and Sweden who also hosts the largest industrial sectors in the Nordics. For the Danish VC industry, our analysis shows clearly deviant results from previous research. As our data basis is expanded from data provided by NVCA and Argentum, two Norwegian organizations, we can assume a possibility of data skewness towards Norway and especially away from Denmark. However, Sweden is presented as undoubtedly the largest VC industry both by our results and previous research. If we assume a data skewness towards Norway, it is inexplicable why our results and previous research on the Swedish VC industry corresponds quite accurately with other papers.

Although it is difficult to conclude how much VC capital is optimal for high performance, it is clear that VC industries with more capital, such as Sweden and Norway, have greater potential to fulfil its role as a supporting industry and hence are more mature (Cetindamar, 2003). We consider the results of the Danish VC industry as difficult to conclude, whereas the Finnish VC industry is perceived as relatively small in Nordic context both in our results and previous research. Our results indicate that Sweden is the most mature VC industry and Finland is the least mature VC industry in the Nordics in terms of size.

Diversity

Mature VC industries require a great deal of diversity among the VC firms, especially concerning stage and industry specialization (Cetindamar, 2003). In this way, the mature VC industry is capable of supporting companies in a broad range of stages and industries. Previous data sources (Invest Europe, 2021) show that later-stage VC funds is underrepresented in the Nordic VC industry, whereas Finland and Denmark have performed most investments towards early-stage companies. Strong academic and research communities in the Nordics lead to most investments towards ICT and life science in Europe (Copenhagen Economics, 2019). Additionally, Denmark and Norway are further registered with a strong investment focus towards respectively life science and energy (Invest Europe, 2021).

Our findings, reported in Figure 12, show that later-stage VC funds are underrepresented in the Nordic countries, especially in Denmark. Norway is observed with most funds in all stages compared to Nordic peers and is perceived as the most diverse VC industry in terms of stage diversity. Our results in Figure 13, state clearly that generalist and ICT are the most common industry specializations for VC funds in the Nordics, whereas we observe smaller allocations of VC funds with a specialization towards life science, energy or cleantech. Both Norway, Finland and Sweden could be perceived as diverse in terms of industry specialization, whereas Denmark is perceived as quite uniform towards ICT.

Interestingly, Nordic countries' presence of later-stages VC funds is lower than European peers and much lower than US. Previous literature perceives the Nordic VC market as less mature than in US and Europe, which can be linked to the underrepresented number of later-stage funds. Surprisingly, Norway is reported in our results with a higher number of VC funds in every stage. As previous literature claims Sweden to be at least as diverse as Norway in terms of stage specialization, it clearly exists a gap between our data and previously used data sources. However, Denmark's low presence of later-stage funds aligns with previous literature, and we can determine that the Danish VC industry can be perceived as less mature in terms of stage diversity.

The findings of less maturity in the Danish VC industry continues when we observe the Danish VC funds' industry specialization, where Denmark is reported with far less generalists than the Nordic peers. However, the majority of Danish VC funds have an industry specialization towards ICT and somewhat towards life science. Copenhagen Economics (2019) finds that countries with large ICT and life science sectors have strong VC markets – with US being a prime example. This way, one can argue whether the benefits of specialization towards ICT and life science is more important than overall diversity in industry specialization.

Competence and experience

The Nordic VC industry was born significantly after that of the US, during the late 1970's and early 1980's. Previous literature on the experience of Nordic VC firms is sparse and reports by the national venture capital associations, NVCA, SVCA, FVCA, and Aktive Ejere, lack the number of VCs operating within the industry. However, few researchers (Hyytinen & Pajarinen, 2001; Christensen, 2000; Karaömerlioglu, 2000) study the formation and first years of the Nordic VC industries. In the 1980s the VC industry began to grow as several new private VC firms were founded. By the mid-1980s, there were about 20 VC firms in Denmark, 5-6 in Norway and some 20 private VC firms in Sweden, accompanied by around 30 regional and government run investment companies (Christensen, 2000; Karaömerlioglu & Jacobsson, 2000). In Finland, the growth lagged a bit the other Nordic countries.

By aggregating our results for the period 1989-2022, presented in Figure 14, we observe that Norway outperforms the rest with 166 funds raised, Sweden and Finland follow with 102 and 103 funds, and Denmark is lagging with 45 funds raised in the period from 1989 to 2022. Additionally, it is observed solid increase in the Norwegian and Finnish VC funds in recent years.

Although the Swedish VC industry was born earlier than the Norwegian VC industry, our results show relatively growth in the Norwegian VC industry during the last twenty years compared with Nordic peers. During this growth period, it is reasonable to assume that Norwegian VC firms and the overall industry have gained competence and experience. Notably, we acknowledge that our analysis is based on limited data before year 2000 and hence we can not interpret our results before 2000 accurately. By considering our results, one can argue that Norway and Sweden possess the most experienced VC industry in the Nordics, whereas Finland and Denmark can be interpreted as the least experienced VC industry.

5.2 Is it Possible to Create a Proxy for Measuring Venture Capital Fund Performance?

In this section we will examine the current state of research on performance proxy research. To find an accurate proxy for performance, we need two key components: an accurate direct measurement for performance, and high-quality representative data. Directly measuring the performance of a VC firm is a surprisingly difficult task. In contrast to most other firms where the stock price of the portfolio companies translates quite directly into the firms' performance, no such measure exists for VC firms. The most widely used direct performance measurement is IRR or TVPI (Diller & Kaserer, 2004). In addition to IRR and TVPI, PME is often used to offset increased performance because of purely macro-economic factors.

The issue with IRR and TVPI is that both measurements are calculated differently between firms. This means there is no standardized methodology for calculating performance for VC firms. This already creates certain limitations for measuring performance using a proxy, because the underlying direct measurements might be inaccurate or different between firms.

5.2.1 Alternate Measures of Performance

Previous literature points out three axes to determine performance by looking at alternative measures. (1) The ability to raise the following fund in a short amount of time, (2) the ability to increase the following fund size and (3) the number of funds raised. All three of the alternative measures has its strengths and limitations.

The ability to raise the following fund in a short amount of time is dependent on several structural characteristics, e.g., stage, industry, age, business model etc. Talent incubators, such as Antler or Katapult, are able to raise funds faster than traditional funds as they utilize innovative business models which allows them to deploy capital more efficiently. The fastest raising funds does not seem to be the best performing, however. The best performing funds raise a following fund in the range of 1 – 3 years, implying a concave relationship between time to raise and performance. This means there are no Top Performers raising faster than 1 year, and no top performer raising slower than 3 years.

The ability to increase the following fund size could reflect limited partners trust in the firm, giving them more capital as a result of good performance in previous funds. Historically, firms that have several unicorns in their portfolio such as Y Combinator, Andreesen Horowitz or Tiger Global has increased their fund size over time as a result of increased investor interest in investing in their funds. The ability to increase the following fund size is heavily dependent on the year the fund was raised, stage focus and industry amongst other factors.

The number of funds closely resemble experience, and smaller firms have shown in our study to raise a new fund faster. A firm with more funds than another firm might not be more experienced, even though the two factors are connected. We have seen in our study that a higher number of funds differentiates the Visionaries from the Commoners. This might be either because of our utilized proxy for performance, or because these funds perform better.

5.2.2 Data Quality

The majority of existing literature concerning venture capital fund performance is based on data provided by principal providers such as VentureXpert, Venture Economics, Reuters Thomson and Prequin (Ljungqvist & Richardsson, 2003; Lerner et al., 2011; Arundale, 2018). The data has been subject for critique as it is only available in the aggregate, rather than on a fund-by-fund basis and is provided on a voluntary basis and thus disposed to selection biases (Kaplan & Lerner, 2014). Of existing research, 32 of 68 articles have used either VentureXpert or Venture Economics as their data-basis, and these databases have been found to be especially inaccurate as discussed in section 2.3.3. Existing research on performance proxies should be read keeping the limitations of the underlying data in mind.

To take the criticism into account, scholars such as Arundale (2018) adapted alternative research methods allowing first-hand data gathering to ensure higher data quality. However, alternative approaches are often constrained by small sample sizes, and the inability to verify that the sample is representative for the entire population. If it were possible to verify that a sample is representative for the entire population, this would make a conclusion on the accuracy of a proxy easier to confirm.

In order to fill the current research gap, we believe that it is possible to take advantage of the governance movement driving increased transparency and openness in the financial sector. We find in our research efforts that it is possible to expand the current state-of-the-art dataset on venture capital funds in the Nordics by a 3x, simply by trawling official information posted on the firm's website. However, we have not been able to gather financial data such as estimated fair market value, exit rates, return on capital or similar. If we did, it would be fair to assume that the data would be skewed due to selection biases from the funds, unless we had data from the full population. This in turn, implies that we are not able to determine the performance by using financial data alone. To bridge the gap, the research community could create an open-source database of VC firms, including free and open data concerning VC funds, fund size, regions invested in, investment stage and investment sector.

5.2.3 Generalizing a Proxy for Performance

Even if researchers manage to create a proxy for performance, there are limitations that should be kept in mind. Proposed proxies such as exit-rates, experience of the investors and increase in subsequent funds might only be applicable in certain contexts. As we have seen, different firms with different strategies might both reach high performance with very different fund behaviors. Firms, such as Antler, raise small funds and invest fast. Later-stage firms typically raise slower and bigger funds. Even sectors matter a lot, as the portfolio companies grow and exit at very different rates for ICT funds and life-science funds as an example.

Of the 68 papers concerning performance determinants used in this study, 44 are from 2008 and earlier. This might imply that even if previous research found an accurate proxy, this research might not be applicable as market conditions have changed sufficiently to render the proxy useless in a modern context. VC fund performance is ever changing, and therefore the recency of studies should be considered when applying previous research.

Because of changing market conditions, differences in ecosystems and differences between firms, finding a general proxy for performance will be difficult. Understanding why a proxy mimics performance in a certain context will therefore be of high importance. A previously accurate proxy might not stand the test of time, as market conditions have changed towards new sectors. Therefore, researchers could rather try to use a proxy to explain the behavior of smaller populations of funds than most researchers have attempted. Making a proxy for the entirety of Europe might be difficult but creating one for a Nordic context might be possible. Even in a Nordic context it might not be possible to find a uniform proxy, as we have seen that the Nordic VC contexts are in different stages of maturity and have different sector focuses. Limiting a proxy to only be applicable for funds in each country or in a given investment stage might therefore be necessary.

5.3 Structural Characteristics

In this section, statistics of structural characteristics influencing VC performance is discussed according to the performance clusters. VC location statistics is concerned in comparison to the structural characteristics to obtain an understanding of internal Nordic differences.

5.3.1 Geographical Focus; is Denmark Lagging Behind?

The Nordic countries are among the most developed economies in the world and a well-educated and skilled population have a natural tendency to focus on innovation and international growth. According to the European Commission's Innovation Union Scoreboard 2021, Denmark, Finland and Sweden are three of the most innovative countries in the world. In addition, the Nordic business sectors score high in different rankings of innovation activity. In the 2018 Global

Entrepreneurship Index ranking, the Nordic countries all rank in the top 25 globally. Denmark is ranked as the sixth most entrepreneurial economy globally, Sweden as number 9, Finland as number 12 and Norway as number 21. There is substantial investment activity between all the Nordic countries. Argentum (2020) reports that Swedish firms account for 41 % of investments in the Nordic venture capital segment, followed by Finnish firms who made 26 % of transactions. Further on, 19 % of investments were made by Danish firms, and 14 % by Norwegian ones.

We observe from our data in Figure 16 that Danish firms primarily fall in the least-performing half of the clusters, such as Defaulted, Laggers and Commoners. On the other side, we clearly see a trend towards both Finland and Norway in the upper half. Across all clusters, the pattern of allocation towards Sweden is more uniform.

The reported results in Figure 16 shows a substantially lower allocation of VCs with a geographic focus towards Denmark in Fast-raisers and Top Performers. As previous research suggests Denmark is the most entrepreneurial economy in the Nordics, these results appear slightly counterintuitive. If we compare these results with the location of VCs, reported in Figure 21, we observe that Denmark has a higher number of VCs in the Fast-raisers cluster than Sweden, whereas Top Performers has a very low allocation of VCs located in Denmark. A plausible explanation is that several VCs located in Denmark invest abroad in other Nordic countries. For instance, Argentum (2020) reports that Sweden and Finland attract more venture investments from other Nordic countries than they do from all other international venture capital funds.

5.3.2 Industry Focus; Higher Risk for Specialization

Most VC firms are specialized in one or more industries. Lockett et al. (2002) argues industry specialization enables the firms to develop their industry-specific capabilities and experience. The chaotic uncertainty and opacity of an emerging technology market may be too high a barrier for firms competing with the most specialized investors (Lockett et al., 2002). In contrast, Aigner et al. (2008) find no advantage of specialization for firms. Risk reduction is important for VC firms to obtain persistent performance. Firms can affect their risk exposure by choosing an appropriate set of industries (which also have their own level of risk) (Buchner & Schwienbacher, 2017). For example, expected risk exposure is likely higher when allocating a larger fraction of funds to a single industry. Industry diversification is usually seen as a mean to reduce and manage risks. Prior research by Humphery-Jenner (2012) documents a positive relationship between industry diversification and performance. In contrast, Cressy et al. (2014) find that industry diversification reduces fund performance. This corresponds with several researchers that encourage VC firms to specialize in industry to reduce risk (Murray & Marriott, 1998; Wang & Ang, 2004).

Our analysis, reported in Figure 17, shows a majority allocation towards both generalist and ICT across all clusters. We observe the highest allocation towards specialized industries (excl. generalists and ICT) among Top Performers and Defaulted compared to peers. These findings align with the results of Humphery-Jenner (2012), which implies higher risk for industry-specialized VCs. A well-established theory in modern finance, Modern Portfolio Theory (Markowitz, 1952), proves the relationship between risk and return – hence higher return implies higher risk. VCs specialized in a single industry has limited investment opportunities, experience the highest risk and therefore most defaults and simultaneously best performance. Additionally, single-industry VCs are more exposed for industry-specific events, both positive and negative, which increases risk of investment.

5.3.3 Stage Focus; Later-stage Funds are Underrepresented

In VC financing, different investment stages have very different risk and return profiles. Early-stage investments, such as investments in seed or venture companies, are associated with higher levels of downside risk and higher upside potential, whereas later-stage investments are associated with lower downside risk and moderate upside potential (Buchner & Schwienbacher, 2017). In contrast with European peers, later-stage VC investments are in general underrepresented in the Nordics, except in Sweden (SVCA, 2020).

We observe in our results, presented in Figure 18, a general skewness toward early-stage investments, which aligns with SVCA's (2020) report on the Nordic VC market. As early-stage investments are associated with higher levels of downside risk, it is reasonable that the cluster Defaulted has the highest allocation of VC firms with a focus toward the seed stage. However, in contrast with previous literature (Buchner & Schwienbacher, 2017), we observe a constant level of allocation of VC firms with a focus toward the seed stage in all clusters above Defaulted. The allocation of venture stage investments, also perceived as early-stage, are high in Bottom Performers and Mid Performers, whereas less in Top Performers. In line with Buchner & Schwienbacher (2017), we observe an increasing allocation towards later-stages, such as growth and buyout, for better performing VC firms. However, overally, we observe minor allocations towards later stages, which aligns with SVCA's numbers on later-stage investments.

As a consequence of the Nordic region's overall lack of focus on later-stage VC investing, many successful Nordic later-stage growth companies must find funding abroad. We observe a large number of top-performing VC's with later stage funds, which might imply that later stage investments in the Nordic have the best return potential. However, it is unclear whether the Nordic region lacks later-stage growth companies and as a consequence lack VCs with later-stage focus, or if the Nordic region lacks enough capital inflow for later stage investing in general. Both cases will result in a lack of VCs with a focus on later-stage growth companies. Furthermore, this could potentially lead to a challenge for VCs in the Nordic region to catch up with high-performing later-stage VCs abroad, for example in central Europe or the US.

5.3.4 Fund Size; Is Bigger Better?

In finance, any restrictions limiting investment opportunities would generally bring reduced performance in terms of risk-adjusted return. In Nordic VC financing, is it the case that larger funds would bring more opportunities and hence better performance? Maula et al. (2006) recommend that funds should have a sufficient fund size to ensure the possibility of follow-on investments to avoid excessive dilution in later rounds and to diversify the portfolio. According to the findings of the research, there may be an ideal, medium-sized fund for improved fund performance. Lerner et al. (2011) finds that medium-sized funds outperform large funds. Kaplan and Schoar (2005) confirm this by suggesting a positive concave relation between fund size and performance. While larger funds have higher public market equivalents, when funds become very large, performance declines. Several previous researchers classify fund size according to Lerner et al. (2011), hence funds with assets under management below $84 million can be categorized as small funds, $84-$364 million as medium-sized and above $365 million as large funds.

Our analysis, presented in Figure 20, reports a considerable higher mean fund size by Top Performers than any peers. However, the standard deviation of fund size of Top Performers is also the proportionately highest. Fast-raisers and Commoners reports a mean fund size respectively of € 75 M and € 56 M, with a corresponding standard deviation of 87 and 68. On the same track, we observe the mean fund size of Slow-raisers and Defaulted at € 101 M and € 49 M. Slow-raisers is observed with the lowest proportionately standard deviation of 59, whereas Defaulted has a standard deviation of 74.

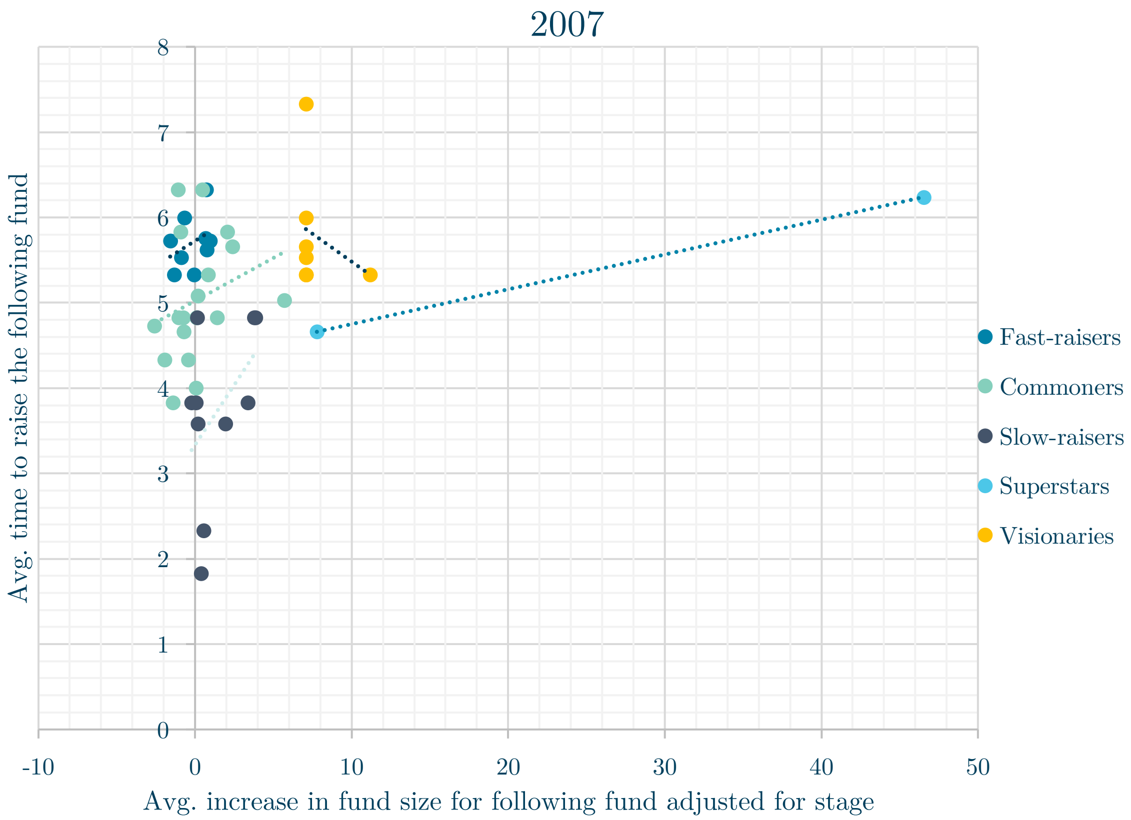
We observe a large difference in mean fund size between Top Performers and any other cluster. It is clear from our analysis that Top Performers have large funds relative to our data. However, when we look at the classifications provided by Lerner et al. (2011), we see that the average for Top Performers imply a fund size of € 426 M, slightly above the medium-sized funds. In combination with a relatively high standard deviation of Top Performers, it is reasonably to say that a great share of Top Performers operates medium-sized funds. In this way, our results are in line with previous research (Kaplan and Schoar, 2005).

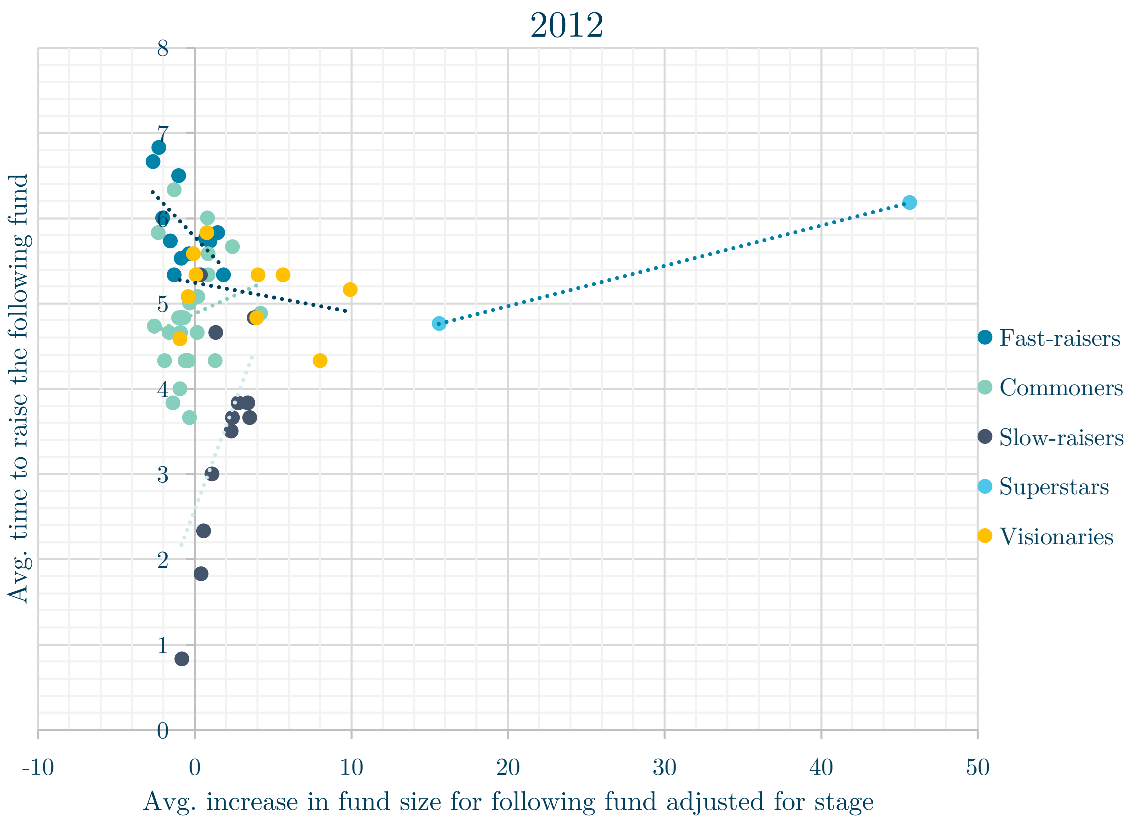
All less-performing clusters than Top Performers have a mean fund size classified as small or slightly medium, according to Lerner et al. (2011). SVB Capital (2010) describes smaller fund sizes, and hence fewer investments force managers to focus on capital efficiency. Plausibly, benefits of operating a large fund does not appear until the firm reaches a certain fund size. This would substantiate our results to why we observe negligeable difference in mean fund size between Defaulted, Slow-raisers, Commoners and Fast-raisers.

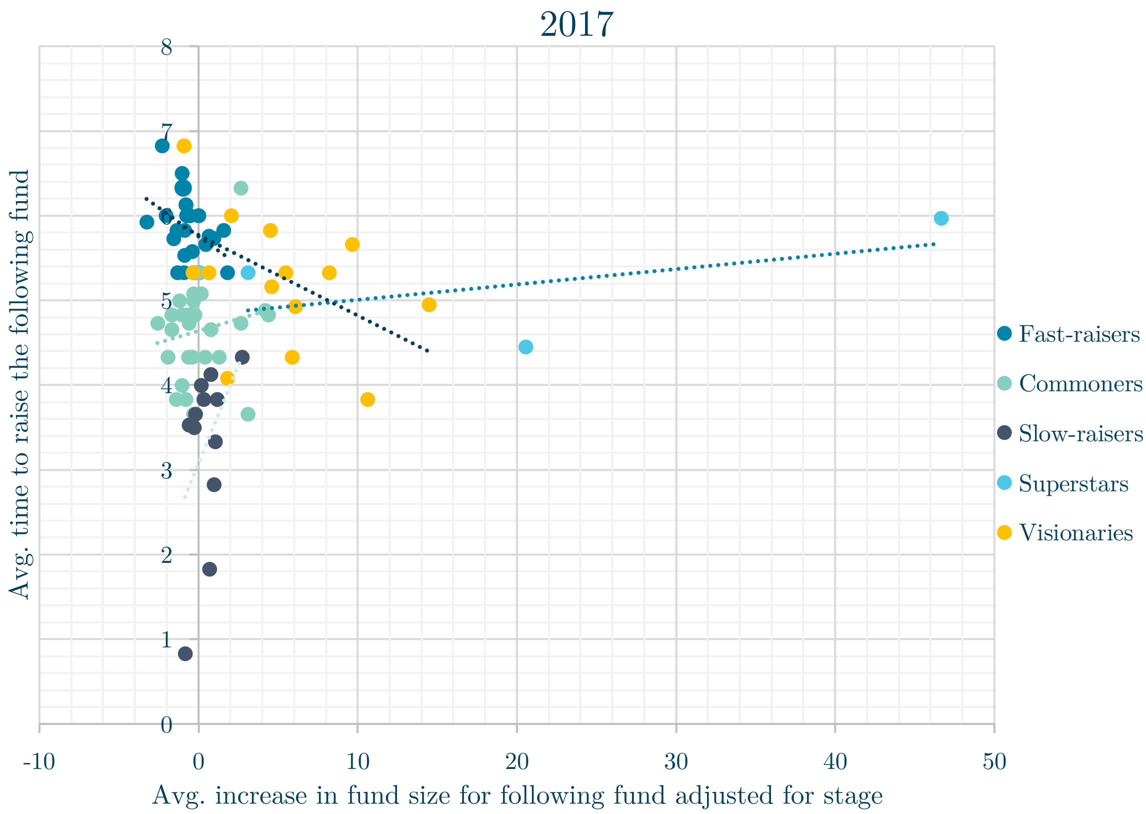
5.4 How has the Nordic VC Ecosystem Evolved Over Time?

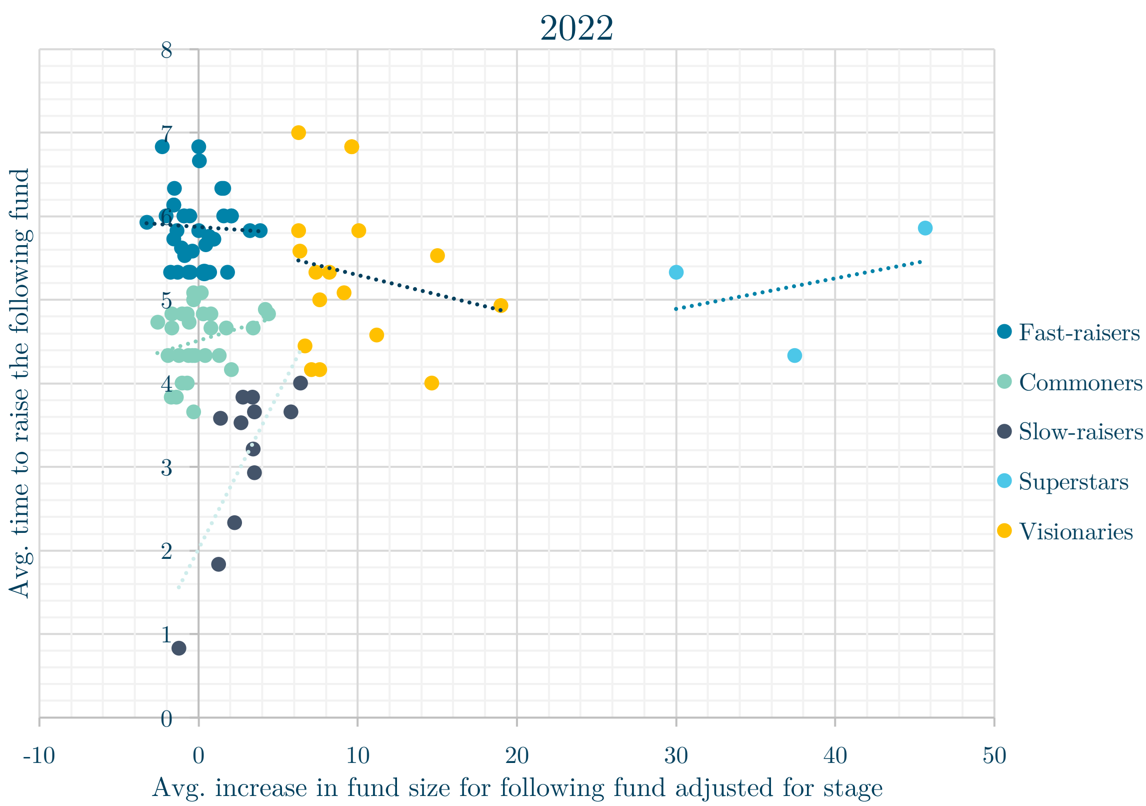
To understand the current context of the Nordic VC ecosystem, we should inspect how it has evolved over time. By doing so we might hope to uncover whether the current performance measurements are seemingly random or having evolved with an apparent trend to reach its current state.

On the following pages is the performance state of the Nordic VC ecosystem plotted in 5-year intervals between 2007 and 2022. The firms are plotted with their most recent fund in the given year, but the datapoints are colored with their 2022-clustering. Thus, we aim to show how the ecosystem of Nordic VC funds have evolved over time to become what it is today. In the coming sections we will explore whether performance is persistent, and whether it was possible to predict today's performance in 2007.


2007
Avg. time to raise the following fund
Avg. increase in fund size for following fund adjusted for stage
Fast-raisers
Commoners
Slow-raisers
Superstars
Visionaries


2012
Avg. time to raise the following fund
Avg. increase in fund size for following fund adjusted for stage
Fast-raisers
Commoners
Slow-raisers
Superstars
Visionaries


2017
Avg. time to raise the following fund
8
7
5
4
3
2
1
0
-10
0
10
20
30
40
50
Avg. increase in fund size for following fund adjusted for stage
Fast-raisers
Commoners
Slow-raisers
Superstars
Visionaries


2022
Avg. time to raise the following fund
8
7
5
4
3
2
1
0
-10
0
10
20
30
40
50
Avg. increase in fund size for following fund adjusted for stage
Fast-raisers
Commoners
Slow-raisers
Superstars
Visionaries

5.4.1 Is Performance Persistent?

Kaplan & Schoar (2005) found that venture capital and buyout funds show persistence in performance. In their study, high-performing VC funds were shown to continue to outperform the market average in subsequent funds. This outperformance is unique in the segment of capital management and is not found in hedge-funds and other stock market instruments. Nanda et al. (2020) also found performance to be persistent despite significant mean reversion. Their study concludes that performance is persistent, even though an overperforming initial fund likely will perform worse over time while still staying above the market average.

In our study, we observe some of the tendencies uncovered by previous studies. We observe that Superstars, especially one single VC firm, perform well over the average already in 2007, whereas Slow-raisers are tightly clustered by this point. Interestingly, Fast-raisers, Commoners and Visionaries are approximately equivalent performing in 2007. In 2017 we are starting to see Visionaries significantly outperforming Commoners, and in 2022 this pattern is strengthened. Further, we observe in 2022 that the Superstars that previously was observed together with the majority of VC firms, are now considerably more high performing. However, both Fast-raisers and Commoners have during the period of investigation roughly experienced a standstill in terms of performance. Another interesting finding is that performance is significantly expanding in the overall Nordic VC industry, and this growth has accelerated in the most recent years. This might indicate that the ecosystem as a whole has become more mature and larger over time and reached a state where there exists a significant performance difference between the different firms.

Also of interest is the fact that all high-performing VCs raise a subsequent fund fast; all high-performing VCs use less than 3 years on average. Only two high-performing VCs raise faster than a year on average. Thus, it seems the VC firms that want to become high-performing should aim to raise a new fund every 1 – 3 years. Noticeably, there is a slight trend for VCs to converge into this range of raising every 1 – 3 years in more recent years.

A notable finding is that outperformance is increasing over time for some VC firms, contrary to the findings of Nanda et al. (2020). There are many factors which might lead to the difference in results. Firstly, the measurement of performance used in our study is dissimilar. Nanda et al. (2020) use IPO rate as a measurement of performance, while our study uses fund growth and fund averages. Another reason might be the difference in the data. Nanda et al. (2020) use VentureXpert as their database, while we have compiled a new database of the Nordics. In addition, the two studies look at a different population of funds, and there might be a different tendency for performance persistence across different regions.

5.4.2 Commoners vs Visionaries

Of very notable interest is the fact that Commoners and Visionaries are clustered together in 2007, however, after 2017 we observe that Visionaries significantly outperform Commoners. This indicates that low-performing VCs persistently stay low-performing, shown by how they are already tightly clustered together already in 2007 and 2012. High-performing VC firms, however, are made over time. Thus, it is of interest to inspect what the structural difference between these clusters are, and whether we can get some hints regarding their gap in performance.

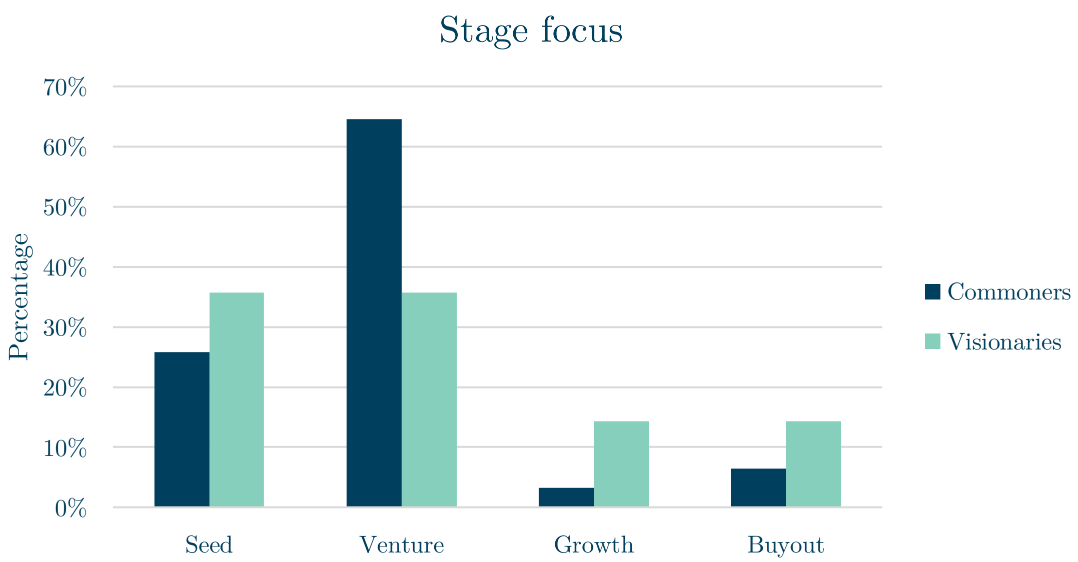


Figure 22: Stage focus of Commoners and Visionaries

We observe the difference in stage focus between Commoners and Visionaries in Figure 22. Clearly, Visionaries are more represented towards later-stages, whereas Commoners invest more in early-stages, especially the venture stage. Buchner & Schwienbacher (2017) state that early-stage investments are associated with higher levels of downside risk and higher upside potential, whereas later-stage investments are associated with lower downside risk and moderate upside potential. Higher levels of downside risk might have resulted in lower performance for Commoners than Visionaries. However, we observe that Visionaries are more allocated towards seed stage, the earliest stage and hence higher downside risk than Commoners. Intuitively, this would have resulted in lower performance for Visionaries, but the relatively high allocation towards both growth and buyout stage might compensate for the high seed stage focus.

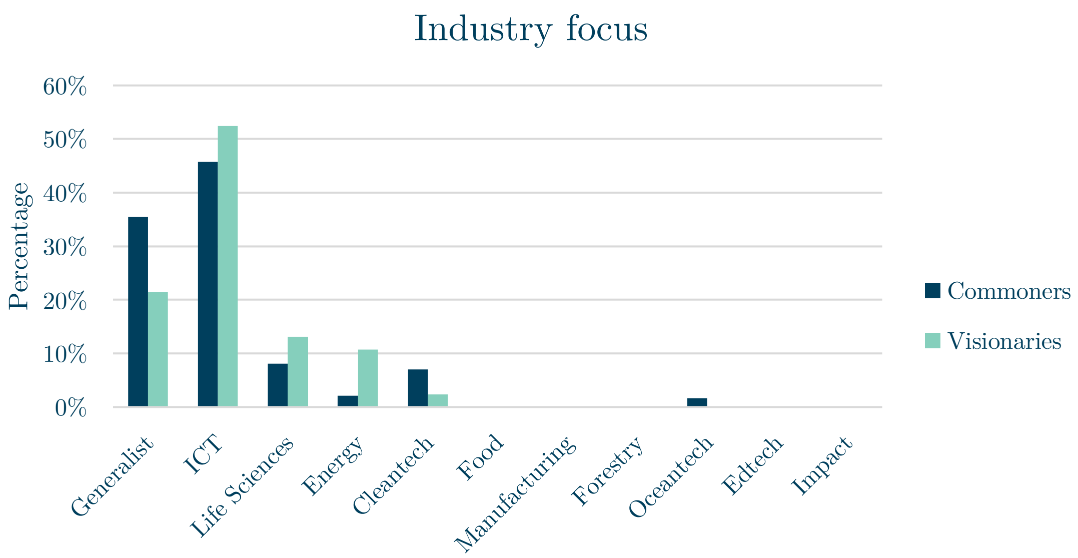


Figure 23: Industry focus of Commoners and Visionaries

Figure 23 presents the difference in industry focus between the clusters Commoners and Visionaries. Both clusters are seemingly focused toward the industries Generalist and ICT, however, Commoners are more focused towards Generalist than Visionaries. As previous mentioned, expected risk exposure is likely higher when allocating a larger fraction of funds to a single industry. Industry diversification is usually seen as a mean to reduce and manage risks

(Humphery-Jenner, 2012). However, also several previous research (Murray & Marriott, 1998; Wang & Ang, 2004) encourage VCs to specialize in industry to reduce risk. As Visionaries perform better and are more specialized than Commoners, this might illuminate that industry specialization versus industry diversification do not significantly impact performance. If so, it would assume that industry specialization would lead to higher performance.

Inspecting the differences between stage and industry for Commoners and Visionaries does not give us a clear causal relationship to their difference in performance. However, looking at number of funds raised imply that higher performing VCs have raised and operated more funds that lower performing VCs. On average, Commoners has raised 4.5 funds, whereas Visionaries has raised 8.1 funds. Number of funds raised could be interpreted as a proxy for experience of the VCs. These results are in line with Kaplan (2005), who found that experience leads to high performance.

6 Conclusion

The purpose of this study was to measure and understand venture capital fund performance in the Nordics. Previous literature in the field of venture capital has identified structural factors that influence firms' performance. Still, few scholarly works have been able to investigate the influence of structural factors as the data availability and quality is unsatisfactory.

This study provides a proxy for measuring venture capital fund performance without concerning data criticized for its shortcomings in previous literature. The proxy is based on two proprietary datasets, which has been expanded by a threefold by trawling public information on firms' website. Furthermore, the study sheds light on the industry's development in the Nordics, by describing performance-based clusters in terms of structural factors. The findings provide practical value for firms, limited partners, and startups. Additionally, it confirms, contradict, and extend existing literature in a new context (the Nordics) and has provided the field with quantitative data for further research and analysis.

6.1 Implications for Practitioners

The Nordic region has a strong industrial position in the high-tech industries, that globally thrive on the back of venture capital investments, and top rankings in international surveys of competitiveness. When studying the journey of the venture capital industry in the Nordics, previous development has been colored by the development in the general economy. Our findings for post 2015, which has been the scope of this study, indicate that the low interest environment in combination with expansive fiscal policies due to Covid-19, stimulated tech stocks to rise, which in turn led to an increased willingness to invest in small cap and venture in order to find attractive investments. As of 2022, the economic environment is challenging, and we predict that the venture capital investments will decline as a result of risk aversion from institutional investors and slumbering funding activity.

Benchmarking against peers is well-known for measuring and evaluating performance (Ramón et al., 2020). **Best practice benchmarking has**

previously not been available for the venture capital industry, and we believe our approach, proposing a proxy based on alternative metrics allows firms to gain increased insight regarding the characteristics of the most successful VC firms. This in turn should be part of the basis for developing plans on how to make strategy improvements or adapt specific best practices, with the aim of increasing financial returns for its investors.

Later-stage venture capital funds with sufficient international scale and experience, deliver the best returns. As previous literature is consistent in its statement that mid-size funds outperform their smaller and larger peers, our findings indicate that this is either not the case for the Nordics or extends the current literature as mid-size funds in the US and UK are in absolute measures, equal to the large-size funds in the Nordics. This might imply there exists an attractive growth opportunity for the existing Nordic large cap-funds to grow even bigger and serve the need for late-stage funding. Currently, this need is typically filled by Asian, central European or US VCs investing internationally such as the Softbank Vision Fund, Accel and Sequoia.

Sector specialization offers increased risk and returns. Both defaulted firms and top performers lead in sector specialization. As the Nordics is small, we believe that the leading firms specialized in a sector can attract the best investment opportunities as well as having the easiest access to capital from limited partners. On the other hand, firms that are not tier 1, is forced to invest in less favorable opportunities and struggle attracting capital from limited partners. In other words, the competition is perceived to be highly concentrated for the sector specialized funds. This does not apply to ICT.

Denmark is lagging as a result of specialization reflecting national comparative advantage in the life science sector, in addition to raising few large-scale funds. The life science sector is known for having longer investment cycles, i.e., demands more time to realize investment opportunity, making Danish funds suffer on our performance proxy. When looking at its Nordic peers, Norway is benefiting on its leading position within the energy sector, the most value creating sector in the Nordics by far. Sweden and Finland on the other hand has developed strong tech and ICT hubs, allowing funds to access investment opportunities with great scalability.

6.2 Implications for Further Research

While several observations have shown indications of consistency with existing literature on venture capital fund performance, some deviations and ambiguities became present during the analysis, and imply a need for further research.

A standard for measuring direct performance must be established. We argue that to properly measure directly performance, researchers and hopefully practitioners should aim to agree upon a set standard for calculating IRR and TVPI, because not doing so will lead to innate differences in reporting of performance from different firms. Thus, quantifying the performance of firms with their current reporting standards is difficult. Even further increasing the difficulty of directly measuring performance is the fact that researchers rarely if ever can verify the accuracy of reported IRR and TVPI because this is private data for the VC firms.

Alternative measures not encompassing direct performance apply differently based upon the business model and strategy of the individual firms. Therefore, we argue these alternative measures should either be applied based on the business model of the VC firm or should be considered another proxy for performance instead of a direct measurement.

Data quality remains a barrier for conclusive research as low-quality datasets still have a huge presence in VC research today. In this study we have aimed to address data-quality issues by manually creating a higher quality dataset for a smaller region by increasing the original number of firms by 42 % and the amount of funds by 104 %. This has however not been without its issues, as many firms report different numbers in different contexts. Therefore, there is a need for an even higher quality dataset to be able to conclusively argue what drives performance. Data integrity is also an issue, as there is no way to confirm the numbers reported by the venture capitalists in a meaningful way. We believe the research community should aim to create an open-source database of VC firms allowing the continuous improvement of the data-basis for research. As data concerning funds, fund sizes, regions invested in, investment stage and investment sector are public and free, this information could be used to create a full-scale database of investors.

Inferential statistics with a representative quantity of high-quality data can let us accurately estimate performance of the Nordic ecosystem. With a complete open-source database, the research community would be able to effectively group or cluster VC firms on other metrics than performance. Such a database would make it possible to ascertain whether a sample-database with more granular hard to collect non-public data such as IRR or TVPI is representative of the entire population of funds, and thus make it possible to quantify performance. Before we have increased the data quality underlying research significantly, it is hard to tell if a good proxy for performance even exists.

Whether proxies are valid for a broader region needs to be further investigated. The findings in this study suggest that there is a strong correlation between the performance proxy and alternative measures such as the number of funds raised. The existing literature in the field of venture capital has predominantly been studying financial returns which offers several pitfalls. In order to bridge the gap between the research conducted in this study and the majority of previous research of venture capital fund performance, the authors thus propose that further research look into the correlation between the proposed proxy and financial returns for a broader region, e.g., US, UK or EU. As the proposed study demands data on fund level, it also implies that further research demands substantial data gathering from publicly available information.

Capital inflow might impact the ecosystem and change the dynamics within the national VC industries in the Nordics. In this study, we have studied the Nordic VC industry separately, without considering capital inflow from other sources, such as investments from international VCs or companies' opportunities to obtain capital from credit institutions or public stock exchanges. An expanded study on these mechanisms could generate insight of the implications on the ecosystem for early-stage companies, the maturity of the VC industries and the dynamics of VC performance.

The Nordic context is underexplored and demands further research to understand how the venture capital industry is developing in comparison to the broader European economies. A finding throughout our study of the meagre existing research on the Nordic VC industry is the contradictory numbers - even when considering the most trivial metrics. We experience that the handful of studies report numbers with variations in the magnitude of a tenfold on the same metric for the same period. In line with the concerns expressed by Lerner and Kaplan (2017), the handful of studies concerning the Nordic VC industry does neither publish a sufficient description of the applied methodology, nor publish the underlying data. We propose that further research investigate the Nordic venture capital industry by both considering the development of startups, the venture capital industry, and the broader entrepreneurial ecosystem. Additionally, we strongly encourage future research to promote transparency by emphasizing the importance of research methodology and making the underlying datasets publicly available.

7 Limitations

While the authors have gone to great lengths to ensure the robustness of their results in this thesis, some limitations to the study should be considered.

Estimated data coverage of 80 % which could cause skewed results. The data coverage is based on the proprietary datasets from Argentum and NVCA, in addition to the independent data gathering, resulting in fund-level data on 421 funds. By considering top-down analysis published by SSB, SCB, DST and STAT we estimate that our data covers approximately 80 % of the funds raised in the Nordics. Especially early-stage seed funds and buyout funds are thought to be underrepresented in our study, due to low publicity for early-stage seed funds making it easier for them to go under the radar, and a higher level of secrecy for the buyout funds.

Impact of other factors on performance. Whilst this study has been investigating structural factors impacting the performance of venture capital funds in the Nordic, there are other factors such as operational and wider environmental factors, that have not been investigated. These include the investment strategy, deal sourcing and the due diligence process, or unemployment rates and gross domestic product trends. Such factors could be investigated in further studies in terms of performance of venture capital funds. In our experience, information regarding operational factors is seldom publicly available, whilst wider environmental factors are more of a low-hanging fruit as the data is available for all Nordic countries.

Limitations regarding the framework. Despite the limitations described above, this study contributes with innovative insights into Nordic venture capital fund performance. Applying the framework for limitations of quantitative research by Almeida et al. (2017), the complexity of the employed techniques is a limitation of the framework used. To reduce the impact of the limitation, we have created a notebook containing a description of the methodology, in addition to the code itself (Attachment B). Furthermore, the specific quantitative analysis applied in this study requires the use of statistical software or knowledge of programming, e.g., use of Python or similar. When considering the framework develop by Almeida et al. (2017), another limitation that appears is that quantitative research often demands a higher degree of structure in theoretical framework and hypothesis compared to qualitative analysis. The same is valid for the flexibility and exploration of the analysis. As there is a gap in the current research concerning the venture capital industry in the Nordics, this study was pushed towards a more exploratory analysis than what would be optimal. As an implication to this, we encourage the academic environment to conduct quantitative research on the field of venture capital performance in the Nordics, to enhance the capabilities of quantitative research in the time going forward.

References

Abell, P., & Nisar, T. (2007). Performance Effects of Venture Capital Firm Networks. *Management Decision, Vol. 45 No. 5: 923-936*.

Aigner, P., Albrecht, S., Beyschlag, G., Friederich, T., Kalepky, M., & Zagst, R. (2008). What Drives PE? Analyses of Success Factors for Private Equity Funds. *Journal of Private Equity, 11. 63-85*.

Aktive Ejere. (2019). *Aktive Ejere.* Retrieved from Kapitalfonde - Årsskrift 2018/19: https://docplayer.dk/157713146-Kapitalfonde-aarsskrift-2018-19.html

Almeida, F., Faria, D., & Quirós, A. (2017). Strengths and Limitations of Qualitative and Quantitative Research Methods. *European Journal of Education Studies*, pp. 369-387.

Amit, R., Brander, J., & Zott, C. (1998). Why Do Venture Capital Firms Exist? Theory and Canadian Evidence. *Journal of Business Venturing, 13: 441-466*.

Argentum. (2020). *State of Nordic PE 2019.* Retrieved from Argentum: https://info.argentum.no/stateofnordicPE2019/sec/3#report-top

Arrow, K., Bolin, B., Costanza, R., Dasgupta, P., Folke, C., Holling, C., . . . Pimentel, D. (1995). Economic growth, carrying capacity, and the environment. *Ecological Economics, Volume 15*, pp. 91-95.

Arundale, K. (2017). Investigating the Characteristics of Top Performing Venture Capital Funds in Europe and USA. *ISBE Conference 2017*.

Arundale, K. (2018). Exploring the Difference in Performance Between UK/European Venture Capital Funds and US Venture Capital Funds. *University of Glasgow*.

Avdeitchikova, S. (2012). The geographic organisation of venture capital and business angels, in Handbook of Research on Venture Capital: Volume 2 A Globalizing Industry. *Edward Elgar Publishing*, pp. 175-208.

Axelson, U., & Martinovic, M. (2013). European Venture Capital: Myths and Facts. *London: BVCA*.

Çetindamar, D. (2003). *The growth of venture capital: a cross-cultural comparison.* New York: Praeger.

Barney, J., Busenitz, L., Fiet, J., & Moesel, D. (1996). New venture teams' assessment of learning assistance from venture capital firms. *Journal of Business Venturing*, 11: 257-272.

Bloom, N., Jones, C., Van Reenen, J., & Webb, M. (2020). Are Ideas Getting Harder to Find? *American Economic Review, 110(4)*, pp. 1104-44.

Boote, D., & Beile, P. (2005). Scholars Before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation .

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*.

Bruton, G., & Ahlstrom, D. (2003). An Institutional View of China's Venture Capital Industry - Explaining the Differences Between China and the West. *Journal of Business Venturing, 18: 233-259*.

Buchner, A., & Schwienbacher, A. (2017). Diversification, Risk, and Returns in Venture Capital. *Journal of Business Venturing, 32, 519-535*.

Buchner, A., Espenlaub, S., & Khurshed, A. (2018). Cross-Border Venture Capital Investments: The Impact of Foreignness on Returns. *Journal of International Business Studies, 49(5): 575-604*.

Bygrave, W. (1987). Syndicated Investments by Venture Capital Firms: A Networking Perspective. *Journal of Business Venturing, 2: 139-154*.

Bygrave, W., & Timmons, J. (1992). Venture Capital at the Crossroads. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Canadian Business Growth Fund. (2021, November 18). *CBGF*. Retrieved from CBGF: https://cbgf.com/news-insights/whats-the-difference-between-seed-venture-and-growth-capital/

Cassis, Y., & Minoglou, I. P. (2005). *Entrepreneurship in Theory and History .* London: Palgrave Macmillan London .

Christensen, J. (2000). Effects of Venture Capital on Innovation and Growth. *Aalborg University*.

Copenhagen Economics. (2019). *Copenhagen Economics.* Retrieved from The role of venture capital for economic growth in the Nordic region: https://copenhageneconomics.com/wp-content/uploads/2021/12/rising-north-report.pdf

Corporate Finance Institute. (2022). *Private Equity Funds*. Retrieved from Corporate Finance Institute: https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/private-equity-funds/

Cressy, R., Malipiero, A., & Munari, F. (2014). Does VC Fund Diversification Pay Off? An Empirical Investigation of the Effects of VC Portfolio Diversification on Fund Performance. *International Entrepreneurial Management Journal, 10, 139-163*.

Cumming, D. (2006). Adverse Selection and Capital Structure: Evidence from Venture Capital. *Entrepreneurship Theory and Practice, 30(2): 155-83*.

Cumming, D., & Walz, U. (2010). Private Equity Returns and Disclosure Around the World. *Journal of International Business Studies, 41(4): 727-754*.

Døskeland, T., & Strömberg, P. (2018). *Regjeringen.* Retrieved from EVALUATING INVESTMENTS IN UNLISTED EQUITY FOR THE NORWEGIAN GOVERNMENT PENSION FUND GLOBAL (GPFG): https://www.regjeringen.no/contentassets/7fb88d969ba34ea6a0cd9225b28711a9/evaluating_doskelandstromberg_05042018.pdf

De Clercq, D., Fried, V., Lehtonen, O., & Sapienza, H. (2006). An Entrepreneur's Guide to the Venture Capital Galaxy. *Academy of Management Perspectives Vol. 20*, pp. 90-112.

Diller, C., & Kaserer, C. (2004). European Private Equity Funds: A Cash Flow Based Performance Analysis. *No. 2004-01, CEFS Working Paper Series*.

Diller, C., & Kaserer, C. (2009). What Drives Private Equity Returns? Fund Inflows, Skilled Gaps And/Or Risk? *CEFS Working Paper No. 2002-2*.

Dimov, D., & Shepherd, D. (2005). Human Capital Theory and Venture Capital Firms: Exploring "Home runs" and "Strike outs". *Journal of Business Venturing, 20(1): 1-21*.

Espenlaub, S., Khurshed, A., & Mohamed, A. (2014). Does Cross-Border Syndication Affect Venture Capital Risk and Return? *International Review of Financial Analysis, 31(C): 13-24*.

European Commission. (2010). *Enterprise and Industry.* Retrieved from Access to finance indicators, Business angels: http://ec.europa.eu/enterprise/policies/finance/data/enterprise-finance-index/access-to-finance-indicators/business-angels/index_en.htm#h2-1

EVCA. (2010). Closing Gaps and Moving up a Gear: The Next Stage of Venture Capital's Evolution in Europe. *EVCA Venture Capital White Paper*.

Fleming, G., Schwienbacher, A., & Cumming, D. (2005). Liquidity Risk and Venture Capital Finance. *Financial Management, Vol. 34: 77-105*.

Florida, R., & Kenney, M. (1988). Venture capital, high technology and regional development. *Regional Studies*, pp. 33-48.

Forgy, E. (1965). Cluster Analysis of Multivariate Data: Efficiency versus Interpretability of Classifications. *Biometrics, 21*, pp. 768-780.

Fränti, P., & Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats? *Pattern Recognition*, pp. 95-112.

Fried, V., & Hisrich, R. (1994). Toward a Model of Venture Capital Decision Making. *Financial Management, 23(3): 28-37*.

Fried, V., & Hisrich, R. (1995). The Venture Capitalist: A Relationship Investor. *California Management Review*.

FVCA. (2020). *FVCA.* Retrieved from 2019 Finnish Venture Capital Acitivity: https://paaomasijoittajat.fi/wp-content/uploads/2020/04/Venture-Capital-in-Finland-2019-korjattu-14.4..pdf

Gjedrem, S. (2000). *Developments in the Nordic financial industry - a central banker's perspective.* Retrieved from Norges Bank: https://www.bis.org/review/r000926a.pdf

Gompers, P. A., Gornall, W., Kaplan, S. N., & Ilya A. Strebulaev. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*, 169-190.

Gompers, P. (1996). Grandstanding in the Venture Capital Industry. *Journal of Financial Economics, 42(1): 133-156*.

Gompers, P., & Lerner, J. (2000). Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuations. *Journal of Financial Economics, 55(2): 281-325*.

Gompers, P., & Lerner, J. (2001). The Venture Capital Revolution. *Journal of Economic Perspectuves, 15, 145-168*.

Gompers, P., & Lerner, J. (2004). The Venture Capital Cycle.

Gompers, P., Kovner, A., & Lerner, J. (2009). Specialisation and Success: Evidence from Venture Capital. *Journal of Economics & Management Strategy, 18(3): 817-844*.

Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2008). Venture Capital Investment Cycles: The Impact of Public Markets. *Journal of Financial Economics, 87: 1-23*.

Gompers, P., Kovner, A., Lerner, L., & Scharfstein, D. (2010). Performance Persistence in Entrepreneurship. *Journal of Financial Economics, 96(1): 18-32*.

Gorman, M., & Sahlman, W. (1989). What Do Venture Capitalists do? *Journal of Business Venturing 4(4): 231-48*.

Gornall, W., & Strebulaev, I. (2015). The Economic Impact of Venture Capital: Evidence from Public Companies.

Harris, R., Jenkinson, T., & Kaplan, S. (2013). Private Equity Performance: What Do We Know? . *Journal of Finance*, pp. 11-44.

Harris, R., Jenkinson, T., & Stucke, R. (2010). A White Paper on Private Equity Data and Research. *UAI Foundation Consortium Working Paper*.

Heffner, C. (2017). *Allpsych.* Retrieved from Research Methods: https://allpsych.com/researchmethods/experimentalvalidity/

Hege, U., Palomino, F., & Schwienbacher, A. (2003). Determinants of Venture Capital Performance: Europe and the United States. *LSE Ricafe Working Paper 001*.

Hellmann, T., & Puri, M. (2002). Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence. *Journal of Finance, 57(1): 169-197*.

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*.

Hsu, D. H. (2002). What do entrepreneurs pay for affiliation? *Working paper*.

Humphery-Jenner, M. (2012). Private Equity Fund Size, Investment Size, and Value Creation. *Review of Finance, 16, 799-835*.

Hyytinen, A., & Pajarinen, M. (2001). Financial Systems and Venture Capital in Nordic Countries: A Comparative Study. *Elinkeinoelämän Tutkimuslaitos*.

Invest Europe. (2021). *Activity Data*. Retrieved from Invest Europe: https://www.investeurope.eu/research/activity-data/

Isaksson, A. (2006). Studies on the venture capital process. *Doctoral dissertation*.

Jackson, J. (2002). Data mining; a conceptual overview. *Communications of the Association for Information Systems*.

Jääskeläinen, M., Maula, M., & Seppa, T. (2006). Allocation of Attention to Portfolio Companies and the Performance of Venture Capital Firms. *Entrepreneurship: Theory and Practice, 30*.

Kameshwaran, K., & Malarvizhi, K. (2014). Survey on clustering techniques in data mining. *International Journal of Computer Science and Information Technologies*, pp. 2272-2276.

Kaplan, S., & Lerner, J. (2016). Venture Capital Data: Opportunities and Challenges. *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*.

Kaplan, S., & Shoar, A. (2005). Private Equity Performance: Returns, Persistence, and Capital Flows. *The Journal of Finance*.

Kaplan, S., & Strömberg, P. (2000). How do Venture Capitalists choose Investments? *University of Chicago, 121: 55-93*.

Kaplan, S., & Strömberg, P. (2003). Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts. *The Review of Economic Studies, 70(2): 281-315*.

Karaömerlioglu, D., & Jacobsson, S. (2000). The Swedish Venture Capital Industry: An Infant, Adolescent or Grown-up? *Venture Capital: International Journal*.

Kasliwal, N., Lade, S., & Prabhune, S. (2012). *Introduction of Clustering by using K-means Methodology .* IJERT.

Kaur, P., Stoltzfus, J., & Yellapu, V. (2018). Descriptive statistics. *International Journal of Academic Medicine*.

Kim, J., & Park, H. (2021). The Influence of Venture Capital Syndicate Size on Venture Performance. *Venture Capital, 23: 1-25*.

Korteweg, A., & Sorensen, M. (2017). Skill and Luck in Private Equity Performance. *Journal of Financial Economics, Volume 124(3)*.

Kraemer-Eis, H., Lang, F., & Gvetadze, S. (2012). European Small Business Finance Outlook. *EIF Research & Market Analysis*.

Kräussl, R., & Krause, S. (2014). Has Europe Been Catching Up? An Industry Level Analysis of Venture Capital Success over 1985–2009. *Eur Financial Management*, pp. 179-205.

Large, D., & Muegge, S. (2008). Venture capitalists' non-financial value-added: an evaluation of the evidence and implications for research. *Journal of Entrepreneurial Finance*, pp. 21-53.

Leleux, B. (2007). The Performance of Venture Capital Investments. *Handbook of Research on Venture Capital: 236-252*.

Lerner, J., Pierrakis, Y., Collins, L., & Biosca, A. (2011). Atlantic Drift: Venture Capital Performance in the UK and US. *Nesta Research Report: June 2011*.

Lincoln, Y., & Guba, E. (1985). Naturalistic inquiry. *CA: Sage*.

Liu, Z., & Zhiqi, C. (2014). Venture Capital Networks and Investment Performance in China. *Australian Economic Papers, 53*.

Ljungqvist, A., & Richardson, M. (2003). The Cash Flow, Return and Risk Characteristics of Private Equity. *National Bureau or Economic Research Working Paper no: 9454*.

Ljungqvist, A., Richardson, M., & Wolfenzon, D. (2007). The Investment Behaviour of Buyout Funds: Theory and Evidence. *AFA 2008 New Orleans Meetings Paper*.

Lockett, A., Murray, G., & Wright, M. (2002). Do UK Venture Capitalists Still Have a Bias Against Investment in New Technology Firms. *Research Policy, Volume 31, Issue 6*.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, pp. 77-91.

Mason, C., & Harrison, R. (2013). *Business angel investment activity in the financial crisis: UK evidence and policy implications.* Retrieved from Accepted for publication in Environment and Planning C: Government and Policy: http://www.gla.ac.uk/schools/business/staff/colinmason/businessangelresearch-recentpublications/#d.en.286602

Mathur, B., & Kaushik, m. (2014). Data Analysis of Students Marks with Descriptive Statistics .

Maula, M., Ahlstrom, J., Haahkola, K., Heikintalo, M., Lindstrom, T., Ojanperä, H., & Tiainen, A. (2006). The Prospects for Successful Early-Stage Venture Capital in Finland. *Sitra Reports, No. 70(37)*.

Maula, M., & Murray, G. (2001). Complementary value-adding roles of corporate venture capital and independent venture capital investors . *Journal of Biolaw and Business*.

McKenzie, M., & Janeway, W. (2011). Venture capital funds and the public equity market. *Accounting and Finance Journal*, pp. 764-786.

McKinsey. (2021). *McKinsey.* Retrieved from A Year of Disruption in the Private Markets: https://www.mckinsey.com/~/media/mckinsey/industries/private%20equity%20and%20principal%20investors/our%20insights/mckinseys%20private%20markets%20annual%20review/2021/mckinsey-global-private-markets-review-2021-v3.pdf

Metrick, A., & Yasuda, A. (2021). *Venture capital & the finance of innovation.* Chichester: Wiley.

Mitchell, M. L., & Jolley, J. M. (2010). *Research design explained: Instructor's edition, 7th ed. .* Wadsworth: Cengage Learning.

Morris, R., Rivas, Y., & Hagdu, D. (2020). *Are markups a good proxy for venture capital investment returns? .* Retrieved from Valor VC: https://valor.vc/venture-capital-returns/

Murray, G., & Marriott, R. (1998). Why has the Performance of Technology-Specialist, European Venture Capital Funds been so Poor? *Research Policy, 27: 947-976.*

Muzyka, D., Birley, S., & Leleux, B. (1996). Trade-offs in the Investment Decisions of European Venture Capitalists. *Journal of Business Venturing.*

Nanda, R., Samila, S., & Sorenson, O. (2020). The Persistent Effect of Initial Success: Evidence from Venture Capital. *Journal of Financial Economics, No. 17-065.*

NVCA. (2020). *NVCA.* Retrieved from Private Equity Funds in Norway 2020: https://nvca.no/wp-content/uploads/2021/04/Private-Equity-Funds-in-Norway-2020.pdf

Pandher, G. (2021). The performance of venture capital investments: failure risk, valuation uncertainty & venture characteristics. *Quantitative Finance*, pp. 929-943.

Parhankangas, A. (2007). An Overview of Research on Early Stage Venture Capital: Current Status and Future Directions. *Handbook of Research on Venture Capital: 236-252.*

Payne, G., & Payne, J. (2004). Key Concepts in Social Research. *SAGE Key Concepts.* Retrieved from SAGE Key Concepts.

Phalippou, L., & Gottschalg, O. (2009). The Performance of Private Equity Funds. *The Review of Financial Studies, 22(4): 1747-1776.*

Prohorovs, A. (2014). Quantitative and qualitative analysis of the informal venture capital in Latvia. *Proceedings of the 2014 International Conference "Economic Science for Rural Development"*.

Quas, A., Marti, J., & Reverte, C. (2021). What money cannot buy: a new approach to measure venture capital ability to add non-financial resources . *Small Business Economics*, pp. 1361-1382.

Rai, P., & Shubha, S. (2010). A Survey of Clustering Techniques . *International Journal of Computer Applications*.

Ramón, N., Ruiz, J., & Sirvent, I. (2020). Cross-benchmarking for performance evaluation: Looking across best practices of different peer groups using DEA. *Omega*.

Röhm, P., Merz, M., & Kuckertz, A. (2020). Identifying corporate venture capital investors – A data-cleaning procedure . *Finance Research Letters*.

Revest, V., & Sapio, A. (2012). Financing technology-based small firms in Europe: what do we know? *Small Bus Econ 39*, pp. 179-205.

Rosenstein, J., Bygrave, W., & Taylor, N. (1993). The CEO, Venture Capitalists, and the Board. *Journal of Business Venturing, 8(2): 99-113*.

Sapienza, H. (1992). When do Venture Capitalists Add Value? *Journal of Business Venturing, 7: 9-27*.

Sapienza, H., Amason, A., & Manigart, S. (1994). The Level and Nature of Venture Capitalist Involvement in their Portfolio Companies: A Study of Three European Countries. *Managerial Finance, 20(1): 3-18*.

Sapienza, H., Manigart, S., & Vermeir, W. (1996). Venture Capitalist Governance and Value Added in Four Countries. *Journal of Business Venturing, 11(6): 439-469*.

Schwarzkopf, C. (2015). *Fostering Innovation and Entrepreneurship.* Springer.

Seppä, M. (2000). Strategy Logic of the Venture Capitalist. *Jyväskylä Studies in Business Economics*.

Seppä, T., & Maula, M. V. (2002). Certification and bargaining power in venture capital: the impact of investor prominence on company valuations. *Working paper*.

Shaw, K., & Sørensen, A. (2019). The Productivity Advantage of Serial Entrepreneurs . *ILR Review, SAGE Journals*, pp. 1225-1261.

Smith, R., Pedace, R., & Sathe, V. (2010). Venture Capital Fund Performance: The Effects of Exits, Abandonment, Persistence, Experience, and Reputation.

Stucke, R. (2011). Updating History .

SVB Capital. (2010). Dialing Down: Venture Capital Returns to Smaller Size Funds. *Venture Capital Update, SVB Capital, May*.

SVCA. (2020). *SVCA.* Retrieved from 2019 Swedish Private Equity Acitivity: https://www.svca.se/wp-content/uploads/2020/06/2019-Swedish-Private-Equity-Activity-Final.pdf

Trochim, W. (2006). *Social Research Methods.* Retrieved from Social Research Methods Knowledge Base: https://www.socialresearchmethods.net/

Vækstfonden. (2019). *From startup to scaleup.* Retrieved from vf.dk: https://vf.dk/media/1891/from-startup-to-scaleup-2019.pdf

Walske, J., & Zacharakis, A. (2009). Genetically Engineered: Why Some Venture Capital Firms Are More Successful Than Others. *Entrepreneurship Theory and Practice, 33*.

Wang, C., & Ang, B. (2004). Determinants of Venture Performance in Singapore. *Journal of Small Business Management, 42: 347-363*.

Wilson, J. (2014). *Essentials of Business Research .* University of East Anglia: SAGE Publications.

Wittenstein, J. (2022). *Meta Loses Top-10 Ranking by Market Value Amid Worst Month Ever* . Retrieved from Bloomberg: https://www.bloomberg.com/news/articles/2022-02-17/meta-platforms-falls-from-ranks-of-10-most-valuable-companies#xj4y7vzkg

Woodward, S. E., & Hall, R. E. (2004). Benchmarking the Returns to Venture. *SSRN*.

Wright, M., Lockett, S., Pruthi, S., Manigart, H., Sapienza, P., Desbrieres, P., & Hommel, U. (2004). Venture capital investors, capital markets, valuation and information: US, Europe and Asia. *Journal of International Entrepreneurship*.

Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. pp. 226-235.

Zarutskie, R. (2010). The Role of Top Management Team Human Capital in Venture Capital Markets: Evidence from First-time Funds. *Journal of Business Venturing, 25(1): 155-172*.

Attachments

A: Overview of previous literature on determinants influencing venture capital performance

B: Jupyter notebook

Appendix A: Overview of Datasets, Research Methods, and Findings

Determinant	Research	Dataset	Research methodology/ Model	Findings
Fund size	Jääskeläinen et al., 2006	Venture Economics	Quantitative/ Poisson regression	Countering effects resulting in a curvilinear inverted U-shaped relationship.
	Kaplan and Schoar, 2005	VentureXpert	Quantitative/ Regression	Positive, but concave. E.g., suggesting decreasing returns to scale.
	SVB Capital, 2010	Preqin	Quantitative/ Central Tendendcy	Negative
	Phalippou and Gottschalg, 2009	VentureXpert	Quantitative/ Regression	Positive and linear, but not significant
	Liu and Zhiqi, 2014	Zdatabase	Quantitative/ Regression	Positive and statistically significant
	Humprey-Jenner, 2012	Preqin, VentureXpert and Execucomp	Quantitative/ Multiple regression models	Negative relationship with performance at 1 % significant
Age of VC firm	Liu and Zhiqi, 2014	Zdatabase	Quantitative/ Regression	U-shaped relationship with performance
	Gompers et al., 2008	Venture Economics	Quantitative/ Regression	Not significant coefficient
	Lerner et al., 2011	Capital IQ (CIQ) Database Dealogic VentureXpert	Quantitative/ Multivariate analysis	Positive relationship with performance
	Kaplan and Schoar, 2005	VentureXpert	Quantitative/ Regression	Positive

	Phalippou and Gottschalg, 2009	VentureXpert	Quantitative/ Regression	Strong positive. Finds that it appears to be the unique explanatory variable for fund performance.
Specialization in investment stage	Swildens and Yee, 2017	Cambridge Associates	Quantitative/ Central tendency	Early-stage funds yield an averaged 21.3% over a 30-year span. Later-stage funds yield an averaged 12.6% over a 30-year span.
	Buchner et al., 2017	CEPRES	Quantitative/ Multivariate analysis	No significant coefficient
Specialization in industry	Buchner et al., 2017	CEPRES	Quantitative/ Multivariate analysis	Finds a positive significant relationship between industry diversification and downside volatility (e.g., higher industry diversification increases the likelihood of picking "losers")
	Humprey-Jenner, 2012	Preqin, VentureXpert and Execucomp	Quantitative/ Multiple regression models	Weak negative correleation with performance
	Cressy et al., 2014	VentureXpert	Quantitative/ QMLE regression	No significant coefficient
	Wang and Ang, 2004	EDB 1997, Questionnaire surveys	Quantitative/ Moderated regression analysis	Positive relationship with performance
	Lockett et al., 2002	BVCA, Questionnaire surveys	Quantitative	No significant coefficient
	Gompers et al., 2008	Venture Economics	Quantitative/ Regression	Finds that poorest performance is associated with unspecialized firms.

Persistence of returns	Kaplan and Schoar, 2005	VentureXpert	Quantitative/ Regression	Finds persistence in both ends of the performance distribution.
	Ljungqvist et al., 2007	Venture Economics	Quantitative/ Multiple regression analyses	Finds persistence for high-performing funds due to high relationships and networks.
	Lerner et al., 2011	Capital IQ (CIQ) Database, Dealogic, VentureXpert	Quantitative/ Multivariate analysis	Finds that a fund raised by a firm whose previous fund performed well is more likely to exhibit superior performance.
Ability to raise following fund	Gompers, 1996	Venture Economics	Quantitative/ Regression	Finds that VC firms that in their first fund who have shown no returns finds it difficult to raise new money.
	Kaplan and Schoar, 2005	VentureXpert	Quantitative/ Regression	Strong positive correlation between the past performance (both as PME and IRR) is strongly related to VC firms' ability to raise future funds.
Ability to increase fund size in following fund	Phalippou and Gottschalg, 2009	VentureXpert	Quantitative/ Regression	Finds that a firm with inferior track record struggles to raise a large fund.
	Kaplan and Schoar, 2005	VentureXpert	Quantitative/ Regression	Strong positive correlation between the past performance (both as PME and IRR) and VC firms' ability to increase fund size.

Determinant	Research	Dataset	Research methodology/ Model
Background and experience of VC firm partners	Walske and Zacharakis, 2009	VentureXpert	Quantitative/ Poisson Regression
	Lerner et al., 2007	Venture Economics	Quantitative/ Regression
	Zarutskie, 2010	VentureXpert	Quantitative/ Regression
	Dimov and Shepherd, 2005	VentureXpert	Quantitative/ Regression
	Abell and Nisar, 2007	Venture Economics, FAME, VentureXpert, Diane	Quantitative/ Regression
	Lerner et al., 2011	Capital IQ (CIQ) Database Dealogic VentureXpert	Quantitative/ Multivariate analysis
	Sapienza et al., 1996	Pratt's Guide to Venture Capital Sources, Questionnaire surveys	Quantitative/ Linear regression
	Gompers et al., 2009	Venture Economics	Quantitative/ Regression
	Walske and Zacharakis, 2009	VentureXpert	Quantitative/ Poisson Regression

Specialization in geographical focus	Buchner et al., 2017	CEPRES	Quantitative/ Multivariate analysis
	Humprey-Jenner, 2012	Preqin, VentureXpert and Execucomp	Quantitative/ Multiple regression models
	Cressy et al., 2014	VentureXpert	Quantitative/ QMLE regression
	Wang and Ang, 2004	EDB 1997, Questionnaire surveys	Quantitative/ Moderated regression analysis
Investment strategy	Arundale, 2018	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis
	Arundale, 2017	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
Reputation and brand	Fleming et al., 2005	Australian Bureau of Statistics	Quantitative/ Regression analysis
	Hsu, 2002	Previous studies	Quantitative/ Comparative statics analysis
	Gompers and Lerner, 1999	VentureXpert	Quantitative/ Multivariate regressions

	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis
Network	Cumming, 2006	Canadian Venture Capital Association	Quantitative/ Linear regression and Box-Cox transformations
	Gorman and Sahlman, 1989	Questionnaire survey	Quantitative/ Probability density function
	Hochberg et al., 2007	Thomson Reuters Venture Economics	Quantitative/ Different regression models
Due diligence	Sapienza et al., 1994	Questionnaire survey	Quantitative/ Linear regression
	Amit et al., 1998	Canadian Venture Capital Association	Qualitative/ Formal model
	Arundale, 2018	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
Investment terms	Kaplan and Stromberg, 2003	Custom dataset from VC firms in authors network	Quantitative/ Cross-sectional regression
Syndication size	Bygrave, 1988	VentureXpert	Quantitative/ Network analysis with graph theory
	Hochberg et al., 2007	Venture Economics	Quantitative/ Different regression models
	Kim et al., 2021	Data collected by professional survey firm	Qualitative/ Fuzzy-set qualitative comparative analysis
Adding value	Bottazzi et al., 2008	Questionnaire survey	Quantitative/ Multivariate regression analysis
	Gorman and Sahlman, 1989	Questionnaire survey	Quantitative/ Probability density function
	Hellmann and Puri, 2002	Questionnaire survey and interviews	Quantitative/ Multivariate probit regression framework

Exits	Axelson and Martinovic, 2013	Venture Economics	Quantitative/ Multiple regression analyses
Luck	Korteweg and Sorensen, 2017	Preqin	Quantitative/ Multiple regression analyses
Ecosystem	Diller and Kaserer, 2009	VentureXpert	Quantitative/ Multiple regression analyses
	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis
	Ljungqvist et al., 2007	Venture Economics	Quantitative/ Multiple regression analyses
Capital inflow into the VC industry	Diller and Kaserer, 2009	VentureXpert	Quantitative/ Multiple regression analyses
	Ljungqvist et al., 2007	Venture Economics	Quantitative/ Multiple regression analyses

Determinant	Research	Dataset	Research methodology/ Model
Background and experience of VC firm partners	Walske and Zacharakis, 2009	VentureXpert	Quantitative/ Poisson Regression
	Lerner et al., 2007	Venture Economics	Quantitative/ Regression
	Zarutskie, 2010	VentureXpert	Quantitative/ Regression
	Dimov and Shepherd, 2005	VentureXpert	Quantitative/ Regression
	Abell and Nisar, 2007	Venture Economics, FAME, VentureXpert, Diane	Quantitative/ Regression
	Lerner et al., 2011	Capital IQ (CIQ) Database Dealogic VentureXpert	Quantitative/ Multivariate analysis
	Sapienza et al., 1996	Pratt's Guide to Venture Capital Sources, Questionnaire surveys	Quantitative/ Linear regression
	Gompers et al., 2009	Venture Economics	Quantitative/ Regression
Specialization in geographical focus	Walske and Zacharakis, 2009	VentureXpert	Quantitative/ Poisson Regression
	Buchner et al., 2017	CEPRES	Quantitative/ Multivariate analysis
	Humprey-Jenner, 2012	Preqin, VentureXpert and Execucomp	Quantitative/ Multiple regression models
	Cressy et al., 2014	VentureXpert	Quantitative/ QMLE regression
	Wang and Ang, 2004	EDB 1997, Questionnaire surveys	Quantitative/ Moderated regression analysis
Investment strategy	Arundale, 2018	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis

	Arundale, 2017	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
Reputation and brand	Fleming et al., 2005	Australian Bureau of Statistics	Quantitative/ Regression analysis
	Hsu, 2002	Previous studies	Quantitative/ Comparative statics analysis
	Gompers and Lerner, 1999	VentureXpert	Quantitative/ Multivariate regressions
	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis
Network	Cumming, 2006	Canadian Venture Capital Association	Quantitative/ Linear regression and Box-Cox transformations
	Gorman and Sahlman, 1989	Questionnaire survey	Quantitative/ Probability density function
	Hochberg et al., 2007	Thomson Reuters Venture Economics	Quantitative/ Different regression models
Due diligence	Sapienza et al., 1994	Questionnaire survey	Quantitative/ Linear regression
	Amit et al., 1998	Canadian Venture Capital Association	Qualitative/ Formal model
	Arundale, 2018	Questionnaire survey	Qualitative/ Multi-theoretical conceptual framework
Investment terms	Kaplan and Stromberg, 2003	Custom dataset from VC firms in authors network	Quantitative/ Cross-sectional regression
Syndication size	Bygrave, 1988	VentureXpert	Quantitative/ Network analysis with graph theory
	Hochberg et al., 2007	Venture Economics	Quantitative/ Different regression models
	Kim et al., 2021	Data collected by professional survey firm	Qualitative/ Fuzzy-set qualitative comparative analysis

Adding value	Bottazzi et al., 2008	Questionnaire survey	Quantitative/ Multivariate regression analysis
	Gorman and Sahlman, 1989	Questionnaire survey	Quantitative/ Probability density function
	Hellmann and Puri, 2002	Questionnaire survey and interviews	Quantitative/ Multivariate probit regression framework
Exits	Axelson and Martinovic, 2013	Venture Economics	Quantitative/ Multiple regression analyses
Luck	Korteweg and Sorensen, 2017	Preqin	Quantitative/ Multiple regression analyses
Ecosystem	Diller and Kaserer, 2009	VentureXpert	Quantitative/ Multiple regression analyses
	Aigner et al., 2008	Confidental dataset from a major fund-of-funds investor in Europe	Quantitative/ Markov transition and linear regression analysis
	Ljungqvist et al., 2007	Venture Economics	Quantitative/ Multiple regression analyses
Capital inflow into the VC industry	Diller and Kaserer, 2009	VentureXpert	Quantitative/ Multiple regression analyses
	Ljungqvist et al., 2007	Venture Economics	Quantitative/ Multiple regression analyses

K-Means VC Analysis

May 27, 2022

Library Requirements

[1]:
```
# Jupyter notebook imports
from IPython.display import Image

# Data science imports
import pandas as pd
import numpy as np

# Plotly
import plotly.express as px
import plotly.graph_objects as go

# Matplotlib
import matplotlib.pyplot as plt
from matplotlib.pyplot import figure

# ML libraries
from kmodes.kmodes import KModes
from sklearn.cluster import KMeans
from sklearn.preprocessing import Normalizer
from sklearn.metrics import silhouette_score
from sklearn.preprocessing import StandardScaler

# Image I/O library
import imageio

# Operative system libraries
import os
from os import listdir
from os.path import isfile, join

import warnings
warnings.filterwarnings('ignore')
```

1 1 Cleaning the VC data

1.0.1 1.1 Reading the raw VC data

[2]:
```
df = pd.read_csv('./data/nordic_vc.csv')

# Rename
df = df.rename(columns = {'Unnamed: 0': 'cont'})

# Drop unnamed columns
df = df[df.columns[df.columns.isin(['Year', 'Fund Manager', 'Fund', 'Size (M␣
↪EUR)', 'Country', 'Stage', 'Industry'])]]
```

1.0.2 1.2 Removing funds not in the nordics

[3]:
```
# Getting all non-nordic funds
non_nordic_df = df[~df['Country'].isin(['Norway', 'Sweden', 'Denmark',␣
↪'Finland'])]

# Removing all non-nordic funds
df = df[df['Country'].isin(['Norway', 'Sweden', 'Denmark', 'Finland'])]
```

1.0.3 1.3 Removing recent funds (created after 2016) with only one fund raised

[4]:
```
num_funds_df =  pd.DataFrame(df['Fund Manager'].value_counts())
num_funds_df = num_funds_df.reset_index()
num_funds_df = num_funds_df.rename(columns = {'Fund Manager': 'count', 'index':␣
↪'Fund Manager'})


df = pd.merge(df, num_funds_df,on='Fund Manager')
print("Amount of datapoints:", len(df))

# Funds with only 1 fund, created after 2016
singular_fund_df = df.loc[df['count'] == 1]
singular_fund_df = singular_fund_df.loc[singular_fund_df['Year'] >= 2017]
singular_fund_list = list(singular_fund_df['Fund Manager'])

print("Amount of recent singular funds:", len(singular_fund_df))

# Removing the funds
df = df[~df['Fund Manager'].isin(singular_fund_list)]
```

```
Amount of datapoints: 417
Amount of recent singular funds: 23
```

1.0.4 1.4 Removing dead funds

[5]:
```
unique_funds_df = pd.read_csv('./data/unique_vc_firm.csv')

# Rename
unique_funds_df = unique_funds_df.rename(columns = {'Fund manager': 'Fund␣
↪Manager'})

# Drop unnamed columns
unique_funds_df = unique_funds_df[unique_funds_df.columns[unique_funds_df.
↪columns.isin(['Fund Manager', 'Status'])]]

analysis_df = pd.merge(df, unique_funds_df,on='Fund Manager')

print("Statistics for VC funds:")
print(analysis_df['Status'].value_counts())

discontinued_df = analysis_df.loc[analysis_df['Status'] == 'Discontinued']

analysis_df = analysis_df.loc[analysis_df['Status'] != 'Discontinued']
```

```
Statistics for VC funds:
Active          302
Discontinued     43
Merged            4
Name: Status, dtype: int64
```

1.0.5 1.5 Removing rows without values

[6]:
```
nan_df = analysis_df[analysis_df.isna().any(axis=1)]

analysis_df = analysis_df.dropna(axis=1)
```

2 2 Measuring performance

2.0.1 2.1 Setting the values we will use for performance comparison

2.1.1 Adding stage mean

[7]:
```
# Applying the metric
analysis_df = df

# Getting means for each phase
stage_list = analysis_df['Stage'].unique()

stage_means = {}
for stage in stage_list:
    stage_df = analysis_df.loc[analysis_df['Stage'] == stage]
```

```python
    stage_means[stage] = round(stage_df['Size (M EUR)'].sum() / len(stage_df),␣
↪2)

def apply_stage_mean(x):
    stage = x['Stage']
    return stage_means[stage]

# Adding stage means
mean_raised_series = analysis_df['Stage'].map(stage_means)
analysis_df['stage_mean'] = mean_raised_series.values
df['stage_mean'] = mean_raised_series.values
```

2.1.2 Finding the average time to raise a new fund for each fund manager

[8]:
```python
# Inverted x-axis metric
def avg_time_to_raise(x):
    return round(((x['Year'].max() - x['Year'].min()) / len(x)), 2)

# avg_to_raise
avg_to_raise_series = df[['Year', 'Size (M EUR)', 'Fund Manager']].
↪groupby("Fund Manager").apply(avg_time_to_raise)

avg_to_raise_df = pd.DataFrame(avg_to_raise_series)
avg_to_raise_df = avg_to_raise_df.rename(columns = {0: "avg_to_raise"})
analysis_df = pd.merge(analysis_df, avg_to_raise_df,on='Fund Manager')
```

2.1.3 Finding the mean fund size for each fund manager

[9]:
```python
def mean_raised(x):
    return round(x['Size (M EUR)'].mean(), 2)

# mean_raised
mean_raised_series = df[['Year', 'Size (M EUR)', 'Fund Manager']].groupby("Fund␣
↪Manager").apply(mean_raised)

mean_raised_df = pd.DataFrame(mean_raised_series)
mean_raised_df = mean_raised_df.rename(columns = {0: "mean_raised"})
analysis_df = pd.merge(analysis_df, mean_raised_df,on='Fund Manager')
```

2.0.2 2.2 Calculating the two different performance measurements

2.2.1 Finding the % delta between sequential funds

[10]:
```python
analysis_df['delta_raised'] = 0

# Updating mean raised for all rows
for i in range(len(analysis_df)):
```

```
    # Check if the fund manager is the same
    if analysis_df['Fund Manager'].iloc[i] == analysis_df['Fund Manager'].
↪iloc[i-1]:

        ## Setting delta time raised
        fund, year = analysis_df['Fund Manager'].iloc[i], analysis_df['Year'].
↪iloc[i]
        fund_df = analysis_df.loc[analysis_df['Fund Manager'] == fund]
        fund_df = fund_df.loc[np.array(fund_df['Year']) <= np.array(year)]
        analysis_df["avg_to_raise"].iloc[i] = round(((fund_df['Year'].max() -␣
↪fund_df['Year'].min()) / len(fund_df)), 2)

        ## Setting the delta raised variable
        prev_fund = analysis_df['Size (M EUR)'].iloc[i-1]
        curr_fund = analysis_df['Size (M EUR)'].iloc[i]

        analysis_df['delta_raised'].iloc[i] = round((curr_fund / prev_fund) -␣
↪1, 2)
```

2.2.2 Inverting the y-axis

[11]:
```
# Switching the y-axis
avg_to_raise_max = analysis_df['avg_to_raise'].max()
analysis_df['avg_to_raise'] = avg_to_raise_max - analysis_df['avg_to_raise']
```

2.2.3 Finding the % delta between a fund and the stage mean

[12]:
```
def stage_mean_delta(x):

    sum_stage_mean_delta = 0
    x = pd.DataFrame(x)
    all_deltas = []

    for i in range(len(x)):
        prev_fund = x['Size (M EUR)'].iloc[i]
        stage_mean = x['stage_mean'].iloc[i]
        all_deltas.append((prev_fund / stage_mean) - 1)
        sum_stage_mean_delta += (prev_fund / stage_mean) - 1

    return sum_stage_mean_delta

# delta_stage_mean
mean_raised_series = df[['stage_mean', 'Size (M EUR)', 'Fund Manager']].
↪groupby("Fund Manager").apply(stage_mean_delta)

mean_raised_df = pd.DataFrame(mean_raised_series)
mean_raised_df = mean_raised_df.rename(columns = {0: "delta_stage_mean"})
```

```
analysis_df = pd.merge(analysis_df, mean_raised_df,on='Fund Manager')


# Getting the stage mean
analysis_df['delta_stage_mean'] = (analysis_df['Size (M EUR)'] /␣
 ↪analysis_df['stage_mean']) - 1
```

2.0.3 2.4 Adding the calculated deltas

[13]:
```
def calc_metric(x):
    x = pd.DataFrame(x)

    # Sum the deltas in % raised
    delta_raised_sum = x['delta_raised'].sum() - x['delta_raised'].iloc[0]

    # Sum the relative fund size
    delta_stage_mean = x['delta_stage_mean'].sum()
    return delta_raised_sum + delta_stage_mean

# Getting the metric to use in k-means
mean_raised_series = analysis_df[['mean_raised', 'delta_raised',␣
 ↪'delta_stage_mean', 'Fund Manager']].groupby("Fund Manager").
 ↪apply(calc_metric)

mean_raised_df = pd.DataFrame(mean_raised_series)
mean_raised_df = mean_raised_df.rename(columns = {0: "metric"})
analysis_df = pd.merge(analysis_df, mean_raised_df,on='Fund Manager')
```

[14]:
```
analysis_df['metric'] = 0

# Updating mean raised for all rows
for i in range(len(analysis_df) - 1):

    # Check if we should accumulate the sum
    if analysis_df['Fund Manager'].iloc[i] == analysis_df['Fund Manager'].
 ↪iloc[i-1]:

        new_metric = analysis_df['delta_raised'].iloc[i] +␣
 ↪analysis_df['delta_stage_mean'].iloc[i]
        old_sum = analysis_df['metric'].iloc[i-1]

        analysis_df['metric'].iloc[i] = round((new_metric + old_sum), 2)

    # If this is the first metric-entry for this fund manager
    else:
```

```
        new_metric = analysis_df['delta_raised'].iloc[i] +␣
↪analysis_df['delta_stage_mean'].iloc[i]
        analysis_df['metric'].iloc[i] = round((new_metric), 2)
```

2.0.4 Plotting the distribution of fund managers

Getting the newest metric

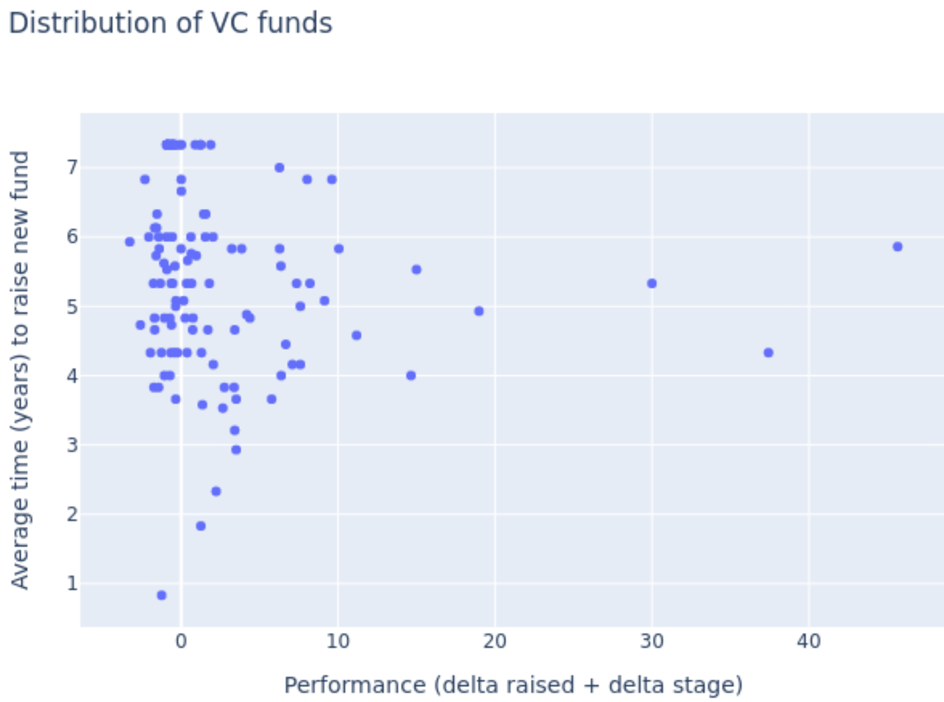
[15]:
```
def get_newest(x):
    x = pd.DataFrame(x)
    newest_metric = x['metric'].loc[x['Year'] == x['Year'].max()]
    return newest_metric

newest_metric = analysis_df[['metric', 'Fund Manager', 'Year', 'avg_to_raise']].
↪groupby("Fund Manager").apply(get_newest)
newest_metric = pd.DataFrame(newest_metric)


subset_df = analysis_df.iloc[newest_metric.index.droplevel(0)]


_analysis_df = analysis_df
analysis_df = subset_df
```

[16]:
```
analysis_df = analysis_df.loc[analysis_df["avg_to_raise"] != 0]
```

[17]:
```
fig = px.scatter(analysis_df, x="metric", y="avg_to_raise",
                 labels={
                     "metric": "Performance (delta raised + delta stage)",
                     "avg_to_raise": "Average time (years) to raise new fund",
                 },
                title="Distribution of VC funds")

fig.write_image('./plots/plot_2_4.png')
```

[18]:
```
Image(filename='./plots/plot_2_4.png')
```

[18]:



3 3.0 Running K-Means

3.0.1 3.1 Setting the clusters we won't calculate

3.1.1 Setting funds with no delta as own cluster

[19]:
```
no_delta_df = analysis_df.loc[analysis_df['metric'] == 7.33]
no_delta_df['cluster'] = -1

analysis_df =  analysis_df.loc[analysis_df['avg_to_raise'] != 7.33]
```

3.1.2 Setting outperformers as own cluster

[20]:
```
outperformers_df = analysis_df.loc[analysis_df['metric'] > 20]
outperformers_df['cluster'] = -2
analysis_df = analysis_df.loc[analysis_df['metric'] < 20]
```

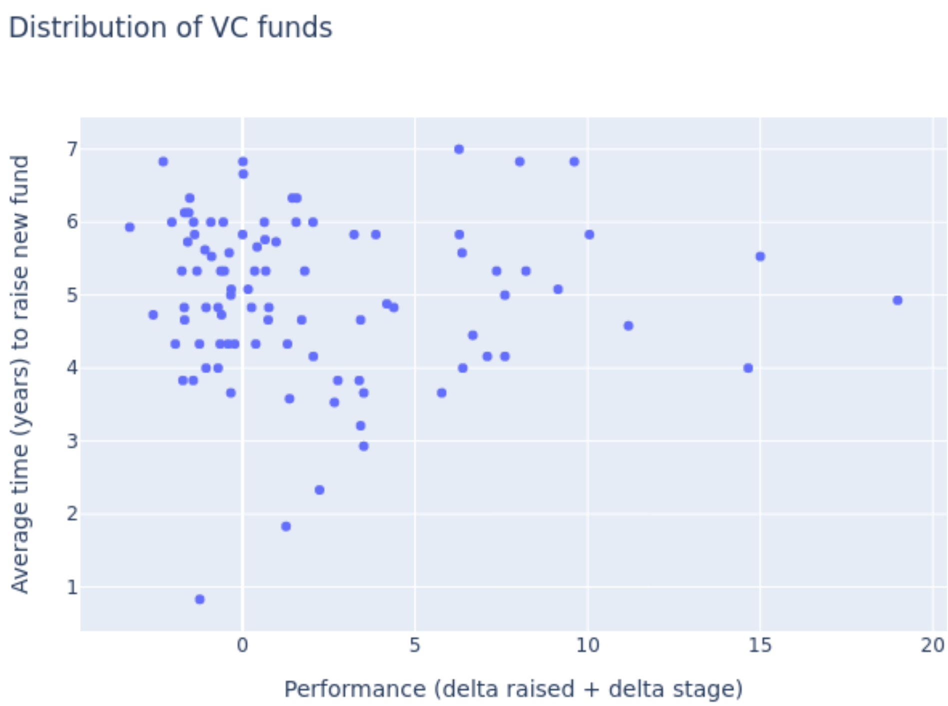
3.1.3 Plotting the new distribution

[21]:
```
fig = px.scatter(analysis_df, x="metric", y="avg_to_raise",
                 labels={
                     "metric": "Performance (delta raised + delta stage)",
                     "avg_to_raise": "Average time (years) to raise new fund",
```

```
                    },
                  title="Distribution of VC funds")

fig.write_image('./plots/plot_3_1_3.png')
```

[22]:
```
Image(filename='./plots/plot_3_1_3.png')
```

[22]:



3.1 3.2 Preparing the data for k-means

3.2.1 Wrangling the data to correct format

[23]:
```
k_means_df = analysis_df
k_means_df = k_means_df

# Get the array to fit
X = k_means_df[["metric", "avg_to_raise"]].fillna(0).to_numpy()
```

3.2.2. Normalizing and standardizing the data

[24]:
```
# Set switches
normalize = False
standardize = True
plot_normalized = True
```

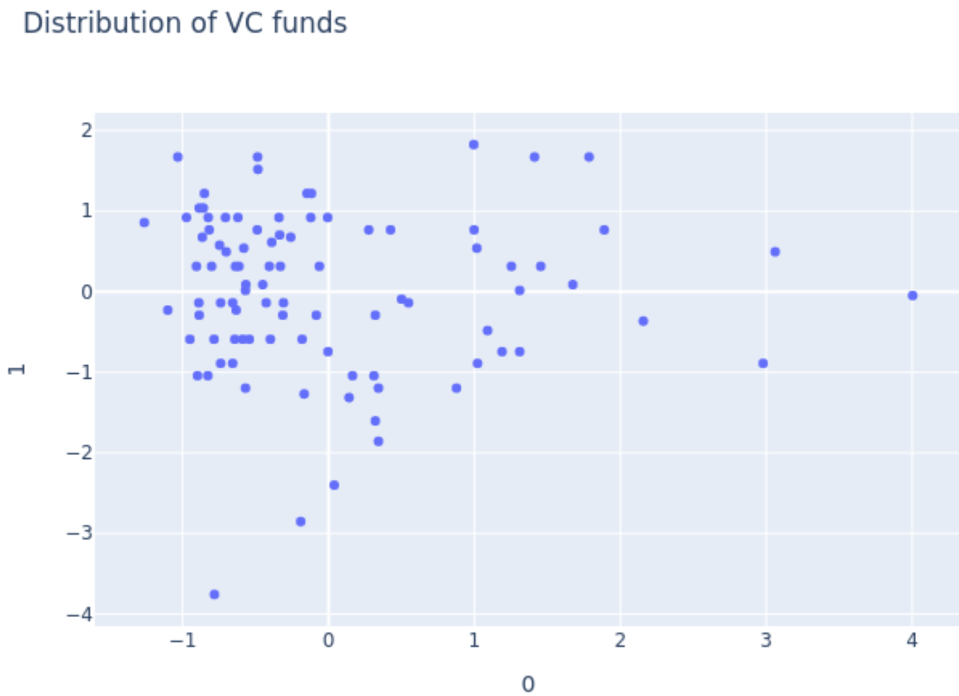
```
remove_outliers = True

if normalize:
    transformer = Normalizer()  # fit does nothing.
    X = transformer.transform(X)

if standardize:
    scaler = StandardScaler()
    scaler.fit(X)
    X = scaler.transform(X)
```

3.2.3 Plotting the resulting data

[25]:
```
if plot_normalized:

    x_normalized = pd.DataFrame(X)

    fig = px.scatter(x_normalized, x=0, y=1,
                 labels={
                     "metric": "Mean fund size raised",
                     "avg_to_raise": "Average time (years) to raise new fund",
                 },
                title="Distribution of VC funds")

    fig.write_image('./plots/plot_3_2_3.png')
```

[26]:
```
Image(filename='./plots/plot_3_2_3.png')
```

[26]:



3.1.1 3.3 Finding the optimal number of clusters

[27]:
```
sil = []
kmax = 10
best_sil_score = 0
# dissimilarity would not be defined for a single cluster, thus, minimum number␣
↪of clusters should be 2
for k in range(2, kmax + 1):
    kmeans = KMeans(n_clusters=k).fit(X)
    labels = kmeans.labels_

    score = silhouette_score(X, labels, metric="euclidean")

    if score > best_sil_score:
        best_sil_score = score
        best_k = k
    sil.append([score, k])

arr = np.array(sil)
sil_df = pd.DataFrame(arr, columns=["Silhouette Score", "Num Clusters"])
```

```
fig = px.line(sil_df, x="Num Clusters", y="Silhouette Score", title="Num␣
 ↪Clusters")

print(f"Optimal number of clusters is {best_k} with a silhouette score of␣
 ↪{round(best_sil_score, 2)}")


fig.write_image('./plots/plot_3_3.png')
```

```
Optimal number of clusters is 2 with a silhouette score of 0.44
```

[28]:
```
Image(filename='./plots/plot_3_3.png')
```

[28]:

3.1.2 3.4 Running K-Means

[29]:
```
km = KMeans(
    n_clusters=4,
    init='random',
    max_iter=300,
    random_state=0
```

```
)

y_km = km.fit_predict(X)


k_means_df['cluster'] = y_km
k_means_df['cluster'] = k_means_df['cluster'].astype(str)


analysis_df['cluster'] = k_means_df['cluster']
analysis_df['cluster'] = analysis_df['cluster'].fillna(-1).astype(str)


if remove_outliers:
    analysis_df = analysis_df.loc[analysis_df['cluster'] != "-1"]
```

4 4.0 Inspecting the results

4.1 Plotting the distribution of VC funds

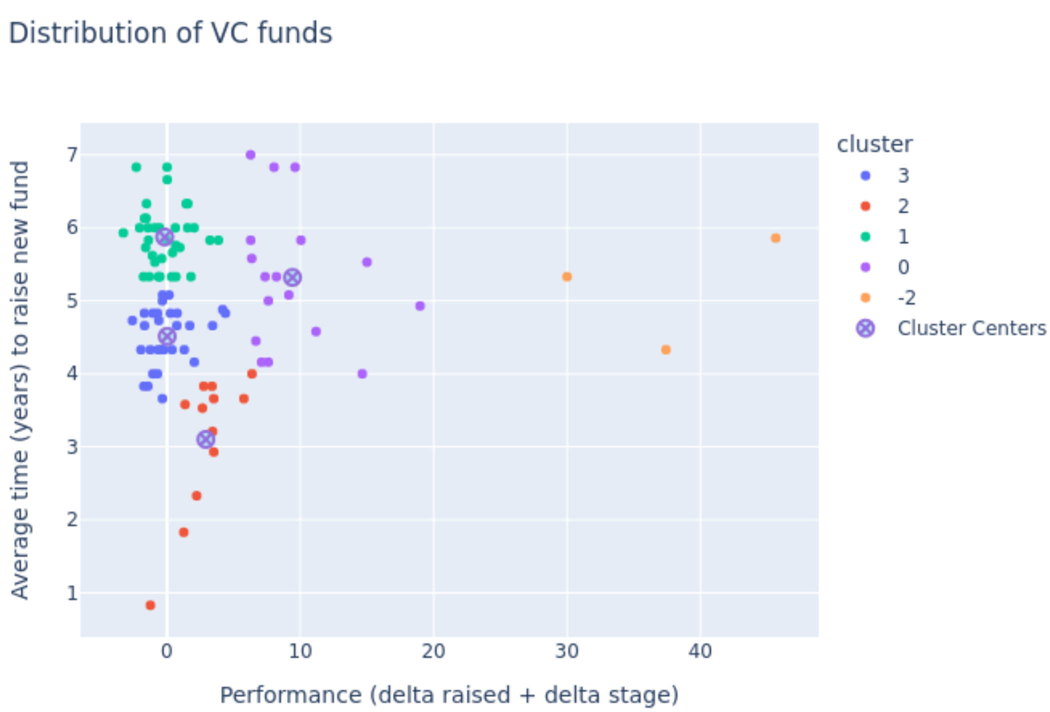
[30]:

```
if standardize:
    std_arr = scaler.inverse_transform(km.cluster_centers_)
    y = std_arr[:,1]
    x = std_arr[:, 0]
else:
    y = km.cluster_centers_[:, 1],
    x = km.cluster_centers_[:, 0]


analysis_df = analysis_df.append(no_delta_df)
analysis_df = analysis_df.append(outperformers_df)

fig = px.scatter(analysis_df, x="metric", y="avg_to_raise",
                color='cluster',
                 labels={
                     "metric": "Performance (delta raised + delta stage)",
                     "avg_to_raise": "Average time (years) to raise new fund",
                 },
                title="Distribution of VC funds")

fig.add_trace(go.Scatter(y = y,
                        x = x,
                        name = 'Cluster Centers',
                        mode="markers",
                        marker_symbol = "circle-x",
                        marker=dict(
                            color='rgba(135, 206, 250, 0.5)',
```

```
                                size=10,
                                line=dict(
                                    color='MediumPurple',
                                    width=2
                                ))),
                            row=1, col=1)

fig.write_image('./plots/plot_4_1.png')
```

[31]:
```
Image(filename='./plots/plot_4_1.png')
```

[31]:



4.1.1 Saving the old dataframe

[32]:
```
old_analysis_df = analysis_df
analysis_df = _analysis_df
```

4.0.1 4.2 Rerunning the plot for all years between 2000 - 2022

[33]:
```
import plotly as plotly

if standardize:
    std_arr = scaler.inverse_transform(km.cluster_centers_)
```

```
    y = std_arr[:,1]
    x = std_arr[:, 0]
else:
    y = km.cluster_centers_[:, 1],
    x = km.cluster_centers_[:, 0]

# Inverted x-axis metric
def avg_time_to_raise(x):
    return round(((x['Year'].max() - x['Year'].min()) / len(x)), 2)

def get_newest(X):
    x = pd.DataFrame(X)
    newest_metric = x['metric'].loc[x['Year'] == x['Year'].max()]

    # When two funds exist for the same year, choose the best metric
    if len(newest_metric > 1):
        newest_metric = newest_metric.loc[newest_metric == newest_metric.max()]

    return newest_metric

for year in range(2000, 2023, 1):

    # Getting only the funds before the current year
    subset_df = analysis_df.loc[analysis_df['Year'] <= year]

    # Getting only the newest metric
    newest_metric = subset_df[['avg_to_raise', 'metric', 'Fund Manager',␣
 ↪'Year']].groupby("Fund Manager").apply(get_newest)
    newest_metric = pd.DataFrame(newest_metric).reset_index()
    newest_metric = subset_df.filter(items = list(newest_metric['level_1']),␣
 ↪axis=0)

    # Merge the newest metric with the cluster values
    newest_metric = newest_metric.merge(old_analysis_df[['Fund Manager',␣
 ↪'cluster']], on="Fund Manager", how='left')

    # Drop Fund Managers with nan values. THis is investors with only one fund,␣
 ↪and they are not analyzed
    newest_metric = newest_metric.dropna(axis=0)

    # Sorting the dataframe on cluster value
    newest_metric['cluster'] = pd.to_numeric(newest_metric['cluster'])
    newest_metric = newest_metric.sort_values(by=['cluster'], ascending=False)

    # Get only most recent fund (again)
    newest_metric = newest_metric.loc[newest_metric['Year'] <= year]
```

```
    # Creating the plots
    fig = px.scatter(newest_metric, x="metric", y="avg_to_raise",
                    color='cluster',
                    color_continuous_scale=px.colors.sequential.Turbo,
                    labels={
                        "metric": "Performance (delta raised + delta stage)",
                        "avg_to_raise": "Average time (years) to raise new␣
 ↪fund",
                    },
                    title=f"Distribution of VC funds in year {year}")

    fig.add_trace(go.Scatter(y = y,
                            x = x,
                            name = 'Cluster Centers',
                            mode="markers",
                            marker_symbol = "circle-x",
                            marker=dict(
                                color='rgba(135, 206, 250, 0.5)',
                                size=10,
                                line=dict(
                                    color='MediumPurple',
                                    width=2
                                ))),
                            row=1, col=1)

    # Saving to disk
    fig.update_layout(yaxis_range=[-1,8])
    fig.update_layout(xaxis_range=[-7,50])
    fig.write_image(f"./images/{year}_plot.png")
```

4.2.1 Saving image snapshot from all years between 2000 - 2022

[34]:
```
# Paths
mypath = "./images/"
gif_name = "vc_time_series.gif"

if os.path.exists(mypath + gif_name):
    os.remove(mypath + gif_name)

filenames = [mypath + f for f in listdir(mypath)]
filenames.sort()

# Appending images to list
images = []
for filename in filenames:
    images.append(imageio.imread(filename))
imageio.mimsave(mypath + gif_name, images, fps=1.3)
```

[35]:
```python
print("Images and GIF saved to disk succesfully.")
```

```
Images and GIF saved to disk succesfully.
```

4.2.2 Save resulting data to disk

[36]:
```python
# Switch for writing the dataframe to disk
write_to_disk = True
if write_to_disk:
    analysis_df.to_csv('growth+stage.csv', encoding="utf-8-sig")
print("CSV saved to disk succesfully.")
```

```
CSV saved to disk succesfully.
```

NTNU
Kunnskap for en bedre verden