Norwegian School of Economics

Bergen, Spring, 2019

Predicting Success in Early-Stage Venture Capital Investment

Hans Jacques Lindahl Falck

Supervisor: Karin S. Thorburn

Master Thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible − through the approval of this thesis − for the theories and methods used, or results and conclusions drawn in this work.

Contents

1. Abstract

In this thesis, I examine predictors of early-stage success (ESS) for early-stage venture capital (ESVC) investments, observable at the point of initial investment. I use a self-collected sample of 30 Norwegian portfolio companies, which all received early-stage funding from a single Norwegian ESVC between 2010 and 2014. A company is characterized as having achieved ESS if it succeeds in raising a pre-defined level of additional capital or increased sales sufficiently within three years of the initial ESVC investment. Cross-sectional regressions show that a higher number of full-time employees (FTEs) at the point of initial ESVC investment significantly improves the likelihood of ESS. The probability of success is also significantly higher for firms with experienced founders, defined as having both prior startup experience and a PhD. Understanding these relationships may help ESVCs reduce risk in their investment decisions, improving expected long-run returns.

"Technological change is increasingly seen as the primary driver of productivity growth, adaptation by financial intermediaries may be equally important to realizing the benefits of these new technologies."
(Ewens, Nanda, & Rhodes-Kropf, 2017)

2. Acknowledgements

I begin by thanking my father, **Thomas Scheen Falck**, for the monumental support and patience he has shown me throughout my life.

I am extremely thankful to my guidance-councillor, **Karin S. Thorburn**, for the extensive guidance and support she has shown me throughout this process. I am especially thankful for her efficient guidance, sharpening the document substantially. I am certain that this thesis would never have come into existence had it not been for her wise counsel.

I am thankful to the ESVC Firms whose partners were willing to share their time and expertise with me during the data-collection process. Without their positive attitude towards the project, this thesis would not have a representative sample to conduct empirical analysis on[1].

I am further thankful to all the founders & CEOs of the portfolio-startups who were willing to schedule 15- to 60-minute interviews with me, so that I might collect the required data-sample. I apologize for every time an interview exceeded its planned limits.

I am also very thankful to Åsmund Røyset, Torstein Sandaa-Johansen & Magnus Romstad Stavne; for being willing to comment om the one-page progress-reports I sent them on a weekly basis for the last twenty-four weeks of writing. Had it not been for their continued support, I would in all likelihood still be writing this thesis.

I am thankful to friends and family who have patiently supported me throughout this process. Your kind words and gestures have meant more to me than may have seemed apparent at the time.

I would like to thank **Maria Benedicte Færevaag** for her continued patience and support – I apologize for taking so long.

My thanks to you, the reader. I hope that this document may be of value in some shape or form.

[1] Though I would like to name the people who have generously aided me in the data-collection process, a prerequisite for cooperation has been the anonymity of involved parties. This applies both to the involved ESVCs as well as the firm CEOs & founders.

3. Introduction

> *"Understanding how capital markets affect the growth and survival of newly created firms is perhaps the defining question of entrepreneurial finance."*
> (Robb & Robinson, 2009)

In this thesis, I examine predictors of early-stage success (ESS) for early-stage venture capital (ESVC) investments, observable at the point of initial investment.

I analyse a hand-collected sample of 30 Norwegian portfolio companies, which all received early-stage funding from a single Norwegian ESVC between 2010-2014[2]. The data has been gathered through a combination of interviews with firm CEOs & founders, analysis of annual reports and dialogue with ESVC partners. The larger population from which the sample is drawn is defined as: *Norwegian startups, operating with headquarters in Norway, with Norwegian-private citizens or firms as majority owners, who have received funding from Norwegian ESVCs.* The sample is drawn from three entire ESVC portfolios, totalling 32 investments. Of these, 31 investments match the population definition. The response rate for these firms is 97%.

I argue that entire ESVC investment portfolios, such as the ones studied here, may be viewed as random population samples of ESVC-backed firms: An empirical study of nascent entrepreneurs indicated that only 19% of observed entrepreneurs had an ambition to grow large, while the remaining 81% had a stated ambition to stay small (Gelderen, Thurik, & Bosma, 2005). As ESVC investment is a pre-requisite for sample inclusion, there is no need to further consider firm growth potential. The ESVCs will only invest in firms deemed to hold potential for significant growth and are the market's most competent judges of this potential.

A company is characterized as having achieved ESS if it succeeds in raising a pre-defined level of additional capital or increased sales sufficiently within three years of the initial ESVC investment. Using ESS as a success criterion provides value as a predictor of absolute success for ESVCs, argued by Ewens et al. (2017) and in *Chapter 5*.

[2] 87% of sample funded between 2010-2012.

Cross-sectional regressions show that a higher number of full-time employees (referenced as FTEs) at the point of initial ESVC investment significantly improves the probability of ESS. The probability of success is also significantly higher for firms with experienced founders, defined as holding both prior startup experience and a PhD. I test a total of six cross-sectional regression models, based on different uncertainties in the recording of the determinants.

If discovered relationships could be reliably quantified, the resulting reduced risk in ESVC investment decisions could improve capital efficiency. The hope is that this may contribute to increasing the total volume of available venture capital, at a time when the Norwegian startup economy is in rapid growth (Abelia, 2017).

3.1 A Broad Overview

Chapter 4 presents an overview of academic literature, focusing on arguments relating to the conducted empirical analysis.

Chapter 5 presents a theoretical argument for the value of ESS, building upon a decision framework presented by Ewens et al. (2017), representing the investor's game tree. I argue that ESS creates a significant improvement in the probability of liquid returns on investment for the ESVC.

Chapter 6 elaborates on the data collection process, while **Chapter 7** presents the structure of the empirical analysis. The research question and hypotheses are formalized here.

Chapter 8 contains the results of the empirical study and comments on each hypothesis in light of the results. **Chapter 9** contains brief discussions and a consideration of internal and external validity of the study.

Chapter 10 contains future research steps. They are organized at a macro, methodological and micro level.

The conclusion is written in **Chapter 11**.

The Appendix elaborates upon the collected data. It includes tables of descriptive statistics for collected variables not included in the empirical analysis.

4. Literature Review

4.1 Introductory Remarks

Studying the academic literature has provided both general inspiration for the conducted empirical study and specific arguments that the thesis leans upon. These arguments are presented below:

4.2 Different Definitions of Success

Distinguishing between the different definitions of success for the startup used in different studies, makes it easier to understand the differences in insights that they provide. The distinction is important, as it clearly sets the aim of this thesis apart from similar papers.

Some example definitions follow: Gelderen et al. (2005) define success for the nascent entrepreneur[3] as starting their business; Robb & Robinson (2009) define success for the nascent firm[4] as survival through the first year of operations; Bernstein et al. (2015) define success for young startups as securing early-stage funding.

The success criterion in this thesis is ***Early-Stage Success (ESS)***, defined as an un-calculable but significant improvement in the probability of tangible returns for ESVCs[5]. This improvement is measured three years from the point of initial investment by ESVCs.

4.3 Differences in Employed Methodology

Through the use a randomized field experiment, Bernstein et al. (2015) present suggestive causal evidence of specific investor preference towards certain attributes of the startup. The methodology is employed on an original dataset consisting of roughly 17 000 e-mails sent to roughly 4 500 investors on the AngelList platform[6].

[3] A person undertaking activities to create a business is referred to as the nascent entrepreneur, and the founding effort is called the nascent entrepreneurship (Reynolds & White, 1992).
[4] A firm in its first year of operations is referred to as the nascent firm (Robb & Robinson, 2009).
[5] See *Chapter 6.3.1* for detailed definition of ESS.
[6] AngelList is a platform for startups to raise money online, recruit employees, and apply for funding. In 2017 the platform had profiles of over 70 000 startups (Forbes, 2017).

An empirical study done by Kerr et al. (2010) exploits small changes in the collective interest levels of angel investors. They argue that this can lead to a discrete change in probability of funding for otherwise comparable ventures.

Both Bernstein et al. (2015) and Kerr et al. (2010) use the securing of funding as a success criterion, rather than as a pre-requisite for inclusion in their data-sample. This thesis uses securing of funding as the prerequisite for inclusion in the observed population. Kerr et al. (2010) addresses challenges related to creating a representative sample as unsigned deals are not formally archived, a problem Bernstein et al (2015) solve through use of their methodology.

Bernstein et al. (2015) are left with the problem of considering whether or not investors are right to focus their attention as their results indicate that they currently do. They address this issue, deferring to the ongoing debate between what factors predict future success of an early-stage firm – the idea or the human capital – as discussed by Kaplan et al. (2007). They further defer by explaining how their methodology and the early developmental stage of sample firms prevents them from answering the question directly.

4.3.1 Comments

It is my hypothesis that the framing presented by Kaplan et al. (2007) is misleading, as it implies that either the idea or human capital should be generally dominant. One of the two factors may correlate stronger with success than the other, though it is possible that the secret to minimizing investment risk lies within objectively understanding which combination of characteristics (both human and non-human) most accurately predict future success for early-stage startups.

This thesis defines ESS as its success criteria and the securing of funding as the pre-requisite for inclusion in the sample. These methodological alterations aim among other things to address: ***(1)*** The problem of representative samples faced by Kerr et al. (2010) and ***(2)*** the inability to directly answer whether or not investors are right to focus their attention as they currently do, faced by Bernstein et al. (2015). The methodology employed by this thesis is further discussed in *Chapter 7*.

4.4 The Primary Source of Early-Stage Growth

While the active role of the VC is found to significantly improve the probability of success, it is secondary to the underlying potential of the startup (Sorensen, 2006).

Robb & Robinson (2009) clearly demonstrate that nascent firms with access to relatively high amounts of outside debt (funds without active ownership) are significantly more likely to: ***(1)*** Be in their sample's top revenue group, and ***(2)*** to have hired employees. These results[7] show that even without professional active ownership, the trademark of early stage equity investors, increased capital is linked to increased growth. This result further compliments Sorensen (2006), who's study implies that VC-skill is not the dominant factor when considering investment portfolio success-rate. If a disciplinary effect of debt on management behaviour (Jensen, 1986) does not cancel out the value of active ownership associated with outside equity (Sorensen, 2006) (Cornelli & Yosha, 2003), debt may provide an additional boost to the probability of success for the young firm.

Hochberg et al. (2014) argue that in a large sample of U.S. VC firms, 67.8% of observed VCs lack skill. VCs subsequently go out of business after raising on average 2.7 funds in their lifetime. They argue that the rarity of skill is due in part to high idiosyncratic risk and the large percentage of failed individual portfolio firms.

4.5 Rise of the "Spray & Pray" VC Investment Strategy

The **"Spray & Pray"[8]** investment strategy is a concept presented by Ewens et al. (2017), referring to an observed shift in investment strategy by VCs. The shift has been from larger investments in firms, including governance[9]; to smaller investments into a larger number of firms, including little to no governance[10].

Ewens et al. (2017) argue that this is a consequence of the technological shock created by the introduction of the Amazon Web Services (AWS). AWS eliminated startups need to invest up-front in server-infrastructure, causing significant Capital Expenditure (CapEx) to shift

[7] Only 5% of their sample containing over 4 000 firms held outside equity during the observed period.
[8] The term "Spray & Pray" may be analogous to a term condescendingly used to describe the strategy poorly skilled riflemen use when firing upon their target. The term describes the opposite of a carefully aimed shot: Rapidly firing multiple shots in the appointed general direction, praying that some of the projectiles make contact with the target.
[9] Exemplified by the number of investments where the VC takes (a) board seat(s).
[10] This shift has specifically been observed in tech-startups.

towards later rounds of financing. In affected sectors, VC's make more frequent use of their relatively cheaper abandonment options. This can be seen, as no significant increase in the number of follow-on investments is observed in the study. Investors are demonstrably required to consider and analyse a larger number of potential investments. It is possible, I argue, that the increased number of investments implies that less thorough due diligence is being conducted per startup.

There are more younger and in-experienced founders in light of the technological shock (Ewens, Nanda, & Rhodes-Kropf, 2017). These founders may require more guidance and mentorship than startups run by more experienced entrepreneurs. Gelderen et al. (2005) find that inexperienced founders benefitted significantly from advice and information, though the results are not directly comparable due to the differing definitions of success.

Increased prevalence of the "Spray & Pray" methodology is synonymous with a reduction in the importance of active ownership to early-stage investment. Sorensen (2006) finds that already before the 2006 technological shock studied by Ewens (2017), the underlying potential of the startup is twice as important in predicting success as the direct influence of the VC.

The finding helps to explain the rise of new financial intermediaries, who provide cost effective and scalable mentorship to the startup.

4.6 Current Early-Stage Investor Preferences

Bernstein et al. (2015) find suggestive causal evidence that startups with experienced founding teams have a significantly improved probability of attracting early-stage investment. Experienced investors respond strongly to **positive information about the founding team**, though not to firm traction or existing lead investors. Inexperienced investors react strongly to positive information in all three categories.

This finding shows that investors are currently placing more weight on the founding team than the idea. This contradicts Kaplan et al. (2007) conclusion, that more weight should be placed on the firm idea, due to substantial management turnover during the lifetime of the startup.

Bernstein et al.'s findings (2015) offer a clear example of how investors are currently behaving. The curiosity driving this thesis is not related to current investor behaviour, but to what objectively measurable factors may correlate with success in the early-stage startup.

Further research might compare the results of the conducted analysis with current investor methodology, possibly creating an argument for the altering of current investor behaviour.

4.7 The Staging of VC Investments

This empirical study is predicated on stage investments. It is therefore relevant to reference the regularity of stage-financing in the VC market. The prevalence of stage financing to mitigate investment risk for early-stage investors is clearly documented in the academic literature (Ewens, Nanda, & Rhodes-Kropf, 2017) (Cornelli & Yosha, 2003) (Tian, 2010).

4.8 Concluding remarks

This thesis aims to improve the investor's ability to select the startups most likely to generate liquid returns.

The academic literature considered above suggests that the provision of funds is the primary way in which ESVCs improve the probability of success for startups (Sorensen, 2006) (Robb & Robinson, 2009) (Ewens, Nanda, & Rhodes-Kropf, 2017) (Hochberg, Ljungqvist, & Vissing-Jorgensen, 2014). Consequently, I argue that the most important attribute for VC success is a well-developed mechanism for selecting startups with the highest underlying potential of success.

This thesis empirically tests a mechanism for identifying observable startup characteristics, which correlate with success pre-investment. It may therefore contribute to increasing the efficiency of the Early-Stage Venture Capital market.

5. A Theoretical Argument

5.1 The Value of Measuring Early-Stage Success

Ewens et al. (2017) present in their paper a model for the purpose of simplifying the process of understanding multi-stage financing. I build upon their model to clarify the value of increased quantitative analysis of observable startup characteristics, pre-investment.

They develop the following decision framework:

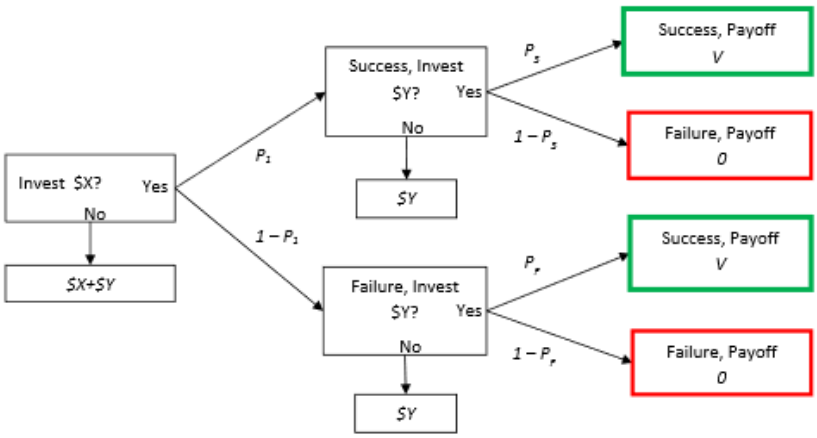

Figure 1: *Extensive Form Representation of the Investor's Game Tree*
(Ewens, Nanda, & Rhodes-Kropf, 2017)

The model (*see Figure 1*) explains the expected returns for the investor given the possible outcomes of a two-stage investment. At the point of initial investment, the investors may choose to invest, or not to invest. If they choose not to invest, they retain $X+$Y but receive no return on investment. If they choose to invest, $X is placed into the firm with a probability P_1 of initial experimentation generating positive information (denoted in the model as success or failure). After the initial experimentation has either succeeded or failed, the investor will again be faced with the choice of a further investment of $Y or retaining this sum. If the investor choses to invest $Y, he will either have done so after initial experimentation has generated positive information, or in spite of it not doing so. The expected probability of success after the second stage investment, is denoted P_S if the initial round of experimentation was a success, or P_F if the initial experimentation was a failure. Success in the second round will lead to a payoff of V.

The aim of this thesis is to measure firm characteristics, known before the investor must choose whether to invest $X, and attempt to correlate them with increased or decreased probability of first-stage success.

5.1.1 Analysis

Success based on e.g. Robb & Robinson (2009) or Bernstein et al. (2015) relates to the ultimate payoff, V. This is difficult to measure within the model parameters, as applying this theoretical framework practically would likely imply adding several steps from initial investment to payoff V. Added complexity arises, as the total amount of years that the funds are invested may well exceed a decade.

Success as defined by this thesis is an indicator of success, rather than Absolute Success (AS). This is disadvantageous as there is still a probability of failure after initial experimentation; it does, however, have the clear advantage of being measurable.

The implications of this may be incorporated into the framework as follows:

P_S & P_F are conditional expectations of a probability, conditional on the success or failure of the initial experimentation. Their denotation is listed below as in the work of Ewens et al. (Ewens, Nanda, & Rhodes-Kropf, 2017).

$$P_S = E[P_2|S]$$

$$P_F = E[P_2|F]$$

$E[P_2]$ is the unconditional expectation of success of the second investment stage.

S & F denote the occurrence of success or failure of the first stage.

Building on this model, I formalize the implied assumption that the true value of P_2 is unobservable (both conditional and unconditional) and change the conditional denotations appropriately:

$$P_S = Unobservable\ true\ probability\ of\ success$$

$$P_F = Unobservable\ true\ probability\ of\ failure$$

$$P'_S \triangleq P_S$$

$$P'_F \triangleq P_F$$

It is assumed that a meaningful conditional estimate of P_2 may only be calculated after the conclusion of the initial experimental phase.

In the model, investors consider staging an investment of \$X, then \$Y in a startup with P_1 probability of initial experimentation being a success. It is assumed that in the second investment stage, $P_S > P_F$. It follows that success after initial experimentation implies a higher likelihood of attaining V than does failure. Therefore, success in initial experimentation, leading to P_S probability of payoff V, ***always*** **maximizes the investors chances of receiving the payoff**. It is thus desirable to maximize the value of P_1, even though success in initial experimentation is only an indicator of future success, not future success in itself.

The probability of receiving payoff V for the investor's investment of $\$X + \$Y$ pre-investment is:

$$P(V) = P_1 * P_S + (1 - P_1) * P_F \quad \text{where } P_S > P_F$$

This implies that the expected value of V is:

$$E(V) = V * P(V)$$

Assume that P_S & P_F are unobservable probabilities (even by approximation) at the point of initial investment. In order to maximize the expected value of V, one must then maximize the value of P_1 , V, or both.

Maximizing V

Maximizing V is an abstract exercise in comparison to the maximization of P_1 . This is because V is an approximated value (though this is no explicitly stated by Ewens et al.). Clarification of notation is again in order:

$$V = Unobservable\ true\ value\ of\ payoff$$

$$V' \triangleq V$$

The true unobservable value of V will vary from startup to startup, as will the observable approximation. This implies that what we are in truth calculating is:

$$P(V') = P_1 * P_S + (1 - P_1) * P_F \quad \text{where } P_S > P_F$$

This implies that the expected true value of V becomes:

$$E(V) = V' * P(V')$$

As P_S and P_F are meaningfully unobservable even by approximation at the point of initial investment, an approximation of V may not be estimated at the point of initial investment. This implies that the probability of payoff $P(V')$ remains incalculable.

It therefore becomes relevant to consider viewing V' as a roughly constant high payoff across a category of startups defined as holding high potential. The implication is that variation in V' across these startups *is not* relevant, but variation in the probability of a payoff occurring *is*:

$$Var(V') \approx 0$$

$$Var(P(V')) \neq 0$$

Within the defined category of high potential startups,
implying a high minimum value for V'

Maximizing V' thus becomes a question of selecting startups that exist within a category defined as *high potential*, rather than filtering individual startups within this category. This thesis does this by only including startups which have been invested in by professional early stage investors managing professionally raised funds.

The remaining uncertainty is defined as variation in the probability of a payoff occurring. As meaningful estimates of P_S & P_F are unobservable within the scope of this thesis, risk at the point of initial investment is minimized by maximizing P_1.

Maximizing P_1

Maximizing P_1 may be done by selecting the startups that have the highest probability of succeeding in the first experimental phase. Clarifying notation follows:

$$P_1 = Unobservable\ true\ probability\ of\ first\ stage\ success$$

$$P_1' \triangleq P_1$$

What fundamentally drives this thesis is a belief that uncertainty in the estimation of P_1' may realistically be decreased. Uncertainty (or variance) in P_1' reduces the probability that ESVCs maximize the P_1 of their investment portfolios. Reduction in the variance of P_1' would therefore reduce risk and increase the efficiency of capital allocation in ESVC portfolios. My argument thus becomes:

The variance of P_1' may plausibly be reduced. Investors should minimize the variation in the estimate of P_1' in order to invest in high-potential startups with the highest P_1.

$$\min_{0<P_1'<1} var(P_1') \;\; \Rightarrow \;\; \max_{0<P_1<1} P_1$$

Lack of startups with high true values of P_1 is an external supply problem. Investors may not through their primary activities increase the supply of startups with a high true probability of success[11]. Reducing the variation of the observable estimate for P_1, however, is a learnable skill, well within the scope of the primary activities of the investor.

I aim to successfully correlate objective and observable factors of individual firms, pre-investment, to success in the first experimentation phase. This will contribute to indicating which startups have the highest true value of P_1. Note that this assumes that $P_S > P_F$, which is not tested empirically by this thesis.

5.2 Concluding Remarks

The direct value of the empirical analysis rests on the assumption that ESS is a significant predictor of AS, as argued by Ewens et al. (2017). It may also be inferred through investor behaviour that they believe this to be true, observed through the prevalence of stage-financing[12]. I have not found any direct test of the validity of this claim, nor am I able to confirm or reject it within the scope of this thesis.

[11] Looking at ways to plausibly increase supply, perhaps by looking deeply at the reasons for startup failure and aiming to formalize methods to avoid those most common on the part of the startup, may contribute to increasing the supply of high-potential startups.
[12] See *Chapter 4.7*.

6. Data Collection Process

> *"…the dearth of data on very early stage firms makes it difficult for researchers to look further back in firms' life histories."*
> (Robb & Robinson, 2009)

6.1 Database Structuring

6.1.1 Introductory Remarks

I argue that a complete ESVC portfolio of firms can viewed as a random sample of firms that receive ESVC investment. The database was therefore constructed by identifying entire investment portfolios, then working to gather data from each individual investment within the portfolio.

The final gathered database consists of three complete portfolios from a single Norwegian ESVC, totalling 32 firms. This includes all failed investments which have been removed from the ESVC's websites. Two firms have incomplete data entries, one of which does not match the population criterion.

Each entry represents the ESVCs first investment into a given startup. They consist of cross-sectional data from the point of initial investment by the ESVC (59 variables per firm). Cross-sectional data has also been gathered 3 years post initial investment (22 variables per firm). The post-investment data has been condensed to the binary variable *Early-Stage Success (ESS)*. ESS is used as the dependent variable in all regression testing. The database considers the entire initial investment, including cases where the ESVC investor is not the only investor in the initial round of financing.

There is idiosyncratic risk at the investor-level. The ESVC may have an investment strategy or general skill-level which produces results that differ markedly from the larger population of Norwegian ESVCs.

6.1.2 The Process

Cold-calling at the ESVC level led to meaningful dialogue with the CFO of an ESVC. Had the ESVC not provided their blessing and been helpful during the data-collection process I would not have succeeded in gathering data from entire portfolios.

A single ESVC provided access to their portfolio-firms. A total of three ESVCs were willing to speak to me through in-person meetings and conferencing-software. The primary benefit of these discussions was related to the structure of the database, its variables, as well as feedback on the hypotheses.

The collection process made use of the following primary tools:

E-mail

Initial contact with portfolio firms was done via e-mail. Upon reaching the individual firm a phone-interview was scheduled.

Phone Interviews

Phone interviews were conducted for all firms in the database, with the exception of certain abandoned firms, for which interviews were instead conducted with a partner who had served as a board member of the firm in question. The interview had two parts:

1. Walk-through of specific questions while inputting responses into the database
2. General conversation relating to the theme of the thesis, receiving general advice on possible challenges, as well as insights they had learned over the course of running their respective firms.

Each part of the interview would usually take roughly 15 minutes, with the first section occasionally exceeding this time period, and the second part varying in both directions of time.

Annual reports

It was necessary to supplement the interview-data with data from annual reports for several reasons, to which I became increasingly aware as the interview process went by:

1. Some data was not recalled or accessibly stored by the CEOs
2. Some of the data provided by the CEOs was an estimate at best
3. Some CEOs were indicating that total interview-time exceeding 30 minutes was more of their time than they were interested in sharing

Roughly 120 annual reports were analysed and double-checked against interview answers. In the cases where they did not match, data from the annual reports were used. This reduced interview-times and improved data-quality.

6.2 Data Sample

The data-sample used in the empirical analysis is limited to Norwegian startups, operating with headquarters in Norway, with Norwegian-private citizens or firms as majority owners, which have received funding from Norwegian ESVCs.

The analysed data-sample makes use of 30 firms and 17 variables, including the ESS variable, defined in *Chapter 7.3.1*. One foreign portfolio firm does not match the population definition. Bar this, the sample contains complete data entries for 97% of the 31 eligible portfolio firms.

6.3 Calculating new fields

The data-sample makes use of two calculated fields, ***(1)*** Early-Stage Success, and ***(2)*** Experienced Founders. They are defined below:

6.3.1 Early-Stage Success

To define whether a firm has achieved ESS or not, a checklist of questions is considered for each firm. If the answer to any of these questions are yes, then the firm is defined as having achieved ESS. The checklist of questions has been considered by academic and market professionals and is included below:

Definitions:

- **Post-investment:** Refers to a point in time three years after the initial ESVC-investment

The minimum requirement for ESS is a YES to <u>at least one</u> of the following questions:

1. Has the firm raised 25MNOK or more post-investment?
2. Has the firm achieved all of the following post-investment?
    a. Raised 10MNOK or more
    b. Raised twice or more of its original investment
    c. Raised capital from new external investors
3. Has the firm achieved all of the following post-investment?
    a. Raised 10MNOK or more post-investment
    b. Increased Valuation by more than 50%
4. Has the firm achieved all of the following post-investment?
    a. Sold goods for more than 2.5MNOK on average per year
    b. Annualized average sales 3y post-/pre-investment have declined by less than 10%
    c. Sold total amount of goods for more than 25MNOK since founding
5. Has the firm achieved all of the following post-investment?
    a. Sold goods for more than 1.5MNOK on average per year
    b. Annualized average sales, 3y post/pre, have increased by more than 200%
    c. Sold total amount of goods for more than 7.5MNOK since founding

Reducing such an uncertain and subjective concept as Early-Stage Success (ESS) of a startup, down to a binary variable requires the use of subjective judgement. In order to efficiently analyse the results of this analysis, one must first familiarize oneself with the definition of ESS used in this thesis.

What Does ESS Attempt to Explain?

ESS describes a state of being for the firm, three years post ESVC-investment, for which the perceived chance of future AS is above a defined threshold.

Lack of ESS is not a direct indication of a firm's failure. The possible conditions of unclear future outlook, negative future outlook, as well as failed investment, are all contained within the category of non-ESS. This thesis focuses on the analysis of firms which achieve ESS and has therefore not made further clarifications of the firms that have not achieved this state.

Concluding Remarks on ESS

A firm has achieved ESS either if it ***(1)*** has raised a significant amount of new capital, with borderline cases holding extra restrictions on increased valuation and the inclusion of new external investors; or ***(2)*** if it has had a stable significant amount of annual sales without significant decline, with borderline cases holding extra restrictions on minimum growth in sales.

I believe that the resulting applied definition of ESS provides a good representation of the state of each individual data-entry 3 years post-investment.

6.3.2 Experienced Founders

Existing academic literature indicates that experienced founders are more likely to be successful than founders who are not experienced (Bernstein, Korteweg, & Laws, 2015) (Gelderen, Thurik, & Bosma, 2005). Experience is, however, not a directly observable variable in the same way as the valuation of the firm at the point of initial investment.

It therefore becomes necessary to create an instrument for experienced founders. I define this though a binary variable, equal to one if the founding team holds both previous founding experience, as well as a PhD.

Roughly two thirds of the data sample founders hold a PhD. Closer study of the impact a PhD holds on the probability of success therefore becomes relevant. Examining whether this combination of experience correlates with ESS is to the extent of my knowledge a novel approach.

6.4 Concluding Remarks

While collecting data I have failed to add a variable observing the number of people in the founding team as well as in management - this has limited the analysis. Another failure related to the structuring of the database is the failure to include a variable indicating whether or not the founder is working full-time on their startup. Gelderen et al. (2005) find that founders working full-time on their startup increase its chance of success. After gathering data and seeing the number of founders not working full-time on their startup, I now believe that there should have been defined a variable checking whether the founders worked full-, part-, or no-time on their startup. This would allow tests regarding whether the finding of Gelderen et al. (2005) remains significant across the changing definition of success.

Certain sample firms were started by non-individual founders, meaning that they were owned entirely by other firms or research groups from nascency.

7. Structuring the Empirical Analysis

7.1 Introductory Remarks

The empirical analysis consists of descriptive statistics and cross-sectional regressions conducted on the data-sample. The structure of the analysis is elaborated upon below.

7.2 Structuring Research Question

The Early-Stages of the capital market is arguably where there is the most risk to investors and firms alike (Kerr, Nanda, & Rhodes-Kropf, 2014) (Cornelli & Yosha, 2003) (Ewens, Nanda, & Rhodes-Kropf, 2017). I argue that this makes performing empirical analysis here particularly important. The aim of this analysis is not to eliminate observed risk, but to contribute to its significant reduction.

The research question empirically tested by this thesis is:

> *Are there statistically significant and relevant correlations between objective and quantifiable variables within startups, known at the point of initial ESVC investment, and Early-Stage Success?*

7.3 Structuring Hypotheses

To test this research question, a set of four hypotheses have been structured, intended to be tested in the cross-sectional regression analysis. The hypotheses were derived from previous academic literature, as well as repeated iterations of rationalizing and seeking feedback from market and academic professionals. An overview of the hypotheses and the rationale behind selecting them follows:

7.3.1 Experienced Founders are Positively Correlated with ESS

This is considered due to the suggestive causal evidence put forward by Bernstein et al. (2015) indicating that experienced founders increase the probability of attracting early-stage investment. Gelderen et al. (2005) also find that experienced founders are more likely to achieve success.

Kerr et al. (2007) show in their paper that management turnover is significant throughout the lifetime of a successful startup. In the observed segment of startup development[13], I hypothesize that experienced founders are correlated with the probability of achieving Early-Stage Success.

Founder experience must be instrumented, as it is not directly measurable. I test this hypothesis through the use of the **"Experienced Founder"** variable[14].

I define a founder as having held significant equity in the firm (above 10% or 2MNOK worth) since its establishment. A founder has also been defined as an individual and not a company. Over the course of the data-collection process, I became aware that different firms (research groups and others) were often a part of the founder mix.

7.3.2 Focus on Diversification is Positively Correlated with ESS

The second hypothesis stems from the argument that diversification reduces idiosyncratic risk, as well as facilitating synergies (assuming that this value outweighs the lost value of specialized focus). It also bases itself on the law of diminishing returns, implying that two engineers may achieve less than an engineer and a BA-Graduate (under the theoretical assumption that their level of skill is held equal).

A startup is faced with considerable amounts of commercial and technological risk (depending on progress and the degree of technological innovation). Consequently, the correct course of action is rarely clear. I hypothesize that diversification may increase the probability of ESS.

This thesis tests for diversification through the use of two variables based on dialogue with market professionals:

1. **Amount of planned independent revenue categories**, defined as the number of individual, independent revenue sources planned at the point of initial ESVC investment. This variable instruments revenue diversification.
2. **Customer portfolio diversification**, defined as the number of paying customers at the point of initial investment. The firm must have sales exceeding 1MNOK in the last twelve months to qualify as having any paying customers. This variable instruments diversification of the customer portfolio.

[13] From the point of initial ESVC investment to three years post-investment.
[14] Defined in *Chapter 7.3.2*.

7.3.3 Wage-Compensated Employees are Positively Correlated with ESS

Gelderen et al. (2005) find that nascent entrepreneurs working full-time on their startup had a significantly higher probability of achieving success than those who did not. Will a similar result hold true if we observe the number of full-time employees at the point of initial investment for the sample firms?

Increasing the monthly costs to the firm may increases the chance of success. The firm may fail quicker than it would have otherwise, though I believe that positive psychological effects may be at play when the sum of monthly wages is increased.

1. The effect of going "all inn" may instil a sense of purpose and urgency to the firm's employees
2. Being compensated with valuable limited resources may instil a sense of responsibility.
3. A firm that compensates its employees with fair wages is at a more advanced stage, than one which does not. Being aware of this fact may have positive psychological effects on the productivity and efficiency of the employees

This thesis does not focus on psychology, though including a consideration of psychological factors which may be in play at the level of the individual involved in the startup seems fitting.

I test this hypothesis through the use of the "**Wage-compensated Full-Time Employees (FTEs)"** variable. Employees are defined as holding less than 5% equity or equity worth less than 2MNOK.

The number of FTEs is listed in each firm's annual report, eliminating insecurity relating to data-collection.

7.3.4 Size of the Initial ESVC Investment Correlates with the Probability of ESS

Robb & Robinson (2009) show that for nascent firms, access to increased levels of capital correlates with increased growth after twelve months of operations. I hypothesize that the same holds true for the size of the initial ESVC funding round.

I test this hypothesis through a variable measuring the **level of raised capital** in the initial ESVC investment. This variable also functions as a control for the level of capital raised for the individual portfolio investment.

7.4 Structuring Control Variables

When tested in the cross-sectional regression, the hypothesis variables act as controls for each other. There are, however, also specific aspects I would like to control for in order to produce a more homogenous sample. Two sets of control variables are explained below:

7.4.1 The Primary Control Variables

There are two primary control variables which are included in the cross-sectional regression. Their purpose is to create a more homogenous sample:

1. **Valuation:** The valuation of the firm controls for the different stages of progression at the point of initial investment. Though valuations are arguably imperfect at the point of initial ESVC investment[15], I hypothesize that their measurement acts as a valuable control for progress.
2. **The Level of Sales in the Last Twelve Months (LTM):** The level of sales LTM is also a control for progression, directly angled towards interaction with the market. Controlling for interaction with the market may usefully narrow the gap between firms with different estimated times to market, as well as different degrees of technological innovation.

7.4.2 The Alternative Control Variables

The alternative setup of control variables substitutes the level of sales LTM for the estimated time to market at the point of initial investment. The valuation variable is kept unchanged:

1. **Valuation:** Unchanged.
2. **Estimated Time to Market:** The estimated time to market effectively controls for the degree to which the individual investment is primarily faced with R&D[16] or commercial & implementation risk. Controlling for the estimated time to market thus makes it easier to compare the risk profiles of firms facing considerably different developmental challenges.

The estimated time to market is an estimate given by the founder/CEO of each individual sample firm. It is used as a secondary alternative due to the added risk of recording error that stems from the data-collection process. The founder/CEO is asked to recall a non-binding estimate given 4-8 years ago, subject to regular change over time. Given perfect information this variable may have been considered superior to the level of sales LTM.

[15] Claim based on my interviews with founders/CEOs of the sample firms.
[16] Research & Development

Estimated time to market might, given a larger sample, be recorded as a discrete variable, listing different categories of estimated time to market[17]. This might also reduce the recording uncertainty present in the current sample.

7.5 Structuring Cross-Sectional Regressions

A total of **six cross-sectional regressions** have been conducted in two sequences of three to test the four hypotheses and the research question. All regressions use ESS as the dependent variable.

As the second hypothesis includes recording risk[18], the sequence of three regressions is run once for each of the two variables described in *Chapter 8.3.2*.

The variables from the four hypotheses are first regressed without the control variables. The regression is then repeated including the Primary Control Variables. It is then run a third time, substituting the Primary Control Variables for the Alternative Control Variables.

7.6 Concluding Remarks

The variables used in the regression analysis are based on data recorded either from annual reports or interviews + LinkedIn data. I identify an added risk stemming from possible recording errors when the variables are based solely on interview answers. This has manifested itself in the dual sequence of the cross-sectional regressions, in an attempt to reduce the possible impact this risk will have on the final results.

The experienced founder variable is not based on information from annual reports, though it has been supplemented with clear and concise information from LinkedIn. As a consequence of this I consider the risk of recording errors in this variable to be low.

[17] E.g.: (1) Less than one year, (2) 1-3 years, (3), 3-5 years, (4) more than five years.
[18] They are based on interview estimates, rather than data from annual reports.

8. Results

8.1 Introductory Remarks

The results section will, ***(1)*** present descriptive statistics of general sample variables, the hypothesis variables and the control variables, ***(2)*** present and discuss the results of the cross-sectional analysis, ***(3)*** comment on each hypothesis in light of the results, and ***(4)*** end with a consideration of the internal and external validity of the study.

8.2 General Descriptive Statistics

Included below is a selection of descriptive statistics. The Appendix includes a complete overview of all gathered variables not included in the empirical analysis.

General Variables								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	%	*n*	*%*	*n*	*%*	*n*	*%*	*n*
n	*100%*	*30*	*53%*	*16*	*47%*	*14*	*6%*	2
High Tech	73%	22	59%	13	*41%*	*9*	13%	4
Med. Tech	23%	7	29%	2	*71%*	*5*	*-42%*	*-3*
Low Tech	3%	1	100%	1	*0%*	-	100%	1
Cust. LTM?	50%	15	60%	9	*40%*	*6*	20%	3

Table 1: *General Variables*

The General Variables table (*Table 1*) presents the distributions of binary general variables ***(1)*** throughout the entire sample, ***(2)*** the sub-sample of ESS observations, ***(3)*** the sub-sample of non-ESS observations, and ***(4)*** the difference between the ESS and non-ESS distributions. Star(s) next to a difference value indicates the degree to which it is statistically significant[19].

The variables concern ***(1)*** whether the firm considered itself to have a high, medium or low degree of technological innovation and ***(2)*** whether there were paying customers in the last twelve months pre-investment (instrumented by whether sales exceeded 1MNOK in the year of investment).

8.2.1 Year Funded

Variable grouping data-sample by the year it was first funded by the ESVC (*see Figure 13*). The majority of firms are funded from 2010-2012. All are funded after the financial crisis and most before the 2014 price collapse of Brent Crude Oil.

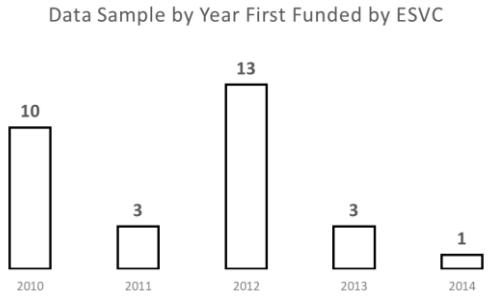

Figure 13: *Column chart showing a count of sample firms by the year that they were first funded by the ESVC.*

[19] $* = p - value < 0.1, ** = p - value < 0.05, *** = p - value < 0.01.$

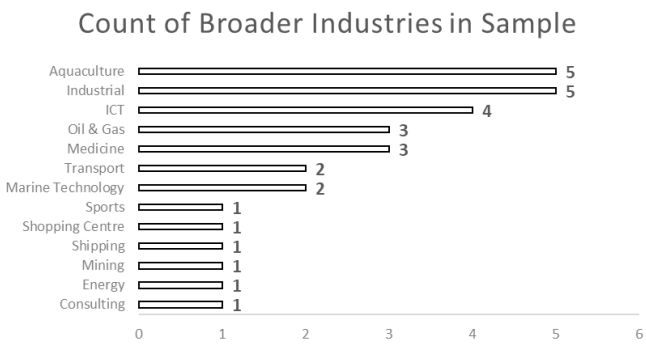

Figure 14: *Bar chart showing a count of the firms in different broader industries in the data-sample.*

8.2.2 Broader Industry

Variable cataloguing different broader industries prevalent in the sample, shown in *Figure 14*.

8.2.3 General Remarks

The general variables are intended to provide general information about the data-sample. Given a larger data sample, they would have been additionally useful as possible control variables, which would have allowed for more specific analysis of the startups.

The general variables gathered by this thesis, clarify the need for strict variable definitions in order to secure that they function as intended.

8.2.4 Founder Background Variables

Separate attention has been given to different founder background variables, in light of the suggestive causal evidence presented by Bernstein et al. (2015), linking strong founding teams to increased likelihood of equity investment. Bernstein et al.'s (2015) results are based on the founding team having attended a prestigious school and/or having relevant work/founding experience. Using the following founder background variables, I hope to observe the founding team with an increased resolution:

Founder Background Variables								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	%	*n*	*%*	*n*	*%*	*n*	*%*	*n*
PhD	63%	19	63%	12	*37%*	*7*	26%	5
Masters	23%	*7*	43%	*3*	*57%*	*4*	-14%	1
Bachelor	10%	3	33%	1	*67%*	*2*	-33%	1
No formal edu.	3%	1	0%	0	*100%*	*1*	-100%	1
Domain XP	33%	10	60%	6	*40%*	*4*	20%	2
Founding XP	37%	11	73%	8	*27%*	*3*	45%	5
BA Graduate	10%	3	100%	3	*0%*	-	100%	3
Experienced Founders	23%	7	86%	6	*14%*	*1*	71%	5

Table 2: *Founder Background Variables*

The Founder Background Variables table (*Table 2)* presents the distributions of binary founder background variables ***(1)*** throughout the entire sample, ***(2)*** the sub-sample of ESS observations, ***(3)*** the sub-sample of non-ESS observations, and ***(4)*** the difference between the ESS and non-ESS distributions. Star(s) next to a difference value indicates the degree to which it is statistically significant[20].

The variables concern different aspects of the founding team: ***(1)*** the highest level of formal higher education (PhD, Masters, Bachelors, or none), ***(2)*** whether there is domain experience in the founding team, ***(3)*** whether the founding team has previous founding experience, ***(4)*** whether the founding team includes a BA graduate, and ***(5)*** whether the founders were experienced per the presented definition in *Chapter 7.3.2*.

As shown in *Table 2*, the sample consists of founding teams which primarily hold a PhD., while roughly a third of observed firms have previous founding or domain experience. The variable instrumenting whether founders are experienced has a clear tilt towards ESS, though the number of observations limits the significance of the results.

Only 10% of the sample founders have a business administration background, though 100% of observed BA Graduate founders achieve ESS. The numbers relating to BA Graduate founders are too small to conduct empirical analysis, though the result does provoke questions regarding the value of BA Graduates in startups.

[20] $* = p - value < 0.1, ** = p - value < 0.05, *** = p - value < 0.01.$

8.3 Hypothesis Variables:

Hypothesis Variables								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Revenue Streams	3.4	*2.0*	4.6	*2.5*	*2.4*	*2.0*	2.2	0.5
Customers LTM	21.5	*1.0*	38.1	*2.0*	*4.4*	*1.0*	33.7	1.0
Wage-Compensated FTEs	5.1	2.5	8.4	4.5	*1.8*	*1.5*	(**) 6.6	3.0
Raised Capital	7,470,313	3,500,000	9,535,294	5,000,000	*5,130,000*	*3,000,000*	4,405,294	2,000,000

Table 3: *Hypothesis Variables*

The Hypothesis Variables table (*Table 3)* presents the distributions of continuous hypothesis variables, considered as the mean & median of ***(1)*** the entire sample, ***(2)*** the sub-sample of ESS observations, ***(3)*** the sub-sample of non-ESS observations, and ***(4)*** the difference between the ESS and non-ESS distributions. Star(s) next to a difference value indicates the degree to which it is statistically significant[21].

Table 3 presents four of the five hypothesis variables[22], find the "Experienced Founder" variable in *Table 2*. See *Table 1* for the distribution between the ESS and non-ESS sub-samples.

An observable pattern in *Table 3* is the presence of large outliers affecting the sample means, as illustrated by the differences between mean and medians. The t-tests conducted on the mean differences between the ESS and non-ESS samples account for this through the inclusion of the standard error term. This makes the significant difference between the Wage-Compensated FTE distributions of ESS and non-ESS firms an interesting finding.

Given the limited sample size, it is possible that recasting some of the continuous variables as discrete variables with 2-4 possible alternatives would reduce the impact of large outliers. The second hypothesis variables are meant to instrument diversification along different axis. One might argue that the marginal reduction in idiosyncratic risk becomes non-significant after a certain number of revenue streams/customers.

[21] $* = p - value < 0.1$, $** = p - value < 0.05$, $*** = p - value < 0.01$.
[22] Described in *Chapter 7.3*.

8.4 Control Variables:

Control Variables								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Valuation (to nearest 000)	22,480,000	11,650,000	28,859,000	11,800,000	15,250,000	11,500,000	13,609,000	300,000
Sales LTM (to nearest 000)	3,724,000	1,000,000	5,439,000	1,200,000	1,895,000	540,000	3,544,000	660,000
TTM	2.9	2.5	2.5	3.0	3.3	2.0	- 0.8	1.0

Table 4: *Control Variables.*

The Control Variables table (*Table 4)* presents the distributions of continuous control variables[23], considered as the mean & median of ***(1)*** the entire sample, ***(2)*** the sub-sample of ESS observations, ***(3)*** the sub-sample of non-ESS observations, and ***(4)*** the difference between the ESS and non-ESS distributions. Star(s) next to a difference value indicates the degree to which it is statistically significant[24].

See *Table 1* for the distribution between the ESS and non-ESS sub-samples.

As seen in *Table 3*, *Table 4* also shows signs of clear outliers (particularly in the ESS distributions). These are observable through the large differences between the mean and median values of the *Valuation* and *SalesLTM* variables. The estimated time to market variable (*TTM*) behaves differently than the two other control variables. There are lower differences between the mean and median, which do not coincide across the different samples.

8.5 Cross-Sectional Regressions (CSRs)

The cross-sectional regressions test the four hypotheses presented and will collectively provide an answer to the research question of the thesis.

The first three regressions are presented using the number of planned revenue streams as the instrument for second hypothesis diversification. The last three are presented using the number of customers in the last twelve months as the instrument for second hypothesis diversification[25].

[23] See *Chapter 7.4* for variable definitions.
[24] $* = p - value < 0.1$, $** = p - value < 0.05$, $*** = p - value < 0.01$.
[25] The methodology is further explained in *Chapter 7.5*.

CSR1: Revenue Stream Diversification + No Controls

```
Model 1: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -3.47295      1.33758      -2.596   0.0094   ***
  ExperiencedFound~     4.58585      2.12733       2.156   0.0311   **
  Revenuestreams        0.304022     0.219692      1.384   0.1664
  Employees10           0.856054     0.395895      2.162   0.0306   **
  RaisedCapital        -1.13862e-07  8.00484e-08  -1.422   0.1549

Mean dependent var    0.533333   S.D. dependent var    0.507416
McFadden R-squared    0.510967   Adjusted R-squared    0.269744
Log-likelihood      -10.13652    Akaike criterion      30.27305
Schwarz criterion     37.27903   Hannan-Quinn          32.51432

Number of cases 'correctly predicted' = 25 (83.3%)
```

CSR1 includes a total of four variables and is run as a logit regression without controls.

Note that the "RaisedCapital" variable should be multiplied by 1", as it represents values that changes in the millions. Its altered coefficient becomes: 0.114.

CSR2: Revenue Stream Diversification + Primary Controls

```
Model  2: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -4.49110      1.93483      -2.321    0.0203   **
  ExperiencedFound~     5.87600      3.18751       1.843    0.0653   *
  Revenuestreams        0.546380     0.431096      1.267    0.2050
  Employees10           1.71942      0.713241      2.411    0.0159   **
  RaisedCapital        -5.81851e-09  3.11986e-07  -0.01865  0.9851
  Valuation            -9.84181e-08  1.71116e-07  -0.5752   0.5652
  SalesLTMpre          -7.10637e-07  3.87918e-07  -1.832    0.0670   *

Mean dependent var    0.533333   S.D. dependent var    0.507416
McFadden R-squared    0.630884   Adjusted R-squared    0.293171
Log-likelihood      -7.650930    Akaike criterion      29.30186
Schwarz criterion     39.11024   Hannan-Quinn          32.43965

Number of cases 'correctly predicted' = 28 (93.3%)
```

CSR2 includes a total of six variables and is run as a logit regression. In addition to the four hypothesis variables, it makes use of the primary control variables: "Valuation" & "SalesLTMpre".

Note that the "RaisedCapital", "Valuation" and "SalesLTMpre" variables should be multiplied by 1", as they represent values that change in the millions. Their altered coefficients become: -0.006, -0.10, -0.712, respectively.

CSR3: Revenue Stream Diversification + Alternative Controls

```
Model  3: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -3.79197      1.87469      -2.023    0.0431   **
  ExperiencedFound~     4.66145      2.59984       1.793    0.0730   *
  Revenuestreams        0.389688     0.249084      1.564    0.1177
  Employees10           0.854891     0.429968      1.988    0.0468   **
  RaisedCapital        -3.47225e-08  2.65408e-07  -0.1308   0.8959
  Valuation            -5.50659e-08  1.30256e-07  -0.4228   0.6725
  TTMarket              0.145896     0.219325      0.6652   0.5059

Mean dependent var    0.533333   S.D. dependent var    0.507416
McFadden R-squared    0.534453   Adjusted R-squared    0.196741
Log-likelihood      -9.649716    Akaike criterion      33.29943
Schwarz criterion     43.10781   Hannan-Quinn          36.43722

Number of cases 'correctly predicted' = 25 (83.3%)
```

CSR3 includes a total of six variables and is run as a logit regression. In addition to the four hypothesis variables, it makes use of the alternative control variables: "Valuation" & "TTMarket"[26].

Note that the "RaisedCapital" and "Valuation" variables should be multiplied by 1", as they represent values that change in the millions. Their altered coefficients become: -0.035, -0.055, respectively.

[26] Estimated Time to Market at point of initial investment.

CSR4: Customer Portfolio Diversification + No Controls

```
Model 4: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -2.07213      0.917705    -2.258   0.0239   **
  ExperiencedFound~     3.21837      1.73454      1.855   0.0635   *
  CustLTM1             -0.0384350    0.0262969   -1.462   0.1439
  Employees10           0.757643     0.349631     2.167   0.0302   **
  RaisedCapital        -8.50966e-08  7.45689e-08  -1.141   0.2538

Mean dependent var   0.533333   S.D. dependent var   0.507416
McFadden R-squared   0.400809   Adjusted R-squared   0.159586
Log-likelihood      -12.41985   Akaike criterion     34.83971
Schwarz criterion    41.84570   Hannan-Quinn         37.08098

Number of cases 'correctly predicted' = 25 (83.3%)
```

CSR4 includes a total of four variables and is run as a logit regression without controls.

Note that the "RaisedCapital" variable should be multiplied by 1", as it represents values that changes in the millions. Its altered coefficient becomes: -0.085.

CSR5: Customer Portfolio Diversification + Primary Controls

```
Model 5: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -4.17215      1.96077     -2.128   0.0334   **
  ExperiencedFound~     5.55470      2.63650      2.107   0.0351   **
  CustLTM1             -0.130480     0.0642658   -2.030   0.0423   **
  Employees10           1.96348      0.816403     2.405   0.0162   **
  RaisedCapital        -4.05880e-07  2.31582e-07  -1.753   0.0797   *
  Valuation             1.17888e-07  8.35777e-08   1.411   0.1584
  SalesLTMpre          -8.09080e-07  4.21778e-07  -1.918   0.0551   *

Mean dependent var   0.533333   S.D. dependent var   0.507416
McFadden R-squared   0.559440   Adjusted R-squared   0.221728
Log-likelihood      -9.131786   Akaike criterion     32.26357
Schwarz criterion    42.07195   Hannan-Quinn         35.40136

Number of cases 'correctly predicted' = 26 (86.7%)
```

CSR5 includes a total of six variables and is run as a logit regression. In addition to the four hypothesis variables, it makes use of the primary control variables: "Valuation" & "SalesLTMpre".

Note that the "RaisedCapital", "Valuation" and "SalesLTMpre" variables should be multiplied by 1", as they represent values that change in the millions. Their altered coefficients become: -0.406, 0.118, -0.809, respectively.

CSR6: Customer Portfolio Diversification + Alternative Controls

```
Model 6: Logit, using observations 1-30
Dependent variable: EarlyStageSuccess
Standard errors based on Hessian

                      coefficient   std. error      z      p-value
  ------------------------------------------------------------------
  const                -3.01720      1.59089     -1.897   0.0579   *
  ExperiencedFound~     3.98994      2.12785      1.875   0.0608   *
  CustLTM1             -0.0628781    0.0422112   -1.490   0.1363
  Employees10           0.940277     0.451610     2.082   0.0373   **
  RaisedCapital        -2.25359e-07  1.83194e-07  -1.230   0.2186
  Valuation             6.01788e-08  6.88752e-08   0.8737  0.3823
  TTMarket              0.0728323    0.188247     0.3869  0.6988

Mean dependent var   0.533333   S.D. dependent var   0.507416
McFadden R-squared   0.419490   Adjusted R-squared   0.081778
Log-likelihood      -12.03263   Akaike criterion     38.06526
Schwarz criterion    47.87364   Hannan-Quinn         41.20305

Number of cases 'correctly predicted' = 23 (76.7%)
```

CSR6 includes a total of six variables and is run as a logit regression. In addition to the four hypothesis variables, it makes use of the alternative control variables: "Valuation" & "TTMarket"[27].

Note that the "RaisedCapital" and "Valuation" variables should be multiplied by 1", as they represent values that change in the millions. Their altered coefficients become: -0.225, 0.060, respectively.

[27] Estimated Time to Market at point of initial investment.

8.6 General Comments

Brief comments will first be presented for CSR 1-3, then for CSR 4-6. The primary cross-sectional regressions, CSR2 & CSR5, are then compared. Insignificant results will be discussed in relation to the consideration of the four hypotheses in *Chapter 9.7*.

8.6.1 CSR1-3:

There persists a positive correlation between the **number of FTEs** and the probability of achieving ESS. The same is true for the binary presence of experienced founders. While the number of FTEs is consistently significant at the 5% level, **experienced founders** vary in significance between the 5- and 10% level. The presence of these variables indicates an increase in the probability of achieving ESS.

The **constant term** remains significant at the 1- to 5% level with a negative coefficient, indicating an initial tilt towards probability of non-ESS.

The **level of sales LTM** is an instrument to control for whether the firm is primarily faced with technological or commercial risk in the observed period. The variable is statistically significant at the 10% level with a slightly negative coefficient in CSR2. This implies that sample firms with higher levels of sales in the last twelve months pre-investment are slightly less likely to have achieved ESS within three years. A possible interpretation is that the commercial risk is statistically significantly higher than the technological risk in the observed three-year period. When holding slopes constant at the mean, an increase in sales LTM of 1MNOK is estimated to reduce the probability of ESS by 3.58%[28].

8.6.2 CSR4-6:

Replacing the instrument for diversification, the pattern from CSR1-3 is again observable for the **number of FTEs** and the presence of **experienced founders**. These two variables thus remain statistically significant positive predictors of ESS through all 6 iterations of the cross-sectional regressions.

[28] This effect is subject to diminishing returns as the CSRs are run as logit regressions.

The **constant term** remains significant through all three regressions, with comparable negative statistics. The constant term thus becomes a statistically significant indicator of an initial tilt towards non-ESS through all six regression iterations.

CSR5 additionally yields statistically significant results for the **customer diversification variable**, the **level of raised capital variable**, as well as the **level of sales LTM** pre-investment. These are variables likely to increase as the firm enters an implementation/commercialization phase. All coefficients are negative, possibly indicating that the implementation/commercialization phase holds more immediate risk in the observed three-year period than does the phase of technological development.

Collinearity issues are found in CSR5 & 6 between the "CustLTM", "Employee10" and "Valuation" variables. Replacing the instrument used for testing the second hypothesis with its counterpart remedies all cases of collinearity.

8.6.3 CSR2 & 5:

The focus on the primary control regressions is due to the mentioned recording error risk present in the use of the TTM estimate. The alternative control regressions also both explain a significantly lower portion of the sample variance than do the primary control regressions, when considering the adjusted R^2.

The variations in the first and second sets of regressions do not impact the significance of the number of FTEs or the presence of experienced founders.

A difference between the two iterations is the significance of the diversification instrument in CSR5. It is possible that customer diversification captures some of the same effect as the risk variable ("SalesLTM"). I argue that the number of customers is likely to rise as the company enters the commercialization phase, thus possibly creating overlap.

Due to the collinearity issue between the "CustLTM", "Employee10" and "Valuation" variables, the CSR2 becomes the primary regression analysed when considering the four hypotheses.

8.7 Comments on Hypotheses

The considered results are achieved in a regression run alongside three other hypothesis variables. The results are additionally controlled for the difference in progress by individual portfolio investments through the use of the "Valuation" and "SalesLTM" variables[29].

Hypothesis 1: Experienced Founders are Correlated with ESS

After having concluded the cross-sectional regression analyses, **I am forced to reject the null hypothesis that this is not the case at the 10% level of significance.** The variable is found to hold statistical and practical significance in all six iterations of the analysis. The presence of experienced founders **acts as a significant predictor** of whether or not a sample firm will have achieved Early Stage Success three years post-investment. This supplements the results of Bernstein et al. (2015), who present suggestive causal evidence that early-stage investors are more likely to invest in firms with strong founding teams. It is a new finding, as Bernstein et al. (2015) make no conclusive comment on whether this focus is good or bad. This finding indicates that experienced investors appear to be right to focus their attention as Bernstein et al. (2015) observe that they do.

Holding the variable slopes constant at their means for CSR2, having an experienced founder team is estimated to increase the probability of achieving ESS by 18.1%. The result supplements the findings of Bernstein et al. (2015), who propose suggestive causal evidence of a link between experienced founding teams and success in attracting early-stage investment.

Hypothesis 2: Focus on diversification is positively correlated with ESS

After having concluded the cross-sectional regression analyses, **I cannot reject the null hypothesis that this is not the case**. Diversification is tested along two dimensions, the only iteration which produced significant results (CSR5) suffered from multiple strong collinearity issues. In the remaining five iterations none of the diversification variables produced statistically significant results.

[29] Described further in *Chapter 7.4*.

Several possible alternative diversification instruments gathered, unfortunately suffer from clear cases of OVB[30] – the variables were therefore excluded from the analysis. The second hypothesis has clearly shed light upon the need to be strict when defining variables, the lack of which may have contributed to non-significant results.

Hypothesis 3: The number of wage-compensated employees[31] is positively correlated with ESS

After having concluded the cross-sectional regression analyses, **I am forced to reject the null hypothesis that this is not the case at the 5% level of significance.** The variable is found to hold statistical and practical significance in all six iterations of the analysis. The number of wage-compensated employees **acts as a significant predictor** of whether or not a sample firm will have achieved Early-Stage Success three years post-investment. This supplements the findings of Gelderen et al. (2005) who find that nascent entrepreneurs who work full-time on the startup are more likely to succeed. It is a new finding, as Gelderen et al. (2005) observe nascent entrepreneurs aiming to create a firm, while I observe ESVC funded startups aiming to create value over a three-year period.

Holding the variable slopes constant at their means for CSR2, increasing the number of FTEs by one is estimated to increase the probability of achieving ESS by 8.7%[32].

Hypothesis 4: The Size of the Initial ESVC Investment Correlates Positively With ESS

After having concluded the cross-sectional regression analyses, **I cannot reject the null hypothesis that this is not the case**. In the observed sample, the level of raised capital appears not to be a significant predictor of ESS.

This does not support the findings of Robb & Robinson (2009), who find that nascent firms with higher access to capital are more likely to success. The results are, however, not directly comparable. This is both in terms of median investments per firm, and as Rob & Robinson (2009) primarily measure the effect of debt, while I am measuring the size of equity investment.

[30] Listen in *Chapter 8.8.2*, paragraph relating to OVB.
[31] Being a wage-compensated employee is defined as holding less than 5% equity or equity worth less than 2MNOK in the observed portfolio firm.
[32] This effect is subject to diminishing returns as the CSRs are run as logit regressions.

8.8 Possible Implications of Results

From the perspective of the firm, the results do not become practically applicable until one can argue that a causal effect is present. This follows as the firm would have to alter its own behaviour – for example by hiring more employees than originally planned. If there is no causal connection between dependent and independent variable, the firm would be increasing its costs without improving its chances of success. This limits the usefulness of the results for the early-stage firm.

From the ESVC perspective, the results become practically applicable when a strong claim of correlation is made with a representative sample. While ESS does not guarantee AS, I argue[33] that achieving ESS is equal to an unidentified, significant improvement in the expected probability of AS. The ESVC might also use the results to improve corporate governance for portfolio investments, though this would require a causal connection to be present.

8.9 Concluding Remarks

The results may be interpreted as indicative of present correlations. They are not robust enough to argue for the presence of causal connections. More analysis would have to be conducted to consider the true internal validity of the study.

The results appear useful in the following ways: ***(1)*** The results indicate that the methodology may be useful, assuming that ESS firms have a higher probability of achieving AS than non-ESS firms, ***(2)*** future research need not be as exploratory in nature and ***(3)*** the implied null hypothesis of the research question has been rejected. The research question of this thesis is:

> *Are there statistically significant and relevant correlations between objective and quantifiable variables within startups, known at the point of initial ESVC investment, and ESS?*

Having concluded the empirical research, **I am forced to reject the null hypothesis that this is not the case at the 5% level of significance**. This does not necessarily imply that the null hypothesis is wrong, though the empirical analysis conducted forces its temporary rejection.

---

[33] See *"The Value of Measuring Early-Stage Success", Chapter 5.1*.

9. Discussions

9.1 Early-Stage Success

Consideration of the ESS variable is critical in order to be able to properly evaluate the usefulness of the employed methodology. I therefore consider, ***(1)*** the theoretical value of ESS, ***(2)*** the practical definition of ESS, employed by this thesis:

9.1.1 Considering the Theoretical Value of ESS

Startups, from their first interaction with professional equity investors to their transformation into a mature commercial enterprise, is a considerably heterogeneous population. I argue that sub-sampling this population into stages should reduce its heterogeneity, allowing for increased resolution of study. Though this in parallel reduces the external validity of the results, the increased decision relevant information for ESVCs makes further study an interesting prospect.

9.1.2 Considering the Practical Definition of ESS

A caveat relating to the definition of ESS[34], is in my opinion the somewhat arbitrary quantitative requirements of sales/raised capital to be considered an ESS firm[35]. Furthermore, the two variables used to construct the ESS instrument cannot themselves be empirically tested for correlation with ESS.

9.2 Data Scarcity

The definition of the ESS variable opens for detailed study, which may be quantitatively tested within a practical framework. If high-quality representative data-samples cannot be obtained, however, the framework is rendered impotent.

Data collection of sufficiently high quality appears hard to attain. I base this statement on the references made in the observed academic literature (Bernstein, Korteweg, & Laws, 2015)

[34] See *Chapter 7.3.1*.
[35] The quantitative requirements have been considered reasonable by independent academic and market professionals.

(Robb & Robinson, 2009) (Gelderen, Thurik, & Bosma, 2005) (Kerr, Lerner, & Schoar, 2010), as well as my own experiences.

An implication of data scarcity is that it may negatively impact the volume of conducted research. I hope that this risk may be partially mitigated by the research done by this thesis, hopefully increasing the motivation to gather required data in spite of its scarcity.

9.3 Internal & External Validity

9.3.1 A Discussion of Internal Validity

Several threats to internal validity have been identified:

The data-sample is of limited size: This limits the number of variables that may be regressed simultaneously and renders some inadequate for analysis, increasing the risk of Omitted Variable Bias (OVB). The direct risk of OVB is that the current results of the empirical analysis may be misleading, as an omitted variable may alter the statistical significance of the included variables. OVB is the largest identified threat to internal validity.

Recording risk: As the sample investments primarily took place between 2010-2012, there is exists recording risk for variables based solely on interviews conducted in late 2018.

The population may have changed significantly since the observation period: The sudden 2014 fall in Brent Crude Oil prices impacted Norway more profoundly than the average western market, as oil is Norway's primary export good. Combined with the growing focus on the environmental impact of fossil fuels, this was complicit in increasing the national focus on a Norwegian post-oil economy. Norway has since seen a divestment over time from the oil industry and a significantly increased focus on the startup sector (Abelia, 2017).

Lack of ESVC controls: In order to isolate the impact of different startup variables from the effects of governance, the following three things should be controlled for: ***(1)*** ESVC general skill level, ***(2)*** ESVC learning effect over time, ***(3)*** ESVC specialization.

(1) Variance in ESVC general skill level implies that the behaviour of a single Norwegian ESVC may not be representative of the larger population. On the contrary, using data from only one ESVC partly mitigates the risk that the study is capturing the effects of different ESVC skill levels, as opposed to inherent differences in the individual startup. If data from

multiple ESVCs were to be aggregated into a single sample, a control for variance of ESVC skill should be included.

(2) The ESVC may improve its skill over time through a learning effect. Splitting the sample in half by the year of funding, then testing the difference in ESS probability, indicates that there is no statistically significant learning effect in the sample. Given a larger data-sample, a control for this possible effect should be included.

(3) The thesis does not control for ESVC specialization, that certain startups may be a better fit for certain ESVCs. This deals a significant blow to the internal validity of the study, when considering whether the results of the sample are representative for the larger population as a whole.

9.3.2 A Consideration of External Validity

I consider different relevant comparable **geographical markets**, ranked by their perceived comparability, as ESVC-funded startups in:

1. Scandinavian countries
2. Nordic countries
3. Northern-European countries
4. Northern-European & North American countries
5. Northern-European, North American & EU countries
6. Western countries
7. Non-Western countries

Considering the validity of the study in the same market, but over longer **periods of time**, other concerns become more relevant. Ewens et al. (2017) argue that a consequence of the Amazon Web Services rollout was the development of the "Spray and Pray" investment strategy[36]. Significant alterations of ESVC investment strategies may reduce or eliminate the external validity of this study over time.

The empirical analysis is **ecologically valid**, as it is entirely based on observational data. There has been no experimentational recreation of the population over time.

[36] See *Chapter 4.5*.

10. Future Research Steps

Though half of the hypotheses failed to reject their null hypotheses, each failure revealed valuable insights relating to the improvement of future research. This section considers future research steps at the macro-, methodological- and micro-level. It is not an exhaustive list. Not all research questions are elaborated upon or followed by suggestions for hypotheses.

10.1.1 Macro-Level Future Research Steps

Does Early-Stage Success (ESS) correlate with Absolute Success[37] (AS)? This would be difficult to test empirically, as there might be more than a decade from the point of ESS, until one can consider whether AS has occurred. The test would indicate whether ESS correlates significantly with AS, also observing the significance of its coefficient. Significant results would empirically complement the theoretical argument that ESS is a valid predictor of AS.

Consideration of institutional cooperation. The empirical literature indicates that meaningful data is scarce concerning startups (e.g. (Bernstein, Korteweg, & Laws, 2015) (Robb & Robinson, 2009)); this has been reflected by my own experience. A meaningful contribution to future research may be the further development of institutional cooperation between academic and market institutions.

Efficient categorization of startup data, in both the Seed and Venture Capital markets, would be a valuable indirect contribution to future research. An example might be the establishment of an institutional cooperation between The Norwegian School of Economics (NHH) and the Nordic Venture Network[38]. A top-down approach to the collection of data might effectively create a useful database for future research.

10.1.2 Methodological Future Research Steps

These future research steps are aimed at improving the methodology employed in this thesis, following results may then be compared with results using the present methodology. Recommended methodological alterations are listed below:

[37] Absolute Success: Liquid returns on investment for ESVC investors.
[38] The club of leading venture capital investors in the Nordic region: https://nordicventurenetwork.com/

Address internal validity issues. The discussed internal validity issues should be addressed. Significant results should then be re-tested in order to observe whether or not the results are impacted by the alterations.

Stricter variable definition. Further significant results may have emerged, had certain variables in the sample been more strictly defined. Effort should be paid to ensure that necessary variable strictness keeps significant results from being lost. This may be applied both to testing of new variables, as well as to variable re-testing.

Increased sample size. The thesis employs a sample of 30 observations. Increasing this would facilitate further sub-sampling and controlling of the data, increasing flexibility in the analysis. There were several instances during the empirical testing, where the data indicated that further sub-sampling might have produced valid contributions to the research question.

10.1.3 Micro-Level Future Research Steps

These future research steps primarily aim to suggest new research questions and hypotheses. Based on what research questions and hypotheses are tested, different alterations to the methodology may be needed. In any scenario, the observed threats to internal validity should be addressed.

The future research steps are structured in five categories, listed below:

Advisory Services

Future research steps relating to the possible impact of different advisory services.

Does the use of an Advisory Board (AB) correlate with an increased probability of achieving ESS? The database shows that less than 25% of observed firms had made use of any forms of ABs at the point of initial investment. With a larger data-sample one may empirically test whether the use of ABs correlates with ESS. AB exemplifies a variable that would benefit from a stricter variable definition.

Does the use of financial advisory services correlate with the probability of ESS? The database shows that less than 20% of observed firms had employed external professional financial services at the point of initial investment. With a larger data-sample one may empirically test whether the use of external financial advisory services correlates with ESS. Future study would benefit from stricter variable definition.

Founders

Future research steps relating to attributes of the founders of target firms.

Are founders working full-time in the startup positively correlated with ESS? At the point of initial investment, database founders range from being an FTE, to complete non-involvement without stock in the company. To test whether increased founder involvement correlates with ESS, a variable measuring founder involvement in the firm could be defined. This would have implications for how ESVCs should value the founder when considering their initial investment into a target firm. This research question builds on comparable findings in existing academic literature[39].

Do specific experienced founder traits correlate more or less with ESS? The empirical study conducted by this thesis implies that previous founding experience may correlate positively with ESS. Can a sub-sample of founders with previous founding experience find more specific traits that correlate more or less with ESS? E.g.: ***(1)*** Previous success, ***(2)*** Number of previous firms founded, ***(3)*** number of previous firms run, ***(4)*** founder age.

Is 'Domain Experience' a statistically significant predictor of ESS? This may be either a standalone variable, or e.g. in combination with previous founding experience, forming an alternative definition for "Experienced Founder". One may then test whether education or work experience is the most important predictor of ESS. Domain experience may benefit from stricter variable definition.

Study based around the value of the founding team. If different traits within the founding team can be found to correlate significantly with an increased probability of ESS, then an argument can be made for the valuation of these traits. Examples of hypotheses to test are listed below:

- Does being a Business Administration Graduate founder have a significant impact on the probability of ESS?
- Is there a statistically significant difference in the probability of achieving ESS between firms founded by individuals, versus firms founded by other firms or research groups?

Ownership/Valuations

Future research steps relating to the possible impact of different levels of ownership and/or valuations.

[39] FTE Founders are more likely to succeed in the nascent entrepreneurship stage than founders who are not FTE (Gelderen, Thurik, & Bosma, 2005).

Does the ESVC share of the firm, post-money at initial investment, correlate with ESS?

Employment

Future research steps relating to attributes and the impact of employees in target firms.

Do different classes of employees impact the probability of ESS differently? Classifying different groups of employees would allow testing for whether all employee classes significantly improve the probability of ESS. Different possible classifications of employees: ***(1)*** Founding Employee, ***(2)*** Management Employee, ***(3)*** Non-Management Employee ***(3.1)*** Administrative Employee, ***(3.2)*** Operational Employee.

Diversification

Future research steps relating to the possible impact of diversification

Do the number of people in the founding team or management correlate with the probability of ESS? In attempting to answer this research question, one might test several hypotheses, some examples given: ***(1)*** One versus several team members, ***(2)*** splitting the distribution into quartiles and testing each quartile against the entire sample average.

Does having multiple Areas of Expertise in the founding team and/or management correlate with ESS?[40] A stricter definition of Areas of Expertise should be employed before attempting to answer this question.

Do firms with both sexes represented in the founding team and/or management have an increased probability of achieving ESS?[41]

[40] Requires controlling for the number of people in the founding team or management.
[41] Requires controlling for the number of people and AoEs in the teams.

11. Conclusion

The Norwegian startup economy is in rapid growth (Abelia, 2017). Meanwhile, technological advances are shown to alter investment strategies and increase the prevalence of early-stage investments with limited governance (Ewens, Nanda, & Rhodes-Kropf, 2017).

The primary source of early-stage growth appears to be funding combined with the underlying potential of the firm, rather than the added value of governance (Sorensen, 2006). This implies that effective methods for identifying startups with high underlying potential is a primary method of increasing ESVC returns over time. Observation of experienced early-stage investor preferences may aid less experienced investors to increase returns (Bernstein, Korteweg, & Laws, 2015). Such observation fails, however, to answer the question of whether investors are right to focus their attention as they do.

The here conducted empirical analysis indicates possible correlations between pre-investment startup traits and Early-Stage Success (ESS). Though the results provide limited direct decision-relevant information for ESVCs, further improved research may reach more robust conclusions. If future discovered relationships could be reliably quantified, the resulting reduced risk in ESVC investment decisions could improve expected long-run average returns.

The hope is that this thesis may contribute to increasing the total volume of available venture capital in Norway. This would be useful at a time when the Norwegian startup economy is experiencing rapid growth (Abelia, 2017).

12. Bibliography

(2017, October 19). Retrieved from Forbes: https://www.forbes.com/sites/jeffkauflin/2017/10/19/12-websites-to-jump-start-your-career-in-2018/

Abelia. (2017, June). *Growth in the Norwegian Startup Market*. Retrieved from https://www.abelia.no/politikk/valg-2017/arbeidsplasser/nyheter/1500-vekstbedrifter-som-lofter-norge/

Bernstein, S., Korteweg, A., & Laws, K. (2015). Attracting Early Stage Investors: Evidence from a Ranomized Field Experiment. Stanford Graduate School of Business.

Cornelli, F., & Yosha, O. (2003). Stage Financing and the Role of Convertible Securities. *Available at SSRN: https://ssrn.com/abstract=322460 or http://dx.doi.org/10.2139/ssrn.322460*.

Ewens, M., Nanda, R., & Rhodes-Kropf, M. (2017). Cost of Experimentation and the Evolution of Venture Capital. Harvard Business School.

Gartner, W. B. (1985). A Conceptual Framework for Describing the Phenomenon of New Venture Creation. *Academy of Management Review 10(4)*, 696-706.

Gelderen, M. v., Thurik, R., & Bosma, N. (2005). Success and Risk Factors in the Pre-Startup Phase. *Small Business Economics*, 365-380.

Harris, R. S., Jenkinson, T., Kaplan, S. N., & Stucke, R. (2014). Has Persistence Persisted in Private Equity? Evidence from Buyout and Venture Capital Funds. Darden Business School.

Hart, O. D., & Moore, J. (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy, 98*, 1119-1158.

Hochberg, Y., Ljungqvist, A., & Vissing-Jorgensen, A. (2014). Informational Holdup and Performance Persistence in Venture Capital. *Review of Financial Studies 27*, 102-152.

Jensen, M. C. (1986, May). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *The American Economic Review, 76*(2), 323-329.

Kaplan, S. N., & Schoar, A. (2005). Private Equity Performance: Returns, Persistence and Capital Flows. *Journal of Finance 60*, 1791-1823.

Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2007). . .

Kaplan, S. N., Sensoy, B. A., & Strömberg, P. (2007). Should Investors Bet on the Jockey or the Horse? Evidence from the Evolution of Firms from Early Business Plans to Public Companies.

Kerr, W. R., Lerner, J., & Schoar, A. (2010). The Consequences of Entrepreneurial Finance: A Regression Discontinuity Analysis. *Harvard Business School Entrepreneurial Management Working Paper No. 10-086*.

Kerr, W., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as Experimentation. *Journal of Economic Perspectives, 28*, 25-48.

Korteweg, A. G., & Sorensen, M. (2014). Skill and Luck in Private Equity Performance. Rock Center for Corporate Governance at Stanford Working Paper No. 179.

Mankiw, G. N., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics 107*, 407-437.

Manso, G. (2015). Experimentation and the Returns to Entrepreneurship. Working paper, UC Berkeley.

Myers, S., & Majluf, N. (1984). Financing and investment decisions when firms have information markets do not have. *Journal of Financial Economics, 13*(2), 187-221.

Rajan, R. G. (2012). Presidential Address: The Corporation in Finance. *Journal of Finance, 57*, 1173-1217.

Rajan, R. G., & Zingales, L. (2001). The Firm as a Dedicated Hierarchy: A Theory of the Origins and Growth of Firms. *Quarterly Journal of Economics, 116*, 805-851.

Rauch, A., & Frese, M. e. (2000). Psychological Approaches to Entrepreneurial Success: A General Model and an Overview of Findings. *International Review of Industrial and Organisational Psychology, 15*, 101-143.

Reynolds, P., & White, S. (1992). "Finding the Nascent Entrepreneur: Network Sampling and Entrepreneurship Gestation". *Frontiers in Entrepreneurship Research*. Wellesley: Babson College.

Robb, A. M., & Robinson, D. T. (2009, Feb). The Capital Structure Descisions of New Firms. UC Santa Cruz.

Schumpeter, J. (1934). The Theory of Economic Development. Cambridge, MA: Harvard University Press.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics 90*, 65-94.

Sorensen, M. (2006). How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital. *Available at SSRN: https://ssrn.com/abstract=878692 or http://dx.doi.org/10.2139/ssrn.878692*.

Startup Genome. (2018). *The Global Startup Ecosystem Report.*

Tian, X. (2010). The Causes and Consequences of Venture Capital Stage Financing. *Journal of Financial Economics (JFE)*.

Tiessen, J. H. (1998). Individualism, Collectivism, and Entrepreneurship: A Framework for International Comparative Research. *Journal of Business Venturing 12*, 367-384.

Welch, I. (1992). Sequential Sales, Learning, and Cascades. *Journal of Finance 47*, 695-732.

Wernerfelt, B. (1984). A Resource-Based View of the Firm. *Strategic Management Journal*, 171-180.

13. Appendix

The Appendix includes a collection of all the variables gathered during the data-collection process, not considered in the main empirical analysis. The first section defines each variable, the second presents tables of descriptive statistics for the different classes of variables.

13.1 Definitions

Below are the definitions of all gathered variables, not included in the main empirical analysis. Certain variables are based entirely on interview data. These variables are marked with a star in the variable name.

13.1.1 General

- **Sector** - *Categorical variable, string.* What sector does the firm primarily operate within?
- **Function** - *Categorical variable, string.* Does the firm primarily fulfil an operative or support function in its broader industry and sector?
- **Year Pre-Funding** - *Calculated field, integer.* How many years has the firm existed, pre-funding?
- **Primary Revenue Category** - *Categorical variable, string.* Is the primary expected revenue category of the firm a product, a service or a servitization?
- **Pre-Sales?** - *Binary variable.* Has the firm completed pre-sales exceeding 500 000 NOK?

13.1.2 Financial

- **Total sales** - *Continuous variable, integer.* Total registered sales since the creation of the firm?
- **Total sales 3 years pre-investment** - *Continuous variable, integer*.
- **Total Revenue** - *Continuous variable, integer.* Total registered revenue, including non-sale sources of income.
- **SG&A/Total Costs (TC)** - *Continuous variable, percentage.* Sales, General & Administration (SG&A) over Total Costs (TC) in the year of investment.
- **SG&A/Sales LTM** - *Continuous variable, percentage*. Capped at maximum 2000%
- **SG&A/Revenue LTM** - *Continuous variable, percentage.* Capped at maximum 2000%

13.1.3 Founder Background

- ***Years of founding experience** - *Continuous variable, string*. Accumulated years of founding experience in the founding team. Sub-Sample of firms with founding experience.

13.1.4 Equity Variables

- **Share retained post-investment (pre)** -*Continuous variable, percentage.* Total share of firm retained by founding team post-money at initial investment.
- **Share retained 3 years post-investment (post)** - *Continuous variable, percentage.* Share of firm retained by founding team 3 years post initial investment.
- **Value of founder equity, pre (VFE)** - *Calculated variable, integer.* Value of the founding team's equity post-money at initial investment (equal to Valuation * Founder Share post-money).
- **VFE (post)** - *Calculated variable, integer*. Value of the founding team's equity 3 years post initial investment (equal to Valuation * Founder Share post-money).
- **Founding team gain/loss** - *Calculated variable, integer*. Total founding team gain/loss of equity during the observed three-year period.
- **Percent change in value of founder equity** - *Calculated variable, percentage*. During the observed three-year period.
- **Annual gain/loss, founding team** - *Calculated variable, integer.* Annual gain/loss of founder equity value.
- **Annual percent gain/loss, founding team** - *Calculated variable, percentage*. Annual gain/loss of founder equity value, percent.

13.1.5 Diversification Variables

- **Both sex founders?** - *Binary variable.*
- ***Both sex management?** - *Binary variable.* At the point of initial investment.
- ***Areas of expertise, founders** -*Continuous variable, integer.* Number of independent educational and/or professional founder backgrounds.
- ***Areas of expertise, management** - *Continuous variable, integer.* Number of independent educations and/or professional management backgrounds.
- ***Advisory Board (AB)?** - *Binary variable.* Presence of AB at the point of initial investment.
- ***Number of different areas of expertise, AB** - *Continuous variable, integer.* A sub-sample of firms with AB.
- **Customers Last Twelve Months (LTM)?** - *Calculated variable, binary*. Did the firm have paying customers LTM, which generated total sales LTM exceeding 1MNOK?
- ***Number of customers >10% of total sales LTM** - *Continuous variable, integer.*

13.1.6 Employment Variables

- ***Number of wage-compensated part-time employees** - *Continuous variable, integer.* Minimum of 50% of full employment required. Employees are defined as holding less than 5% or 2MNOK worth of equity in the firm.

13.1.7 Governance Variables

- ***Replaced CEO at point of investment?** - *Binary variable.* Started process to replace CEO at point of initial investment, concluded within 1 year.
- ***Hired BA Graduate at point of investment?** - *Binary variable.*
- **Number of Investor-Appointed Board Members (IABM)** - *Continuous variable, integer.* At point of initial investment.
- ***Hired other support/administrative functions at point of investment?** - *Binary variable.*

13.1.8 Advisory Variables +

- ***Employed external legal advice pre-investment?** - *Binary variable.*
- ***Relevant Patent protection pre-investment?** - *Binary variable.* This is a sub-sample of firms who self-report a high degree of technological innovation.
- ***Employed external financial advice pre-investment?** - *Binary variable.*

13.2 Descriptive Statistics

Below are tables of descriptive statistics for the defined variables not included in the main empirical analysis. Certain variables are simple derivatives of other variables (e.g. Annual Gain/Loss Founding team is the total gain/loss divided by three), these are not included in the statistics. Statistics are not included for variables conditional on other variables (e.g. years of founding experience conditional on founding experience).

Database Variables (Binary)								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	%	*n*	*%*	*n*	*%*	*n*	*%*	*n*
n	100%	30	53%	16	47%	14	6%	2
Function								
Operational	47%	14	57%	8	43%	6	14%	2
Support	53%	16	50%	8	50%	8	0%	0
Primary Rev. Category								
Product	77%	23	52%	12	48%	11	4%	1
Service	13%	4	50%	2	50%	2	0%	0
Servitization	10%	3	67%	2	33%	1	33%	1
Pre-Sales?	43%	13	62%	8	38%	5	23%	3
Both Sex Founders?	13%	4	50%	2	*50%*	*2*	0%	0
Both Sex Management?	27%	8	75%	6	*25%*	*2*	50%	4
Advisory Board?	23%	7	57%	4	*43%*	*3*	14%	1
Customers LTM?	50%	15	60%	9	*40%*	6	20%	3
Replaced CEO?	*33%*	*10*	*50%*	*5*	*50%*	*5*	*0%*	*0*
Hired BA Graduate?	*13%*	*4*	*50%*	*2*	*50%*	*2*	*0%*	*0*
Hired Other Support/Adm.?	*30%*	*9*	*56%*	*5*	*44%*	*4*	*11%*	*1*
Hired Legal Services?	*93%*	*28*	*54%*	*15*	*46%*	*13*	*7%*	*2*
Hired Financial Services?	*17%*	*5*	*40%*	*2*	*60%*	*3*	*-20%*	*-1*

Table 5: *Descriptive statistics of the binary variables gathered but not included in the empirical analysis. Statistically significant differences between the ESS and non-ESS distributions are marked with star(s).* $* = p - value < 0.1$, $** = p - value < 0.05$, $*** = p - value < 0.01$. See *Chapter 11.1* for variable definitions.

Database Variables (Continuous)								
n=30	*All*		*ESS*		*non-ESS*		*Difference*	
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	*Mean*	*Median*
Years Pre-Funding	4.7	*3.5*	5.1	*3.0*	*4.4*	*4.0*	0.7	- 1.0
Total Sales (To Closest 100')	10,400,000	2,400,000	14,600,000	1,750,000	5,500,000	2,700,000	9,100,000	- 950,000
Total Sales 3y Pre (To Closest 100')	7,600,000	2,050,000	10,900,000	1,700,000	3,800,000	2,250,000	7,100,000	- 550,000
Total Revenue (To Closest 100')	11,500,000	3,450,000	16,000,000	4,150,000	6,300,000	3,450,000	9,700,000	700,000
SG&A/TC	85.0%	89.6%	86.5%	90.5%	83.2%	89.1%	3.3%	1.5%
SG&A/Sales LTM	646.2%	222.5%	883.0%	225.1%	291.0%	208.5%	592.0%	16.6%
SG&A/Revenue LTM	314.6%	174.9%	370.9%	143.4%	244.1%	206.6%	126.8%	-63.2%
Share Retained (Pre)	29.0%	27.0%	33.3%	29.5%	24.1%	22.5%	9.2%	7.0%
Share Retained (Post)	13.2%	6.1%	18.5%	11.5%	7.1%	2.5%	(*) 11.4%	9.0%
VFE (Pre) (To Closest 100')	6,000,000	2,907,500	8,600,000	4,716,000	3,100,000	2,787,500	(*) 5,500,000	1,928,500
VFE (Post) (To Closest 100')	8,700,000	2,538,950	15,000,000	5,075,000	1,400,000	328,500	(**) 13,600,000	4,746,500
Founder Gain/Loss (To Closest 100')	2,700,000	- 73,250	6,500,000	237,500	- 1,600,000	- 475,000	8,100,000	712,500
% Δ VFE	47.2%	-4.3%	120.8%	45.8%	-36.9%	-23.6%	(***) 157.7%	1
Number of AoEs - Founder(s)	1.6	1.0	2.0	2.0	1.1	1.0	(**) 0.9	1
Number of AoEs - Management	1.9	2.0	2.3	2.0	1.4	1.0	(**) 0.9	1.0
# Cust >10% LTM Sales	1.0	1.0	1.1	1.0	0.8	-	0.3	100.0%
Employees 50%+	5.9	4.0	9.1	5.0	2.2	2.0	(**) 6.9	3.0
# IABM	1.6	1.5	1.4	1.0	1.8	2.0	- 0.4	- 1.0

Table 6: *Descriptive statistics of the continuous variables gathered but not included in the empirical analysis. Statistically significant differences between the ESS and non-ESS distributions are marked with star(s).* $* = p - value < 0.1$, $** = p - value < 0.05$, $*** = p - value < 0.01$. See *Chapter 11.1* for variable definitions.