Norwegian School of Economics

Bergen, Fall, 2020

Measuring the Effect of Business Incubation in Oslo

An empirical study on performance, survival, and access to public subsidies

Arve F. Eide & Thomas Jelsa

Supervisor: Nataliya Gerasimova

Master Thesis in Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This paper examines the effect of participation in a business incubation program in Oslo, Norway, measured in terms of economic performance, survival rates, and access to public subsidies.

Our research involves incubated companies entering an incubator from 2011 and 2016, matching with comparable companies with similar characteristics. Furthermore, we use the data available for these companies in the period 2011-2018 to analyse the effect incubators have on the incubated companies.

To measure the effect of incubator participation, we construct a representative control group by using coarsened exact matching combined with nearest Mahalanobis distance. We then use difference-in-differences estimation (DiD) to estimate the effect of the incubator program on the incubated companies.

We find that, in terms of performance, the only positive significant effect of incubator participation is on the number of employees. We find no significant effects on value creation or sales revenues. However, we also find some evidence of negative effects on operating profits for the incubated companies. Further, we find no evidence that the group of incubated companies experience higher survival rates or better access to public subsidies, compared to the group of control companies.

Preface

This thesis is written as a part of our MSc in Economics and Business Administration at the Norwegian School of Economics (NHH).

Our interest in business incubators comes from Arve's participation in an incubated startup in Bergen, combined with Thomas' interests in innovation and entrepreneurship. This sparked our interest in performing quantitative analysis to measure if business incubators have a positive effect on the companies they assist.

We would like to thank our supervisor, Nataliya Gerasimova, for the support and feedback along the way. We would also like to thank Sigmund Tveit in NSD, Kellis Akselsen in SNF, Beate Rotefoss in Siva, Mathilde Tuv Kverneland in Arkwright X, and the Simula Garage for providing us with data and information for our thesis. Lastly, we would like to thank Diana Medel Carrero for providing insight into the application process to Innovation Norway, and Arvid Raknerud in Statistics Norway for providing us with helpful input regarding the direction of methodology for this thesis.

Bergen, December 20, 2020

Arve F. Eide

Thomas Jelsa

Contents

Table of Contents

1. Introduction

Entrepreneurship is considered an important source of economic growth by economic policymakers (Wennekers and Thurik, 1999). In 2015, the Norwegian Ministry of Trade, Industries and Fisheries released a new plan to stimulate entrepreneurship. The report highlights the role of new and innovative companies in enhancing economic growth and changing current industries (Ministry of Trade, Industry and Fisheries, 2015). At the same time, the report argues that some challenges are obstructing the successful growth of these companies. Examples of such challenges are lack of capital and unfavourable tax laws. Similarly, Grimsby, Grünfeld, and Jakobsen (2009) describe small and midsize businesses (SMBs) as highly important growth engines in the Norwegian economy. They identify SMBs as both an important segment in creating new jobs and as the most important segment for innovation and transforming industries in the Norwegian economy. However, they find Norwegian tax laws to be unfavourable for small companies.

In an article published in *Dagens Næringsliv,* Erik Hagen — managing partner at Viking Venture — evaluates the decreasing amount of venture capital invested in Norway as damaging for Norwegian startups (Tobiassen, 2015). In the same article, Fredrik Syversen — director of industry development at IKT-Norge — claims that investors in Norway are moving towards startups in the growth stage, making capital less accessible for early-stage startups. Meanwhile, the CEO of Argentum, Joachim Høegh-Krohn (2017), claims that access to capital is not the main problem for Norwegian startups. He argues that low returns on investments in early-stage startups are the reason for low venture investments in Norway, and points to public subsidies and better tax incentives as possible solutions to increase the number of successful startups in Norway (Høegh-Krohn, 2017).

Lack of capital and unfavourable tax laws can result in lower entrepreneurial activity (Keuschnigg and Nielsen, 2003). Indeed, Statistics Norway (2020a) report that only 28.4% of all companies established in 2013 were still operating in 2020. Similarly, Holst (2019) describes how 2018 featured the highest bankruptcy numbers since 1993, representing an increase of 2.8% from 2017. The 2020 Covid-19 pandemic is also predicted to vastly increase Norwegian bankruptcy rates, despite public crisis subsidies having resulted in low bankruptcy rates during the first eight months following the pandemic outbreak (Nilssen, 2020).

Under some circumstances, market mechanisms when left alone fail to achieve the best outcome for the economy (Buigues and Sekkat, 2011). To correct these market failures, public subsidies could be a useful countermeasure. Business incubators are another possible correction for such market failures (Hackett and Dilts, 2004).

The purpose of our thesis is to investigate if business incubators have a role in solving these market inefficiencies by enhancing company performance, survival rates, and access to public subsidies. To measure the different impacts of business incubators on these three fields, we construct representative samples of non-incubated companies with similar company characteristics and run regressions on the differences between the groups.

Our findings suggest that participation in a business incubator program in Oslo has few significant effects on performance. The only positive effect we find is on the number of employees, suggesting that incubator participation results in 1.137 additional employees. However, our findings even suggest that incubator participation results in a 263 360 NOK decrease in operating profits. In terms of survival rate and access to public subsidies, we find no significant effects of incubator participation.

This thesis is organised into seven sections. Section 2 comprises a literature review and the most important definitions and concepts used in subsequent sections. Section 3 outlines our three hypotheses and the research design of the study. Section 4 describes the methodology used in both the matching process and the regressions, before section 5 explains data collection and processing. Section 6 presents the results of our analyses on performance, survival rates, and access to public subsidies. Finally, section 7 offers the concluding remarks of our thesis.

2. Background

We begin this section by reviewing some of the most important literature on business incubator performance. Further, we will provide some background on business incubators and the reason for their existence, before providing a short overview on the business incubation scenes in Norway and Oslo. Lastly, we will offer a description of the most important public funding schemes for startups in Norway.

2.1 Litterature review

Many papers research the success and growth of business incubators. However, there are relatively few papers focusing solely on the economic performance of those companies attending the incubators. Some of the papers include economic performance as a measured performance indicator, while others do not analyse it at all.

The most relevant paper on this topic is a report written by Statistics Norway on behalf of the Ministry of Trade, Industry and Fisheries. The report presents the results of Siva's (selskapet for industrivekst) business gardens and incubation program (Fjærli, Iancu and Raknerud, 2018). Fjærli et al. (2018) use a sample of nearly 3 800 companies that have attended Siva's incubation program or one of their business gardens. To assess the performance of the incubator companies, they choose the following metrics: sales revenues, number of employees, value creation, labour productivity, and return on total capital (ROTC). For each metric the authors consider average additional growth, from entering the incubation program until 3 years later, compared to a control group of companies observed during the same period. In addition, they report more long-term effects (3−5 years after entering the program). To construct a representative control group, matching is used. Their matching procedure is a combination of i) exact matching and ii) propensity score matching, based on the company's total assets in the year of treatment. Exact matching requires the companies to be in the same 2-code industry, age group, and region (Fjærli et al., 2018). Propensity score matching is based on selecting companies for the control group that are most likely to participate in the treated group, based on observable characteristics (Fjærli, et al., 2018). The authors' matching procedure involves the loss of around 1 500 companies from nearly 3 800.

The findings indicate that participation in Siva's programs is connected with significant additional growth in all effect indicators, except for ROTC (Fjærli et al., 2018). However, in the longer time period (3-6 years) few significant additional effects are observed from participating in the program.

While the report by Fjærli et al. (2018) provides a detailed analysis of the financial performance of incubated companies vs. non-incubated companies, some factors reduce the accuracy of their analysis. Their combination of business incubators and business gardens complicates the drawing of conclusions on just one of the two, as the additional growth of the companies could be due to only one. In addition, there is a possibility that some of the companies in their control group have attended other incubators that were not included in Siva's programs. Elsewhere, Colombo and Delmastro (2002) reach a similar conclusion about incubator effectiveness in Italy, as they find that the average general growth rate is 55% for incubated companies and 30% for non-incubated companies.

However, Lukeš, Longo, and Zouhar (2018) produced results that contradict those of Fjærli et al. (2018) and Colombo and Delmastro (2002). In their research on incubated companies in Italy, Lukeš et al. (2018) find that incubation tenancy has a significant negative effect on the sales revenues of innovative startups. Their research also analyses the effect of incubation tenancy on number of employees, where they observe no significant effect. Overall, their research does not find any evidence that justifies public spending on business incubators in the short run. However, their research is only focused on sales revenues and number of employees.

Ferguson and Olofsson (2004) conducted similar research on science parks in Sweden. They included the on-park survival rate as a measure of success, and found that of the 30 companies located in the science parks in 1995, 93.3% were still operating in 2002 — 7 years later. In comparison, in the off-park sample the 7-year survival rate was 66.7%. They also found that of the on-park companies that did not survive, half of them were the result of mergers and acquisitions (M&As). In the off-park sample, on the other hand, only a third of the non-surviving companies were the result of M&As. Thus, a larger share of the non-surviving companies in the on-park sample can be considered successful. Other studies show that being located in an incubator does not necessarily increase the survival rates of the incubated companies. In their systematic review of business incubation research, Hackett and Dilts (2004) found that the level of incubator development and the number of incubated companies are positively related to incubated companies' survival.

Colombo and Delmastro (2002) also analyse startups' access to public subsidies. Their findings suggest that companies located in business incubators or science parks had easier access to public financial funds, with 51% of the on-incubator compnaies receiving public subsidies compared to 33% of the off-incubator sample. This is also mentioned in the report of Fjærli et al. (2018), who found that around 25% of the companies participating in Siva's incubation program receive funding or support from at least one other public scheme. However, Fjærli et al. (2018) did not include the percentage of non-incubated companies receiving grants in their research.

Our contribution to the existing literature on business incubators is to perform a pure economic analysis of incubator effectiveness by examining the incubators in a single city. To the best of our knowledge, no existing research has performed such an analysis, the closest being Fjærli et al. (2018). By analysing the performance of the majority of incubators in a single city, our research will reduce the probability of the companies in the control group participating in another Oslo-based incubator.

2.2 Business incubators

The concept of business incubators can be traced back to Batavia, New York in 1959. The number of incubators grew slowly, and by 1980 there were still only 12 incubators in the United States (Stubberud, 2016). However, from 1980 to 2000 the number of incubators in the United States grew to over 1 000, representing the largest incubator industry in the world (European Commission, 2002, p. 10).

There are many different definitions of a business incubator. We choose to use the definition of Hackett and Dilts (2004, p. 57), presented in their systematic review of business incubation research:

"A business incubator is a shared office-space facility that seeks to provide its incubatees (i.e. "portfolio-" or "client-" or "tenant-companies") with a strategic, value-adding intervention system (i.e. business incubation) of monitoring and business assistance."

2.2.1 Market failure theory

A common theoretical foundation in the incubator literature is market failure theory. Market failure occurs when the competitive transactive space for the production and sale of goods and ideas fails to produce a desired outcome (Hackett and Dilts, 2004). Sources of market failure include externalities, imperfect information, monopoly power, and public goods. Researchers who subscribe to market failure theory believe that structures within the market hinder the successful development of entrepreneurial new ventures, and that incubators are a tool for resolving these market failures (Hackett and Dilts, 2004).

2.2.2 Business incubators in Norway

The number of business incubators in Norway has been growing rapidly during the past few years (Tandsæther-Andersen, 2017). However, there is a lack of registers that list all existing incubators. As a result, we do not know the exact number of active business incubators in Norway. Based on discussions with people from the startup scene in Norway and some partial lists of Norwegian incubators, we estimate that there are around 50 active business incubators.

Selskapet for industrivekst (Siva) is an important player in the Norwegian business incubator infrastructure. Established in 1968, Siva is the Norwegian government's instrument for facilitating ownership, developing companies, and growing industry and knowledge clusters in Norway, with a special focus on facilitating growth in rural areas (SIVA, n.d.a). In 2018, Siva partially owned and supported 34 incubators in Norway, representing a total of 2 081 incubated companies (Siva, n.d.b). Of these 34 incubators, three are located in Oslo (Siva, n.d.c).

In addition to the incubators supported by Siva, many different ownership structures can be found among the Norwegian incubators. Some are fully owned by the Norwegian state or municipalities, while others are non-profit, privately owned, or a combination of the different ownership structures.

2.2.3 Business incubators in Oslo

As in the rest of Norway, there are no official lists or registers of the incubators established in Oslo. However, an article published in the Norwegian startup newspaper *Shifter* lists all of the incubators present in Oslo at the end of 2017. In this article, Tandsæther-Andersen (2017) presents a list of 16 Oslo-based incubators. Those included in the list contain several different focus fields, ranging from technology incubators to incubators focused on immigrants with innovative ideas.

However, through further research and conversations with the various incubators, we found that only 10 of these incubators defined themselves as incubators or have been active since 2010. In addition, through internet searches and conversations with people in the incubator ecosystem in Oslo, we found one additional incubator to add to the list. In total, we found 11 active incubators in Oslo at the time of our research.

Since the establishment of incubators in Oslo, many successful companies have been through their programs. Remarkable AS, Kahoot! AS, and Zwipe AS are some examples of companies considered as successful ventures.

According to the webpages of different Oslo-based business incubators, *innovation, founders, market potential,* and *ambition* appear to be important characteristics among incubated companies. For instance, the Oslo-based business incubator Arkwright X (n.d.) writes the following:

> *"We are always looking for super teams with innovative ideas for how to disrupt the status quo…. More specifically, you need: A unique idea with a credible commercialization potential. You need to have a clear and strong value proposition for your product/solution."*

Although the requirements of approval vary between incubators, it is our impression that most business incubators in Oslo focus on supporting innovative startups with high growth potential.

2.3 Public support schemes

In addition to providing indirect support to startups through Siva, the Norwegian government also supports startups directly, mainly through Innovation Norway and the Research Council of Norway. These two companies are responsible for awarding grants and loans to companies with innovative ideas (Innovation Norway, 2020c; The Research Council of Norway, 2019a).

2.3.1 Innovation Norway

Innovation Norway is the Norwegian Government's most important instrument for innovation and the development of Norwegian enterprises and industry (Innovation Norway, 2020a). In addition to financial services like grants and loans, Innovation Norway provides competence, advisory services, promotional services, and network services.

Two of Innovation Norway's most important funding schemes for startups are grants for market clarification and grants for commercialisation (Innovation Norway, 2020b). These grants are given to entrepreneurs who want to try out and scale innovative ideas. Since 2010, Innovation Norway has awarded 32 077 MNOK in grants and 40 175 MNOK in loans (Innovation Norway, 2020d).

2.3.2 The Research Council of Norway

The Research Council of Norway (RCN) works to promote research and innovation of high quality and relevance. It also aims to generate knowledge in priority areas, to enable Norway to deal with key challenges to society and the business sector (the Research Council of Norway, n.d.). Some of the most relevant funding schemes offered by RCN to startups are the *SkatteFUNN Tax Incentive Scheme* and *Innovation Projects for the Industrial Sector*.

SkatteFUNN is a rights-based tax deduction scheme. All Norwegian companies working with R&D can apply for approval, thus obtaining the right to tax deductions (the Research Council of Norway, 2019b). To qualify for SkatteFUNN, a company must work on improving an existing product or service and dedicate resources towards this goal.

An Innovation project for the industrial sector is defined as a company-driven project with extensive R&D activities (the Research Council of Norway, 2019c). An Innovation project should make a significant contribution to innovation and offer increased value creation for the companies participating in the project, and for the general public, by making new solutions available. In 2020 the funding scale ranges between 2 MNOK and 16 MNOK

3. Hypotheses and research design

This section describes our three hypotheses and the research design used to test them.

3.1 Hypotheses

Based on the existing literature and our own experiences with business incubators, we have developed the following three hypotheses:

3.1.1 Hypothesis 1

The findings of Fjærli et al. (2018) indicate that incubated companies achieve higher growth rates on sales, number of employees, value creation, and labour productivity compared to the control group. Similarly, Colombo and Delmastro (2002) find that the average general growth rate for the incubated companies considered in their study was 55%, compared to 30% for the non-incubated companies. These findings led us to formulate our first hypothesis:

> *Companies attending a business incubation program in Oslo outperform non-incubated companies in terms of growth in sales revenues, value creation, operating profit, and number of employees.*

Based on the existing literature, our hypotheses are likely to be correct at some levels. With this first hypothesis, we expect to find significant additional growth on at least some of the assessment metrics, but not necessarily all of them.

3.1.2 Hypothesis 2

Ferguson and Olofsson (2004) find that companies located in science parks in Sweden have a higher survival rate than off-park companies, as they achieve a 7-year survival rate of 93.3% compared to 66.7% in off-park companies. They also find that of the non-surviving companies, a larger portion of the on-park companies are the result of M&As. This led to our second hypothesis:

> *Companies attending a business incubation program in Oslo survive longer than companies with similar characteristics that have not attended an incubation program in Oslo.*

The second hypothesis is based on findings of Ferguson and Olofsson (2004) and Hackett and Dilts (2004). However, Hackett and Dilts (2004) suggest that the age and size of the business incubator are positively related to survival rate. Thus, since all of the business incubators in our sample started after 2011, we expect to find modest differences between the survival rates of incubated companies and control companies.

3.1.3 Hypothesis 3

Our third hypothesis is based on findings from the literature review and from our own experiences. Colombo and Delmastro (2002) find that 51% of the studied on-incubator companies received public subsidies compared to 33% among the off-incubator sample. Similarly, Fjærli et al. (2018) find that around 25% of incubated companies received public funds. This also correlates with our experience of participating in a business incubator. These factors led to our third hypothesis:

> *Companies attending a business incubator in Oslo have a better chance of receiving public subsidies, compared to non-incubated companies.*

3.2 Research design

Our thesis follows a quantitative approach.

The background, literature review, and the growing startup environment in the Oslo region led to the research design of this master thesis. First, we wish to evaluate and review the performance of incubated companies in comparison to a control group with similar characteristics, at the time when a treated company, i.e. an incubated company, enters an incubator. Here we choose to examine only the Oslo region, to enable greater depth of analysis and to examine the effects within a sub-ecosystem. We will also limit the analysis to incubated companies that entered an incubator during the time period 2011–2016, as we only have accounting data available up to 2018.

The choice of variables used to assess the companies is based on what we believe to be the most important metrics. In turn, this is based on a combination of existing literature and conversations with players in the startup ecosystem in Oslo. The variables utilised to assess the growth of the companies are as follows:

1. *Sales revenues*

   The reported revenue generated through the sales of goods and services.

2. *Value creation*

   Defined as the sum of reported operating profits and salary costs.

3. *Operating profit*

   The reported operating profit.

4. *Number of employees*

   The reported number of employees.

We will also assess the bankruptcy rates and the percentage of companies receiving government funding through Innovation Norway and/or the RCN.

As Fjærli et al. (2018) describe, the challenge in measuring non-experimental situations – like business incubators – lies in predicting what would have been the outcome for the incubated companies *without* participating in an incubation program, based on historical data. To compensate for not knowing the counterfactual outcomes, research on business incubators has often constructed a control group from similar companies that have not participated in the incubator (e.g. Fjærli et al., 2018, Colombo and Delmastro, 2002).

To measure the incubation performance, we match a representative control group of comparable companies that have not participated in an incubation program to our sample of incubated companies. We gained partial access to all of the 11 active business incubators in Oslo. However, some of the incubators have poor reporting routines; for instance, they lack data on the start period or length of stay for the incubated companies. Hence, we have excluded companies without a given incubation start date or period in our samples. These were also excluded from the control group sample, to avoid including incubated companies in our control sample.

We will perform a quantitative analysis by comparing the performance of the incubated companies with that of non-incubated companies having similar company characteristics. The control sample will be found by using coarsened exact matching on chosen coarsened covariates bins, with 1:1 matching on nearest Mahalanobis distance for exact values on the same covariates. Our approach will decrease the risk of including companies that have been incubated at other Oslo incubators in our control sample.

As described by Coleman (2018), Norway's startup ecosystem is launching numerous incubators and accelerators that are founded and led by passionate entrepreneurs. This developing ecosystem is on the path to rapidly grow a sense of community and cohesion. We are also interested in this topic because the writers of this thesis have an ongoing venture in an incubator. The reason for choosing the Oslo region is because, to our best knowledge, no reports or papers have evaluated the economic performance of incubators specifically in this region. This unique environment and fast-growing ecosystem are therefore of great interest to us personally.

4. Methodology

In this section we explain the methodology used to analyse our three hypotheses. We begin by explaining the matching technique used to match the incubated companies with similar non-incubated companies. We then describe the methodology used to assess the performance, survival, and access to public subsidies of the two groups.

4.1 Choice of variables for the preprocessing technique

With a total of 214 different variables to control for in SNF — Centre for Applied Research at NHH — database when selecting a control group having similar company characteristics, we identify definite variables that are being influenced for companies that are incubated and those that are not. Since the effect of the treatment is interpreted as the difference between these groups, ceteris paribus, i.e. all else equal, the selected variables are of great importance. We choose to use a similar methodology to Fjærli et al. (2018) for identifying a representative control group, with some modifications. During the matching procedure we choose variables that should be close or equal for the incubated companies and control companies. As described later in section 5, we only include companies in the Oslo region when matching the control companies. We match the treated and control companies on the following variables:

Table 1: Matching variables

Variable	**Description**
Year of Incorporation	The year in which the company was started
Matching Year	The year in which the incubated company entered the incubator
2-digit NACE Code	Industry code (e.g. J-61 is telecommunication)
Rating Code	Risk rating performed by *Dun and Bradstreet (D&B)*
Total Income	The reported total income of the company in the year of matching
Number of Employees	The reported number of employees in the year of matching
Total Debt	The reported total debt of the company in the year of matching
Total Capital	The reported total capital of the company in the year of matching

The matching year must be equal for both groups, as we are comparing the two groups over a period of time. Similarly, the year of incorporation should be equal in both groups, as the experience and market conditions might correlate with the age of the companies.

Market conditions differ across industries, so it is necessary to exactly match the 2-digit NACE code when finding similar market characteristics between the two companies in the same business area. Using only sector information or a 1-digit NACE-code, e.g. technology or transport, would result in an overly vague match between two companies in the same sector.

The rating code refers to an external rating of the company performed by the independent company Dun and Bradstreet. It thus represents an unbiased evaluation of the market conditions for a company and its degree of liquidity from an objective perspective. The rating code is a number between 0–9, where a lower number represents a higher risk and a higher number represents a lower risk. However, the number 9 represents a bankrupt or closed company (Berner, Mjøs and Olving, 2016).

The total reported income and total employees are used to ensure matching companies at similar growth stage and with a similar organisation size. Similarly, the total debt and total capital are used to match companies having similar capital structures, and hence similar risks and incentives.

After the exclusion process, which is described later in section 5, the data on the incubators is merged with the company level, yearly accounting data provided by SNF. This produces panel data. Panel data, or longitudinal data, refers to cross-sections of information about unique companies across time-series. In balanced panel data, the number of time periods, T, is the same for all individuals, c. Observations in panel data involve at least two dimensions; a cross-sectional dimension, i, and a time series dimension, t. It can also include more complicated clustering (Antweiler, 2001, Davis, 2002). Otherwise, it is unbalanced. Since some of the incubated companies do not operate in the same periods, available data is often referred to as unbalanced panel data (Baltagi, 2005, p.165).

4.2 Matching procedure

Iacus, King, and Porro (n.d.) describes matching in the following way:

> "Matching is a nonparametric method of preprocessing data to control from some or all of the potentially confounding influence of pretreatment control variables by reducing imbalance between the treatment and control groups".

The dataset retrieved from SNF includes a total of 47 829 unique companies from over an 8-year time frame that the incubated companies can match with. These 47 829 companies exclude all incubated companies, including the incubated companies with missing data in our initial incubator sample of 630 companies.

4.2.1 Transforming the data

To our knowledge, limited research has performed regression analyses on unbalanced panel data[1]. Dettmann, Giebler, and Weyh (2019) argue that in the case of unbalanced panel data, a flexible difference-in-differences (DiD) approach in terms of time after treatment, instead of specific accounting year analysis, reduces the time bias and matches potential partners for every treated unit to those observed at the individual matching year, for example, the treatment start year.

When preparing our dataset, we encountered an important decision regarding the start of an incubation process. To measure the effects of the incubation program, using available data on pre-, under-, and post-treatment years would be ideal. However, many of the companies only have accounting data from the first year that they appeared in an incubator. Therefore, we define the year of entry into an incubator as year zero.

Because we are analysing different hypotheses, we opt to produce two different datasets; a *growth sample*, to examine growth rates in a continuous three-year period, and a *duration sample,* to compare the survival rates and access to government grants between the treated and control group. Since one of the purposes of this study is to consider post-treatment

[1] In 2019 (Dettmann, Giebler and Weyh) a new revised method to deal with unbalanced panel data, which included a package for Stata, flexpaneldid, presented a modification on the matching and difference-in-differences approaches similar to that of Heckman, Ichimura, Smith, and Todd (1998). The available data are not sufficient to use these stata commands, as it does not possess any of the required characteristics, namely pre-treatment, treatment, post-treatment available. We acknowledge that the research exists, but the method cannot be applied to our thesis.

performance and growth, we only investigate companies having accounting data from SNF for a continuous period of three years in our growth sample. Based on research by Statistics Norway (2020a), only around ~35% of startup companies survive after three years, which could drastically reduce our final samples.

Accordingly, we choose to transform the growth sample in order to have correct and balanced data when comparing the descriptive statistics; and to use the duration sample for comparing survival rates and government grants between the treated and control companies spanning different time periods.

For transforming the unbalanced panel data in the growth sample, which is sectioned into different time periods based on continuous years, we decide to transform it into years after treatment. Year zero is therefore the start of treatment, if the company is incubated. We thus transform the unbalanced dataset into balanced panel data. With such data, which enables examination of the treatment effect after year zero, a dataset results with which we can monitor the combined outcome variables over different time periods. Comparing the treated company with a control company that does not participate in an incubation program, but that has the exact same age, accounting year, and 2-digit NACE code, thus provides a representative group of outcomes for those not being incubated, over multiple time periods.

The companies in the sample may be either incubated or non-incubated, and we do not observe counterfactual outcomes for any of the companies. Economic performance is reliant on the economic environment and hence the place and time of observation (Heckman, LaLonde and Smith, 1999). Ignoring this fact would result in comparing a treated company with a control company from a different year having a different length of experience (i.e. age), resulting in a time bias comparison.

4.2.2 Coarseened Exact Matching

Coarsened exact matching (CEM) is a matching method whereby one segments some covariates of the population and finds matches on intervals instead of on one exact number. According to Iacus et al. (n.d.), the CEM method meets the congruence principle and is robust to measurement errors. Therefore, we use the variables chosen as described in section 4.1. We use CEM combined with the nearest Mahalanobis distance, for both the growth and duration samples of the analysis, with different specifications of the coarsened covariates.

The growth rate sample is matched with exact matching on all coarsened covariates interval bins, only matching with companies having available continuous accounting data over a period of three years. By using the CEM procedure we obtain stratums for each treated company. Stratums are sections of companies having similar baseline characteristics, i.e. in the same exact matched coarsened bins on all of the chosen covariates. Thus, each treated company has no, one, or multiple matched companies in each stratum. After an exact matching is performed on the coarsened covariate bins, the stratum is used to determine the nearest Mahalanobis distance in terms of real values, i.e. not the interval bins, to match the treated company with its closest control company 1:1.

The duration rate sample is also matched with exact matching on some of the coarsened bins, that being start year, accounting year, and 2-digit NACE-code, but without the limitation of having a continuous period of three years. The other covariates are matched, within each stratum, to the closest similar control company having the nearest Mahalanobis distance. This is done to obtain more matches and to avoid bias in terms of only choosing *survival companies*, i.e. companies that survive for more than one year. Therefore, this sample will not be used to measure economic performance, as the main focus of the duration sample is to look only at the survival of companies and government grants with similar characteristics, accompanying loosened constraints on the matching criteria.

According to Rippollone, Huybrechts, Rothman, Ferguson, and Franklin (2019), four steps are necessary when implementing CEM. Letting X be the vector of observed covariates, we:

Step 1: Temporarily coarsen the covariates in X

Step 2: Implement exact matching with the coarsened data

Step 3: Eliminate unmatched units, and pass on the original (coarsened) values

Step 4: Estimate the ATT in the matched data set.

These four steps will be explained in detail in the four next sections.

Step 1

When coarsening the covariates in x, we ensure that the units having the same value for the coarsened covariate bins are substantially indistinguishable. Subcategorization of the covariates chosen to be able to define two similar companies is done by dividing the values of the covariates at different intervals. Furthermore, we examine the histograms of these covariate values for the treated companies to find distinguishable and adequate intervals. This ensures the removal of observable differences between the treated and control groups. In our analysis, the treated group has a significantly smaller population than the control group, which contains all companies in the Oslo region between 2011–2016. If no control companies exactly match the coarsened group covariates of a treated company, depending on the chosen covariates, this means that the treated companies in question are too unique to be compared. That is, in terms of the chosen variables, matching with a less similar company increases variance and standard deviation, resulting in interference within the results.

As an example, the e-ink tablet company Remarkable AS was incubated at Startuplab in 2016 and passed 1 MNOK in pre-sales during their first opening day in November 2016 (Øyvann, 2017). From SNF's accounting numbers we observe that Remarkable AS had in their first accounting year: a debt of 14 MNOK, capital of -2 MNOK, total income of 250 000 NOK, and age zero, with no employees. In our control sample, no other control company seems to exist in the Oslo region that matches this in 2016. It is reasonable to suggest that having a debt to customers to deliver their products before they were ready to ship, combined with a reported low total income, increases the rationale of why they are not matched with any company from the control group when they have such a high debt value in their first year.

Step 2

Using the covariates of X, the growth and duration sample datasets finds matches for the treated and control companies using the R package "Matchit". This package includes different matching models to match two entities originating from two groups. Using exact matching on the coarsened interval bins, and distance = Mahalanobis, replace = false, and ratio = 1, with the non-coarsened values as covariates, we find the closest match within each stratum for each treated company, for both the growth and duration samples, by utilising the Mahalanobis distance. For both samples we match exactly using accounting year, 2-digit NACE-code, and age. For the growth sample we extend the exact match to be performed on employee-, total income-, capital-, rating code-, and debt-coarsened interval groups.

The Mahalanobis distance is measured as:

$$D^2 = (x-m)^T C^{-1}(x-m), \tag{1}$$

Where D^2 is the Mahalanobis distance, x equals the vector of data, m is the vector of mean values of independent variables, C^{-1} equals the covariance matrix of independent variables, and T indicates that the vector should be transposed (McLachlan, 1999).

The Mahalanobis distance solves the multidimensional problem, as it measures the distance between points based on being closest to each other in distance. In comparison, the use of a propensity score, which is a popular matching procedure to preprocess data for causal inference, takes a multidimensional dataset and creates a one-dimensional score (0.0−1.0) based on the probability of it being treated. However, the Mahalanobis distance is measured in the actual covariate space. When using more than three covariates, the Mahalanobis matching distances become too complex to represent in a dimensional space (McLachlan, 1999).

Considering Mahalanobis matching by itself, a research paper by Baltar, Sousa, and Wesphal (2014) proved that a hybrid between propensity score matching and Mahalanobis distance finds better matches than through the individual calculations. In our thesis, where we use a combination of CEM with Mahalanobis distance to our data, we provide a sample of the closest match to treated companies with control companies that are similar, as well as in terms of matching them on exact coarsened interval bins.

By matching with exact coarsened bins on the chosen covariates, we are not matching companies without exact matches, i.e. we are pruning any stratum with 0 treated and 0 control units. Furthermore, in the 1:1 matching process used in this thesis, the matching is done without replacement. This means that once a treated company has been matched with a control company, the control company is not returned to the pool of potential matches for treated companies and cannot be selected again. Hence, the same individual control company cannot be selected as a match for multiple treated companies. Using a control multiple times can induce bias that has an unobservable effect on testing the comparison between the treatment and control companies.

Meanwhile, CEM assumes that the matching variables contain all of the confounders, or that matching on the variables that we have will result in matches on the confounders that we do not have. When using CEM, we have to assume that any grouping, i.e. coarsening the covariates into bins, results in errors that are within tolerable limits. Essentially, we are finding the treated companies' doppelgangers in Oslo.

Step 3

After matching the treatment group with the control group, unmatched stratums are eliminated. If the unmatched companies are included in the final analysis, it could bias the exposure effect estimates (Petersen, Porter, Gruber, Wang and van der Laan, 2012).

Step 4

While CEM matching provides the utility to have multiple control company matches, in this thesis only one treated company is matched with one control company. The two matched companies are therefore 1:1 matched, and weighted similarly against each other in further analyses. The matching thus becomes unbiased and is easier to analyse. However, it is important to note that for these results to have a causal meaning, the parallel path assumption must hold. This assumes that companies which have not been incubated would have developed in the same way as the incubated companies, had they been incubated. Correspondingly, it also assumes that those companies who were incubated would have developed in the same way as the non-incubated companies, had they not been incubated. These assumptions are not certain but are necessary for counterfactual analysis, as we are conducting in this thesis. The average treatment effect on the treated companies will be analysed in the results, section 6.

4.3 Difference in differences

The difference-in-differences (DiD) method compares the *changes* in outcomes over time between a population that is treated and a population that is not (Gertler, Martinez, Premand, Rawlings, and Vermeersch, 2011). The DiD estimator estimates the counterfactual outcome by calculating the change in outcomes for the treated group, i.e. the difference in mean after treatment minus the entry period mean of the treated group, and then subtracting this difference in means after treatment minus the entry period mean of the control group. In this way, DiD computes the difference between two differences in two different groups.

The further application of DiD relies on the common trend assumption that the two groups would have common trends if the treated group had not received the treatment. If this assumption holds, the unobservable company characteristics between the two groups will not influence the estimates. Through this assumption, one could say that without the treatment the outcomes would need to increase or decrease at the same rate in the treated group and control group, respectively.

In the CEM with coarsened bins using the nearest Mahalanobis distance, we assume that we do not have any unmeasured confounders or omitted variables in the DiD estimation process. Due to the choice to have a control group of companies, we assume that the treatment assignment is not independent of the potential outcomes. Therefore, the requirements for difference-in-difference may be violated. A natural experiment does not exist that excludes a subclass of companies from the treatment of an incubation stay. We have therefore chosen the next best alternative, which is to use the companies that did not apply to the incubators, or that applied but failed to obtain treatment in one of the incubators in Oslo as the control group. The matched control group must therefore have similar statistical characteristics to our chosen covariates, to be deemed as representative control individuals for the treated companies.

When assessing the balance between the covariates in the treated and control groups, we will – in addition to the means of the groups – also view a comparison between the histograms in the matched year. We want to examine if the covariates are similar to each other in the total matched sample. In our statistical matching analysis using DiD estimation, we will thus be able to control for **observable** covariates that influence the selection of an incubated company, but not for the unobservable conditions when selecting these companies.

4.3.1 Group time Average Treatment effect on the Treated

Because we have three periods in our growth sample, the Average Treatment effect on the Treated (ATT) needs to be modified to fit our research design, from the normal two-periods - two-group Average Treatment effect on the Treated. We will therefore focus on the average treatment effect of the companies which are members of the group g in a time period t. (Callaway and Sant'Anna, 2019) The group time Average Treatment effect on the treatment can be explained by the equation;

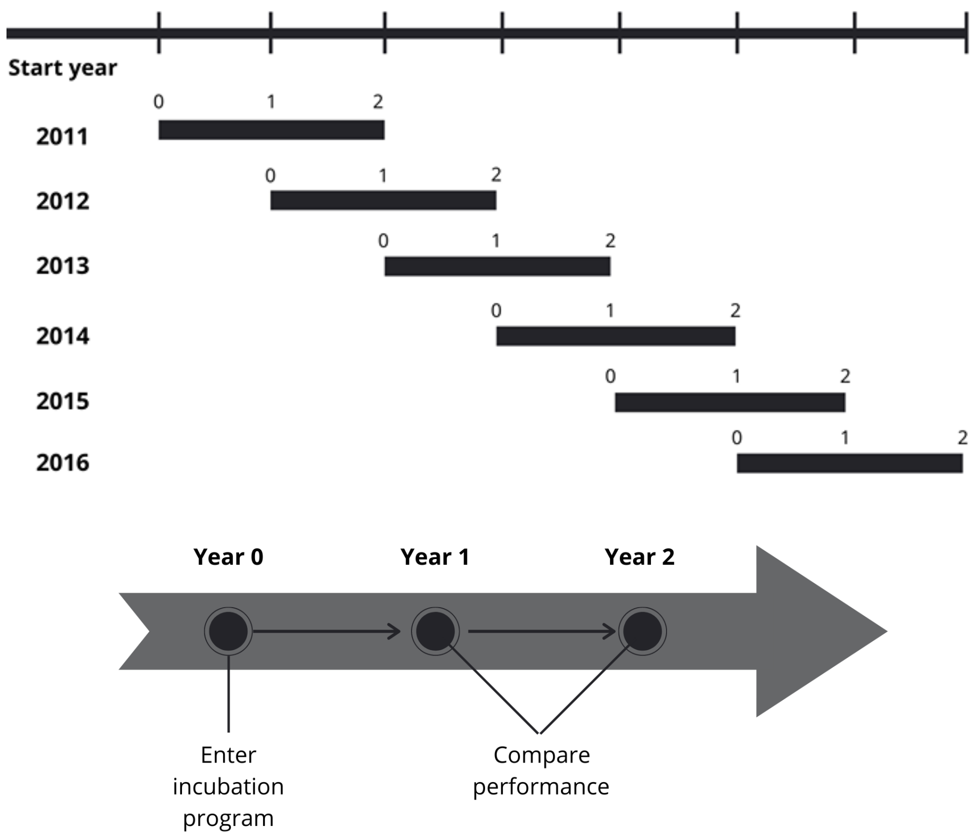
$$ATT(g,t) = E[Y_t(g) - Y_t(0) \mid G=g\ ], \tag{2}$$

The above equation is the average effect on the treatment in the participation of individuals in group g at time period t. If we assume that the parallel trends assumption holds, the group time ATT is identified and can be interpreted to be causal in effect, based on the observable covariates.

To estimate the DiD for causal analysis, we create a *DiD* variable that is multiplied with the treatment variable (0/1) in year one and year two after entry into the incubator, leaving year zero with a value of zero for both groups. This means that the DiD variable is zero for all time periods for the control group, while being zero, one, and one, for the treated group. Even if we do not have a pre-treatment effect by definition, we have entry and post-entry data available, and can still analyse the effect of being in an incubator and the outcome of the periods after treatment in this analysis.

Figure 1 is outlining how the unbalanced panel data is gathered as *one* balanced panel data in the growth sample. Year 0 is the year the treated group enters an incubation program, where the control group has matched with similar characteristics. Year 1 and Year 2 are thus the years we are comparing the performance of the companies in the further analysis.

Figure 1*: *Unbalanced panel data transformation to balanced panel data



The regression analysis for the tables in the next section would normally have values in a two-group, two-period analysis to be interpreted as follows:

- The intercept equals the mean for the control group at time zero
- The treatment coefficient would equal the treatment-control difference at time zero.

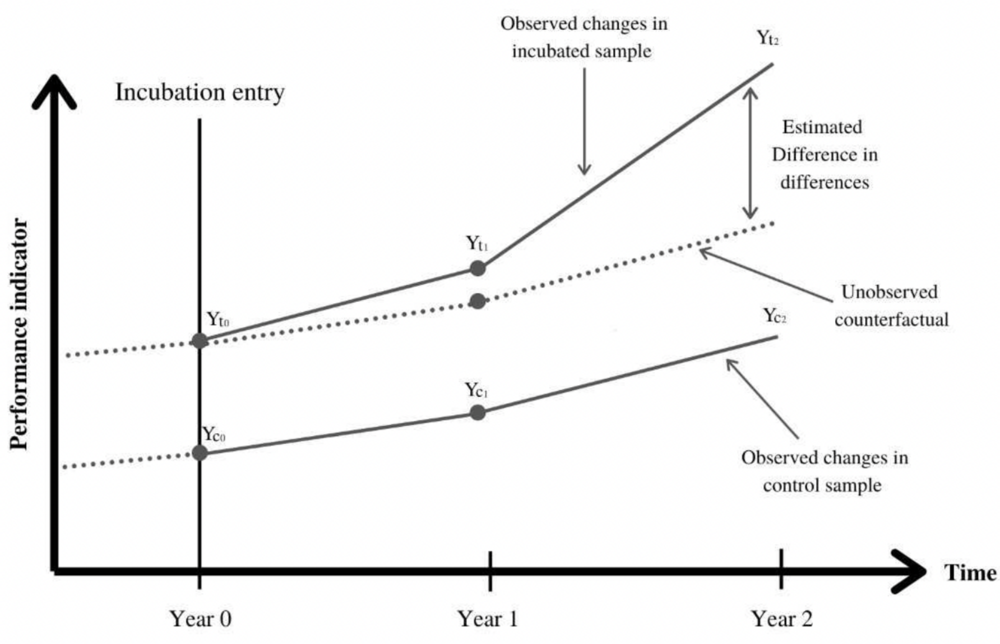
However, we have a multi-period analysis in which time zero is the referent category, and the dummy indicators will reflect the DiD effect for time zero relative to time *t*. The intercept does not equal the mean of the control group, due to our analysis having three periods, and can be interpreted as the best regression line between all three time periods. The regression model we create is in the form of:

$$Y_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D_{ct} + \epsilon_{ict}, \quad (3)$$

Where α is the intercept in the regression model; γ_c is the dummy variable for the treated companies c; λ is the time-dependent variable of the period, which equals 0 in year zero, 1 in year one, and 2 in year two; δ is the DiD estimator, which equals 1 for treated companies in the post-treatment periods, and zero for the control companies in all periods. The DiD coefficient is the only real interpretive causal value in this model.

Figure 2, shows how the DiD coefficient is estimated in our analysis. Since we do not observe the years before entry, we use time period zero as the baseline for the estimation of the performance indicators.

Figure 2: Multiperiod difference-in-difference estimate



For the analysis of the economic performance in the treated and control group, we have the following regression models with their respective dependent variables;

$$\text{Sales revenues}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta\text{Did}_{ct} + \epsilon_{ict}, \quad (4)$$

$$\text{Value creation}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict}, \quad (5)$$

$$\text{Operating profit}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict}, \quad (6)$$

$$\text{Number of employees}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict}, \quad (7)$$

In section 6 we analyse the same performance measures used in the descriptive analysis to obtain the DiD estimates. Here, the control group values are used to obtain a potential counterfactual outcome for the incubated companies, had they not been incubated, as illustrated in Figure 2. We also assume that once a company is incubated, it will remain treated for the subsequent periods. This assumption is called staggered treatment. With this assumption we interpret that the company does not forget about the treatment experience (Callaway a Sant'Anna, 2019, p.2).

4.4 Subsidy regression analysis

In the subsidy regression analysis in section 6.3, we will utilise both the Ordinary Least Squares (OLS) regression analysis with one explanatory variable, and a Multiple Linear Regression (MLR) with robust standard error coefficient test. The standard equation for OLS and MLR regression is given by equation 8;

$$Y_i = \beta_0 + \beta_q X_i + \epsilon_i, \quad (8)$$

Where Y_i is the dependent variable, the number of observations equals i, , X_i equals the explanatory variables, B_0 equal the y intercept (Constant), B_q is the slope coefficient for (each) X_i, independent explanatory variable(s), and ϵ_i equals the random error term.

In our regression analyses in section 6.3, we will have our OLS regressions be the following:

$$\text{Acceptance of Application to IN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_I + \epsilon_i, \quad (9)$$

$$\text{Hit Rate of Application to IN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_I + \epsilon_I, \quad (10)$$

$$\text{Acceptance of Application to RCN, Year}_i = \beta_0 + \beta 1(\text{Treated Incubator})_I + \epsilon_I, \tag{11}$$

$$\text{Hit Rate of Application to RCN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_I + \epsilon_I, \tag{12}$$

We will also include the MLR regressions for each dependent variable in the same regression table, which will include the independent variables used to match, and other variables we deem likely to affect the dependent variable. The MLR regressions are the following;

$$\text{Acceptance of Application to IN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_i + \beta_2(\text{Extra applications IN})_i + \beta_3(\text{Treated RCN Grant, lag})_i + \beta_4(\text{Equity 100k, lag})_i + \beta_5(\text{Debt 100k, lag})_i + \beta_6(\text{Number of employees, lag})_i + \beta_7(\text{Total Income 100, lag})_i + \beta_8(\text{CEO female})_i + \epsilon_{I,} \tag{13}$$

$$\text{Hit Rate of Application to IN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_i + \beta_2(\text{Extra applications IN})_i + \beta_3(\text{Treated RCN Grant, lag})_i + \beta_4(\text{Equity 100k, lag})_i + \beta_5(\text{Debt 100k, lag})_i + \beta_6(\text{Number of employees, lag})_i + \beta_7(\text{Total Income 100k, lag})_i + \beta_8(\text{CEO female})_i + \epsilon_{I,} \tag{14}$$

$$\text{Acceptance of Application to RCN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_i + \beta_2(\text{Extra applications RCN})_i + \beta_3(\text{Treated RCN Grant, lag})_i + \beta_4(\text{Equity 100k, lag})_i + \beta_5(\text{Debt 100k, lag})_i + \beta_6(\text{Number of employee, lag})_i + \beta_7(\text{Total Income 100, lag})_i + \beta_8(\text{CEO female})_i + \epsilon_I, \tag{15}$$

$$\text{Hit Rate of Application to RCN, Year}_i = \beta_0 + \beta_1(\text{Treated Incubator})_i + \beta_2(\text{Extra applications RCN})_i + \beta_3(\text{Treated RCN Gran, lag})_i + \beta_4(\text{Equity 100k, lag})_i + \beta_5(\text{Debt 100k, lag})_i + \beta_6(\text{Number of employees, lag})_i + \beta_7(\text{Total Income 100k, lag})_i + \beta_8(\text{CEO female})_i + \epsilon_{i,} \tag{16}$$

5. Data

This section describes the process used to collect and prepare the dataset used in the research. The main dataset includes incubation data, retrieved either directly from the incubators or Siva. We use data on incubated companies from the time period between 2011 and 2016. Furthermore, we combine accounting and public grant data from 2011 to 2018 to analyse the performance, rate of survival, and access to public subsidies.

5.1 Sample

As described in section 2.2, we identify 11 active business incubators in the Oslo region between 2011 and 2016. Incubators funded through Siva hold available data from start to end on each incubated company. The same is true for some of the incubators not funded by Siva. However, the data collected from the remaining incubators contains different levels of information. For some of the incubators, we were only granted access to information concerning the names of the incubated companies. This requires manually retrieving the organisation number from an online business register, *www.proff.no*. The process is time consuming and requires a substantial amount of manual work. However, it is strictly necessary to ensure the exclusion of incubated companies from the control sample. In a few cases, we also collected data on the incubation period through direct contact with the incubated companies, via email or phone.

For this research we need to supplement the incubation data with accounting data. We gained access to the necessary accounting data from SNF, where annual datasets are received from *Brønnøysundsregistrene* via *Menon Business Economics* and *Bisnode D&B Norway AS* (Berner, Mjøs and Olving, 2016).

In our analysis we will use two different sets of samples. Both samples include a group of incubated companies and a group of control companies matched on a given set of characteristics. The first sample, called the *growth sample*, only includes incubated companies and control companies with continuous accounting data for a period of three years, including the matching year. The second sample, called the *duration sample,* includes all companies with accounting data for the matching year, independent of the number of years surviving after the match.

5.2 Treatment sample

Given the obtained data, a selection process is necessary. From the total of 630 companies attending an incubation program in the Oslo region from 2011–2020, we must filter multiple variables to fit the research design and matching criteria.

The first stage is to exclude the Oslo Cancer Cluster (OCC), as the value creation and ages of these companies interrupt the mean by an exponential degree, since they are outliers. The companies in OCC are capital intensive, often already established, and have no or very few comparable companies in the Oslo region.

Secondly, our research design requires accounting data on all of the assessed companies. Some of the incubated companies do not have publicly available accounting data and are thus excluded from our research. The lack of accounting data on these companies is primarily due to the companies being registered as sole proprietorships or companies with shared responsibility, which do not have public accounting data. Thus, we only include limited liability companies (AS) or public limited liability companies (ASA) in our analysis. Following the exclusion of companies without organisation numbers and all of the companies in OCC, we are left with a total of 545 companies.

To avoid bias in terms of defining the start of an incubation process, we decide not to include companies without information about the year of entry, as this might disrupt our analysis. This is because incubated companies have different ages when they enter into an incubation program. As a consequence, our data only contains the incubation start year for 393 of these 545 companies. In our research, we also only want to examine the years 2011 to 2016 of incubated companies, where we subset the companies only in this period. This leaves us with a total of 204 companies in the ten incubators from which we have gathered information.

Table 2 shows the distribution of incubated companies with available start year, and the number of new companies in incubators, together with the cumulative numbers of total incubated companies.

Table 2: Evolvement of incubated companies in Oslo

Year	New incubated companies	Change from previous year	Total incubated companies	Increase in total companies
2011	1		1	
2012	9	800.00%	10	900.00%
2013	55	511.11%	65	550.00%
2014	41	-25.45%	106	63.08%
2015	34	-17.07%	140	32.08%
2016	64	88.24%	204	45.71%
Total	204		204	

From Table 2, we can see that the number of new incubated companies has increased during the period of the analysis. There are several possible explanations, including an increasing number of incubators, some incubators gaining more popularity, or an increase in innovative companies. The growth of new incubated companies is likely a result of these and other factors.

In Table 3, 204 companies in the current processing sample are grouped into the region of residence in their first reporting year. For this thesis and further processing, we only analyse companies located in the Østviken and Vestviken regions from 2011 to 2016. We can see in Table 3 that some of the incubated companies in the Oslo region are registered in other parts of Norway. Accordingly, 24 companies are incubated in a Viken incubator, but are registered somewhere else. This means that 11.76% of the gathered sample are externally incubated companies. As we discuss later in the limitations section, this means that a control company might be incubated in another city.

Table 3: Regional locations of companies in Oslo incubators

Region	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**Total**
Innlandet			1	1	3	1	2	2	10
Nord-Norge		1	1	1			1	3	7
Sørlandet			1				1	1	3
Trøndelag			2			3	5	4	14
Vestlandet			1	6	1	1	4	5	18
Vestviken			1	1	1	5	4	3	15
Østviken	1	8	48	32	29	54	88	66	326
Total	1	9	55	41	34	64	105	84	393

Excluding the companies not registered in the Østviken and Vestviken region leaves 180 companies. In the aforementioned selection of regions, we also choose to filter out companies not present in one of the municipalities that is a part of the Oslo region. The municipalities considered by this thesis to be in the Oslo region are as follows:

Asker, Bærum, Oslo, Nittedal, Skedsmo, Rælingen, Lørenskog, Oppegård and Ski.

The exclusion of all companies not located in these municipalities leaves a total of 158 companies, which will be used as the basis for the matching process for both samples.

5.3 Final samples

Given the 158 incubated companies in Oslo, we choose various matching procedures to produce the two different sets of samples used in this thesis.

5.3.1 Growth sample

For the growth sample we use all seven covariates, as stated in section 4.1, to match exactly on the coarsened interval bins, finding the closest match by using the nearest Mahalanobis distance to obtain a 1:1 match.

In the growth sample we found a matching control company for 70 of the 158 incubated companies, which is the total treated sample from the Oslo region in 2011–2016. This means that the CEM method has searched 150 708 control group accounting years in the 8-year time period and found a perfect match with 70 of the 150 708 potential accounting year matches. This corresponds with the result of the matching process applied by Fjærli et al. (2018), where almost 50% of the incubator companies were lost due to the inability of the matching process to find any similar control companies. The growth sample therefore left a total of 140 companies, as shown in Table 4.

Table 4: Growth sample, matched companies

	Control	Treated
All	150708	158
Matched	70	70
Unmatched	150638	88

5.3.2 Duration sample

In the duration sample, we find a match for 150 of the 158 treated companies from the Oslo region in 2011–2016. This increased number of matched companies is a consequence of exact matching on fewer covariates. The duration sample therefore left a total of 300 companies.

Table 5: Duration sample, matched companies

	Control	Treated
All	150708	158
Matched	150	150
Unmatched	150558	8

5.4 Public subsidies

From Innovation Norway’s publicly available database, we retrieve data on 44 154 accepted applications in the time period of 2011–2018 (Innovation Norway, 2020d), of which 30 682 are of the grant variety.

Of this sample, Innovation Norway (IN) provided funding to a total of 17,495 unique companies. To analyse the access to government subsidies, we also asked IN to provide the number of companies who applied but did not receive any type of grant. This is referred to as the rejected applications for the same timeframe.

Subsequently, IN provided data on all 12 163 rejected applications. Of these rejected applications, a total of 3 199 grant applications were rejected. A total number of 2 955 companies applied for but did not receive any grants from IN. With this information, we can look closer at all of the companies that applied and received a grant, and those that did not. We have therefore merged the two datasets into one and can examine the acceptance and rejection rates.

We see from Table 6 that IN accepts 90.56%, 9 out of 10, of the applications of the grant variety. Further, we see that roughly 1 in 10 applications is rejected.

Table 6: Number of grant applications to IN

Applications	Number of applications	Percentage
Accepted	30682	90.56
Rejected	3199	9.44
Total	33881	100

According to Table 7, the number of companies that applied and received the grant is lower, at 84.82%. However, 5.05% of the companies had applied multiple times and both received a rejected and accepted application in the time period between 2011–2018. A total of 1 971 companies, or 10.13%, had sent in one or more applications and only been rejected.

Table 7: Number of companies applied for grants to IN

Applications	Number of companies	Percentage
Accepted	16511	84.82
Accepted and rejected	984	5.05
Rejected	1971	10.13
Total	19466	100

From Data Norway (Digitaliseringsdirektoratet, 2020), a directive of publicly available data collections, we retrieved raw data on 11 065 applications to RCN in the time period between 2011 and 2018. Only 4 272, or 38.61%, of these applications were given a grant, while 6 793, or 61.38%, were rejected.

Table 8: Number of applications to RCN

Applications	Number of applications	Percentage
Accepted	4272	38.61
Rejected	6793	61.38
Total	11065	100

As shown in Table 8, of a total of 1 572 companies that applied, 1 090 made at least one accepted application, while 482 companies made only rejected applications. Out of the 1 090 who received a grant, 542 of these companies also made a rejected application. The information from 2011–2018 used in this analysis is available in percentage form in Table 9.

Table 9: Number of companies applied to RCN

Applications	Number of companies	Percentage
Accepted	548	34.86
Accepted and rejected	542	34.48
Rejected	482	30.66
Total	1572	100

5.5 Ethical reflections on data collection and selection

All of the data retrieved for this thesis should be interpreted as if it was not intended for the purpose of this thesis. The data could be biased, contain incorrect information, or have been manipulated for the purpose of its original intended use. Furthermore, parts of the datasets have been entered manually through written and oral communication, which might result in inaccuracies or erroneous data entry.

For instance, if a company receives a grant from IN, it has three years to accept the grant. If the grant is not collected after three years, the remaining amount returns to IN's budget. However, IN does not return to the registers of previous years to update which companies have accepted the grant. This is due to IN using annual datasets as the basis for reporting in their client report, and otherwise reporting to the owner and authorities (D.M. Carrero, personal communication, November 9, 2020).

The reporting standards of the incubators also differ considerably. Some of the incubators only keep the name of their members, while others keep the exact dates of entry and exit, organisation numbers, and names. It is also a possibility that the incubator only provides data on the members they consider to be successful, to maintain their reputation.

The data collected for our research has varying sensitivity to being made publicly available. We regard SNF, approved IN applications, and all data from RCN to be publicly available information. Rejected applications from IN are, however, not public information, and should be regarded as sensitive. Therefore, any company-specific information on the rejected applications is limited in the public version of this thesis.

6. Results

In this section we present the results and interpretation of our analysis. In particular, we describe the results connected to each of our three hypotheses.

The obtained datasets have been used to answer our three hypotheses. We will use DiD estimation after year zero and perform calculations to explain the differences between the treated and control companies. We will also perform statistical tests to evaluate whether the obtained results are significantly different. In addition, regression models are created to evaluate the performance measures of the companies after year zero. An important assumption of the DiD estimation is the fact that the groups need to have common trends in the absence of any treatment. As we have already established, the dataset on incubators does not have any information on pre-treatment, because the companies are often founded in the same year as their entry into an incubator. For this reason, the regression model assumes that the matched set of companies are similar, in regard to having the same age, same accounting year, and same 2-NACE code within each match. This is the closest possibility to analyse the effect of incubation programs in Norway with the available datasets. Consequently, it is not possible to test whether the common trends assumption holds, but rather continue with the analysis on the basis of the companies being matched within the constraints of year zero and accounting year.

6.1 Firm growth

As described in the literature review, Fjærli et al. (2018) found that companies participating in Siva's incubation programs achieved significant additional growth in the effect indicators of sales, number of employees, value creation, and labour productivity. This was true for the first three years after incubation. These findings led us to formulate the following hypothesis:

Companies attending a business incubation program in Oslo outperform non-incubated companies in terms of growth in sales revenues, value creation, operating profit, and number of employees.

Using the growth sample obtained from section 5.3.1, we look at the covariates that have been coarsened to be in the same bin when matching companies with each other. After matching, the exact accounting numbers are used, and in Table 10 we see that the matching is sufficient in finding 70 companies with similar accounting numbers within the accepted range of +/- 20%. One can see that the numbers do not deviate to a limited extent, which is beneficial for comparison between the treated and control group. For the DiD assumptions to hold in the further analyses, we conclude that the observable covariates do not differ from each other, and that the common trend assumption holds for the estimations in the DiD method.

Table 10: Average matching covariate values in entry year

Treated	Rating code	Revenue	Employees	Debt	Equity	Age
0	2.79	690	0.671	359	211	0.947
1	2.75	611	0.632	416	172	0.947

As explained in the methodology (section 4.2.1) we use a continuous period of three years for each company when comparing the effect of being an incubated company versus a control company. We also see that the overlapping values on the matched characteristics in year zero from CEM are sufficient (see Appendix A).

We will also cluster the standard errors in our regressions, the reason being that we have multiple time periods in the DiD analysis. The treatment is assigned at the individual level, and thus we cluster by individual companies, since the unit of randomisation is individual. The method we apply to our regression analysis is to obtain post-estimation “cluster-robust” standard errors proposed by Arellano (1987) for the fixed effects estimator in linear panel models. Our DiD regressor δ is often highly serially correlated, since it will equal zero followed by a string of ones for an individual company that has entered an incubator in the post-treatment years, i.e. staggered treatment (Cameron and Miller, 2015), as explained in section 4.3.1. As described by Bertrand, Duflo, and Mullainathan (2004), using cluster-robust standard errors in DiD settings is of high importance. The clustering should also not be on an individual year level, since the error for company C in 2014 is likely to be correlated with the error for company C in 2013. Accordingly, we will perform a post-regression and cluster our standard errors on the individual level to obtain the cluster-robust standard errors.

6.1.1 Descriptive statistics

Table 11 displays the different values for the covariates from the growth sample, where we have matched the 70 treated and 70 control companies.

Table 11: Descriptive statistics for the economic performance

	Incubated group			Control group			
$N_{incubated} = N_{control} = 70$	Mean	Median	Mean growth	Mean	Median	Mean growth	Mean difference
Sales revenues							
Year 0	568.27	54.50		760.39	160.50		
p = 0.7893	(1093.13)			(1665.94)			
Year 1	1106.01	124.50	94.63%	1187.19	168.00	56.13%	38.50%
p = 0.5742	(2008.19)			(3014.25)			
Year 2	1912.71	233.50	72.94%	1429.71	120.50	20.43%	52.51%
p = 0.2140	(3714.334)			(3496.32)			
Value creation							
Year 0	232.94	-6.00		352.24	10.00		
p = 0.7704	(1075.92)			(806.13)			
Year 1	367.90	0.00	57.94%	495.76	35.50	40.74%	17.19%
p = 0.6887	(1673.80)			(1379.27)			
Year 2	501.20	0.00	36.23%	492.39	59.00	-0.68%	36.91%
p = 04902	(2431.47)			(1769.35)			
Operating profit							
Year 0	-98.44	-15.00		14.76	0.00		
p = 0.8904	(665.55)			(379.82)			
Year 1	-327.70	-51.00		-9.77	-5.50		
p = 0.9822	(1052.87)			(675.33)			
Year 2	-876.42	-25.00		-111.31	-2.00		
p = 0.9929	(2119.51)			(1461.60)			
Number of employees							
Year 0	0.70	0.00		0.71	0.00		
p = 0.5286	(1.11)			(1.24)			
Year 1	2.03	1.00	189.80%	1.19	0.00	66.00%	123.80%
p = 0.0154	(2.62)			(1.88)			
Year 2	3.47	1.00	71.13%	1.51	0.50	27.71%	43.42%
p = 0.0074	(6.03)			(2.68)			

Standard deviations in paranthesis

P- values: t-test, one-sided, diff <= 0

Table 11 presents the descriptive statistics for the two groups, where each group consists of 70 companies, making 140 companies in total. The table contains the mean, median, and mean growth for the four assessment metrics: *Sales revenues, Value creation, Operating profit, and Number of employees.* It also contains the mean difference between the sample of incubated companies and the sample of control companies.

To test for differences in the means of the two samples, we perform a Welch two-sample t-test on each of the four assessment metrics, for each year. The one-tailed t-test is used because we are interested in whether the incubated companies are performing better than the control companies. The assessment metrics are better if they are higher, i.e. positive in comparison. The p-values for the t-tests are displayed under each of the years, on each assessment metric. We can reject the null hypothesis if the p-value is less than 0.05, i.e. the p-values show that the two means of the groups are significantly different, and we can reject the null hypothesis stating that there are no significant differences between the means of the two groups. Alternatively, we can fail to reject the null hypothesis, i.e. the p-values do not show that the treated group has significantly higher means than the control group.

Therefore, our null hypothesis is that of

$$H_0: m_T <= m_C , \tag{17}$$

and the corresponding alternative hypothesis is

$$H_A: m_T > m_C . \tag{18}$$

The null hypothesis in our analysis is thus whether the mean of group Treated is less than or equal to the mean of group Control, while the alternative hypothesis is that the mean of group Treated is greater than the mean of group Control.

The p-values for sales revenues, value creation, and operating profit shown in Table 11 have p-values greater than 0.05 in each year. This means that we cannot reject the null hypothesis – that the difference in means between the incubator sample and the control are less than or equal to zero. The high p-values close to ~0.9 do not, however, mean that the two groups are similar; but only tell us that the hypothesis and covariate that we are testing for fail to reject the null hypothesis.

However, for the number of employees we see significant p-values in years 1 and 2. Therefore, we can reject the null hypothesis – that the difference in means between the incubated sample and the control sample is less than or equal to the incubated companies. This indicates that the number of employees in incubated companies is significantly higher in years 1 and 2 compared to the control companies. These results are consistent with our hypothesis that incubated companies perform better than the control companies in terms of the number of employees.

6.1.2 Sales revenues

Considering the result of the regression Table 12 for sales revenues, we interpret the regression output from the regression model:

$$\text{Sales revenues}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta\text{Did}_{ct} + \epsilon_{ict} \quad (4)$$

The sign of the regression coefficient means that there is a positive or negative relationship between each of the independent variables on the dependent variable. In this regression, the dependent variable is Sales revenues. The coefficient Constant has a value of 696.690 and is the estimated straight-line equation of the Sales revenues in our regression model in year zero. The coefficient Treated Incubator, -128.419, is the estimated mean change in sales revenues among the treatment group in all periods. It reflects the difference in sales revenues by being an incubated company compared to the control group, not taking into account the post-treatment effects. The coefficient Time Period, 297.486 is the estimated mean change in sales revenues for both the treated and control group after entry year, i.e. the pure effect of the passage of time in the total sample.

The coefficient Difference-in-difference*,* 297.486, estimates that the treated group, after entry year, increases its sales revenues relative to what it would have been, had they not been incubated in the post-period. This can be interpreted as the DiD estimation of how the treated group is performing relative to its unobserved counterfactual outcome.

We can interpret the p-value for the Difference-in-difference coefficient as being not significant, i.e. not rejecting the null hypothesis of differences lower or equal in the means. We can therefore say that there are no statistically significant differences in the means between the two, as well as after adjustment for the cluster-robust standard errors on the individual level. According to the causal effect analysis, there is no evidence that the companies which have been in an incubator have significantly different sales revenues in comparison to its counterfactual outcome. In Figure 3 we can also see that the parallel path assumption holds. In turn, this means that the estimated ATT is estimated to be positive and increase sales revenues by 297.486, on average for the treated companies.

Figure 3: Mean total Sales revenues, treated and control group

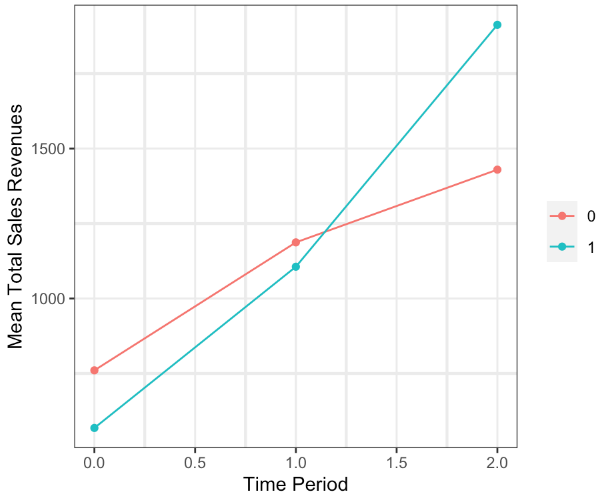


Table 12: Regression results, Sales revenues

	Dependent variable:	
	Sales Revenues	Clustered Standard Errors
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	−128.419	−128.419
	(420.332)	(281.358)
Time Period	429.071**	429.071***
	(201.921)	(113.612)
Difference-in-difference	297.486	297.486
	(494.603)	(234.858)
Constant	696.690**	696.690***
	(273.402)	(242.512)
Observations	420	
R^2	0.024	
Adjusted R^2	0.017	
Note:		*p<0.1; **p<0.05; ***p<0.01

6.1.3 Value creation

Our findings for the regression output for value Creation is interpreted from the regression model:

$$\text{Value creation}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict} \qquad (5)$$

In this regression, the dependent variable is Value creation, while the coefficient Constant has a value of 364.078. The coefficient Treated Incubator, -131.135, is the estimated mean change in value creation among the treatment group in all periods. It reflects the difference in value creation by being an incubated company compared to the control group. The coefficient Time Period, 82.717, is the estimated mean change in the outcome after entry year, i.e. the pure effect of the passage of time in the total sample.

As shown in the regression Table 13, we get a Difference-in-difference coefficient of 77.531 for value creation. The positive coefficient indicates that the treated group after entry into an incubator increases its value creation by this amount. After year zero, this can be interpreted as the DiD estimation of how the treated group is performing in an incubator relative to the counterfactual outcome after entry year. This implies that treated companies in incubators exhibit better performance in terms of growth, i.e. operating profit and salaries.

We can interpret the p-value for the Difference-in-difference coefficient as being not significant, i.e. not rejecting the null hypothesis of no differences in means. We can therefore say that there are no statistically significant differences in means between the two outcomes after entry year, or after adjustment for the cluster-robust standard errors on an individual level. The causal effect analysis can be interpreted as supplying no evidence that companies which have been in an incubator have a significantly higher value creation in comparison to its counterfactual outcome.

Figure 4: Mean total Value creation, treated and control group

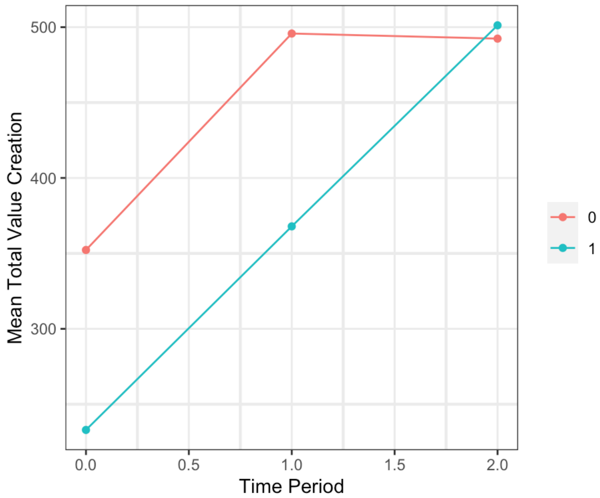


Table 13: Regression results, Value creation

	Dependent variable:	
	Value Creation	Clustered Standard Errors
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	−131.135	−131.135
	(252.793)	(177.474)
Time Period	82.717	82.717
	(121.438)	(70.853)
Difference-in-difference	77.531	77.531
	(297.461)	(155.164)
Constant	364.078**	364.078***
	(164.428)	(114.321)
Observations	420	
R^2	0.003	
Adjusted R^2	−0.004	
Note:		*p<0.1; **p<0.05; ***p<0.01

From Figure 4 we can also see that the parallel path assumption holds, where the treated companies continue their increase in value creation, and where the control group flattens out in time period two.

From the descriptive statistics in the growth sample, as discussed in section 6.1.1, we can also see that the mean number of employees is 3.47 for the treated group, compared to 1.51 for the control group in year two. This is a mean difference of 1.96 employees between the two groups. From Statistics Norway (2020b) we obtain information about the average monthly wage rate in Norway, which is 47 290 NOK, totalling 567 480 NOK a year. The value creation in our analysis is calculated as operating profit plus salaries. Consequently, the difference in salaries between the two groups, based on statistics from Statistics Norway, can be calculated as 1 112 260 NOK on average between the two groups in year 2. One interpretation of these values is that incubated companies employ more people, resulting in increased value creation, but at the same time the operating profit for treated companies in year two could be lower compared to the control group. Another interpretation is that the control companies have higher operating profits, even with fewer employees.

6.1.4 Operating profit

The regression model for operating profit is as follows:

$$\text{Operating profit}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict} \quad (6)$$

In this regression, the dependent variable is Operating profit. The coefficient *Constant* has a value of 124.731. The coefficient Treated Incubator, -219.31, is the estimated mean change in operating profit among the treatment group in all periods. It reflects the difference between being in an incubator program and the control group not taking into account the post-treatment effect. The coefficient Time Period, -160.174 is the estimated mean change in operating profit after year zero, i.e. the pure effect of the passage of time in the total sample.

The Difference-in-difference coefficient, -263.360, estimates that the treated group after entry year decreases their operating profit relative to its potential counterfactual outcome. This can be interpreted as the treated group having on average lower operating profit compared to not being incubated.

We can interpret the p-value for the Difference-in-difference coefficient, presented in Table 14, as being significant on the 10% level after cluster-robust standard errors on individual companies. We can reject the null hypothesis that there is no difference between the means of the treated and its potential counterfactual outcome at the 10% significance level. We can therefore conclude that there exists a significant difference in the means between the two. The incubated companies have a large negative multiplier in the later time periods, which indicate a significant lower operating profit by being incubated.

Developing a product or business to be profitable in the long run may explain why the incubated companies have a lower operating profit, in comparison its potential outcome. Marketing, operating expenses, and depreciation might be higher for the incubated companies, because many features need to be established before realising their long-term plans while they are or have been in an incubator. These plans might differ from those of the control companies, who may not have such great “long-term” plans for their companies. The operating profit is calculated as earnings before interest and taxes, and as we understand it the incubator “pushes” companies to grow, therefore spending more money in the short term to obtain a gain in the long term. Another reason might be that the incubated companies have easier access to investors, that can afford to fund a short term negative operating profit. Unavoidably, there may be expenses and investments in terms of prototypes, testing, and marketing that need to be made while being in an incubator. We cannot conclude that the control group does not experience the same activities; but at a 10% significance level we can say that the treated group differs significantly in terms of their lower operating profit compared to the potenial outcome had they not been incubated after the entry year.

Figure 5: Mean total Operating profit, treated and control group

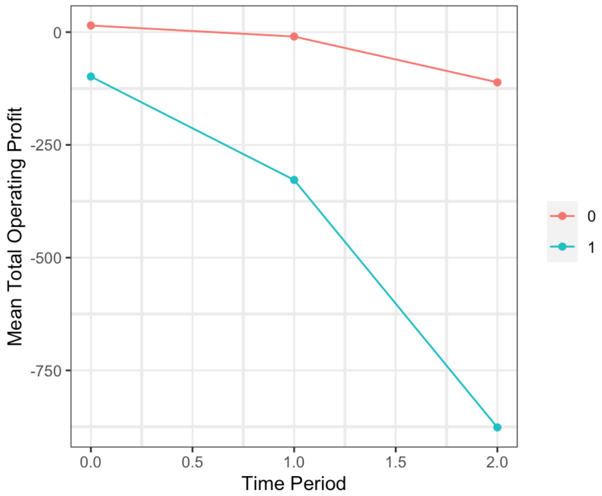


Table 14: Regression results, Operating profit

	Dependent variable:	
	Operating Profit	Clustered Standard Errors
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	−223.174	−223.174**
	(190.944)	(107.181)
Time Period	−160.174*	−160.174**
	(91.727)	(72.450)
Difference-in-difference	−263.360	−263.360*
	(224.684)	(141.544)
Constant	124.731	124.731*
	(124.199)	(64.871)
Observations	420	
R^2	0.051	
Adjusted R^2	0.044	
Note:		*p<0.1; **p<0.05; ***p<0.01

6.1.5 Number of employees

The regression model for Number of employees is as follows:

$$\text{Number of employees}_{ict} = \alpha + \gamma_c(\text{Treatment}_c) + \lambda(\text{Period}_t) + \delta D\text{id}_{ct} + \epsilon_{ict} \qquad (7)$$

In this regression, the dependent variable is number of employees. The coefficient *Constant* has a value of 0.530 and the coefficient Treated Incubator, 0.170, is the expected linear mean change in number of employees among the treatment group in all periods. It reflects the difference between being in an incubator program compared to the control group, not taking into account the effect of post-treatment. The coefficient Time Period, 0.609, is the expected linear mean change in number of employees after year zero for all companies after year zero, i.e. the pure effect of the passage of time in the total sample.

As we can see in Table 15, the coefficient for Difference-in-difference, 1.137, estimates that the treated group after entry year increases their number of employees relative to its potential counterfactual outcome in the same post-period. The estimated Difference-in-difference coefficient may imply that the treated group employs more people relative to what they would have employed, had they not been incubated.

We can interpret the p-value for the Difference-in-difference coefficient as being significant at the 5% significance level, and can reject the null hypothesis of no differences in means. We can therefore say that there are significant differences in the means between the two outcomes in terms of number of employees. The significance level of 5% is the probability of rejecting the null hypothesis when it is in fact true. For a significance level of 0.05, this means that there is a 5% risk of concluding that a difference in means exists when in fact there is no difference.

Figure 6: Regression results, Number of employees

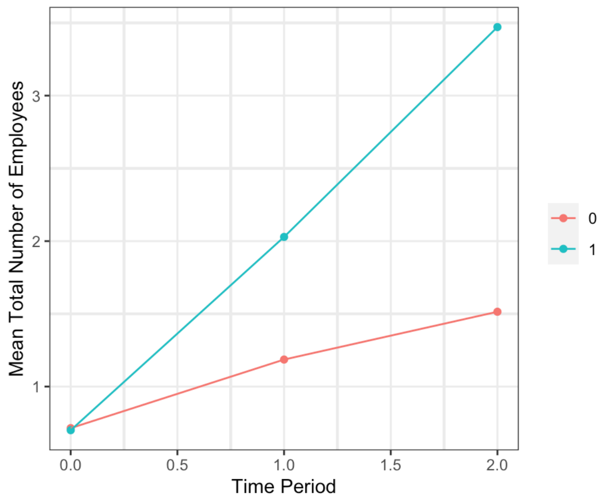


Table 15: Mean total Number of employees, treated and control group

	Dependent variable:	
	Number of Employees	Clustered Standard Errors
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	0.170	0.170
	(0.484)	(0.235)
Time Period	0.609***	0.609***
	(0.233)	(0.144)
Difference-in-difference	1.137**	1.137***
	(0.570)	(0.365)
Constant	0.530*	0.530***
	(0.315)	(0.189)
Observations	420	
R^2	0.082	
Adjusted R^2	0.075	
Note:		*p<0.1; **p<0.05; ***p<0.01

As shown in Figure 6, the means of the two groups after year zero are different, and according to the descriptive statistics presented in section 6.1.1, the t-test statistic for the difference between the means is less than or equal to zero, and the p-value is 0.0074, indicating a rejection of the null hypothesis that the two groups are similar. When comparing the histograms of the number of employees at years zero, one, and two (see Appendix B), we also see that there are more treated companies in the higher count of the number of employees. A possible reason for this significant difference is that we only have 70 companies. Without the one company on the far right with 40 employees, the mean, 3.47, would be lower, at 2.94 as shown in equation 18.

$$70\ x\ 3.47 = 242.9\ |\ \frac{242.9 - 40}{69} = 2.94 \tag{18}$$

With this, we can say that the mean is impacted by individual companies with a higher number of employees than the rest of the treated group in the sample. Nonetheless, the treated group have companies with more employees than the control group, and we can say that there is a significant difference in the means between the incubated companies outcome and its counterfactual outcome at the 5% significance level.

6.2 Survival rate

Differences in survival rates between companies attending incubator programs have been analysed in several research papers (e.g. Ferguson and Olofsson, 2004; Hackett and Dilts, 2004). As described in the literature review (section 2.1), Ferguson and Olofsson (2004) found that 93.3% of the studied companies located in science parks survived their first 7 years, compared to 66.7% in the off-park sample. In their systematic review of business incubation research, Hackett and Dilts (2004) found that the level of incubator development and the number of incubated companies are positively related to the survival of incubated companies. This led us to formulate the following hypothesis:

Companies attending a business incubation program in Oslo survive longer than companies with similar characteristics that have not attended an incubation program in Oslo.

Statistics Norway (2020a) finds that only 28.4% of all companies established in 2013 survived until 2018. As shown in Table 16 and table 17, the companies in our duration sample have a substantially higher survival rate. A possible reason is that we are only considering the matching year, i.e. entry year into an incubator as year zero, in our samples, while some of the companies may be older than a year.

Table 16 and table 17 show the survival rate of incubated companies and non-incubated companies, respectively. The survival rates are calculated for the first five years after matching, or until 2018, which is the last year of our accounting data. At the bottom of the two tables, we calculate the weighted average survival rate for each year after matching.

Table 16: Percentage survival rate of incubated companies, t years after entry

Year	Year 1	Year 2	Year 3	Year 4	Year 5	Number of Companies
2011	100%	100%	100%	100%	100%	1
2012	100%	100%	100%	100%	80.00%	5
2013	95.45%	88.64%	81.82%	77.27%	72.73%	44
2014	100%	92.86%	82.14%	82.14%		28
2015	95.83%	91.67%	79.17%			24
2016	100%	97.92%				48
Weighted Average	98.00%	93.33%	82.35%	80.77%	74.00%	

As shown in table 16 and 17, the weighted average survival rate is higher for the incubated companies in years 1, 2, and 4 after matching. However, in year 3, the control sample has a higher weighted average survival rate. In year 5 the two groups have the same weighted average. Thus, these results might indicate that the short-term effect on survival is positive for incubated companies. We see from the 5-year rate that there are no differences between the incubated group and the control group.

Table 17: Percentage survival rate of control companies, t years after entry

Year of matching	Year 1	Year 2	Year 3	Year 4	Year 5	Number of Companies
2011	0.00%	0.00%	0.00%	0.00%	0.00%	1
2012	100%	80.00%	80.00%	80.00%	80.00%	5
2013	95.45%	86.36%	81.82%	77.27%	75.00%	44
2014	96.43%	96.43%	89.29%	85.71%		28
2015	95.83%	95.83%	87.50%			24
2016	100%	95.83%				48
Weighted Average	96.67%	92.00%	84.31%	79.49%	74.00%	

However, the survival rate used in this thesis does not take the reason for the exit of the companies into account. Ferguson and Olofsson (2004), in their study on science parks in Sweden, found that of the on-park companies that did not survive, half were a result of M&As. In the off-park sample, on the other hand, only a third of the companies that did not survive were the result of M&As. Thus, the survival rates witnessed in our study do not offer full disclosure of the success or failure of the companies.

In their systematic review of incubation research, Hackett and Dilts (2004) found that the level of incubator development and the number of incubated companies are positively related to incubated companies' survival. The business incubators in our sample were all founded between 2011 and 2016, which might imply that the incubators are too early in their development to positively affect the survival rate of the incubated companies.

Nonetheless, there we find no clear evidence that participation in an incubator program results in higher survival rates, compared to non-incubated companies.

6.3 Access to public subsidies

In their paper on incubator effectiveness in Italy, Colombo and Delmastro (2002) found that companies located in business incubators or science parks had easier access to public financial funds. Their analysis revealed that 51% of the on-incubator companies received public subsidies, compared to 33% in the off-incubator sample. Fjærli et al. (2018) also found that 25% of the companies participating in Siva's incubation program received funding or support from at least one other public scheme.

The findings in these papers combined with our own experiences led to the development of our third hypothesis:

Companies attending a business incubator in Oslo have a better chance of receiving public subsidies, compared to non-incubated companies.

For the analysis of access to public subsidies, we use the same duration sample as we used to compare the survival rate in section 6.2. However, we now include only those companies that applied for a grant to either IN or RCN, to compare similar companies seeking funding for their business. The analysis will focus on years zero to four, since some of the grants are only provided to companies younger than 5 years, and prioritise companies younger than 3 years. Most of the grants that do not specify company age are, however, targeted at companies younger than 5 years (Innovation Norway, 2020e).

Given the duration sample from the matching procedure, we have a total of 300 companies, with 150 in each sample. The number of companies from the duration sample that have applied for one of these grants is shown in Table 18.

Table 18: Companies applied to public subsidies

Firm	Companies applied to IN	Companies applied to RCN
Treated	80	41
Control	14	7
Total	94	48

From the total sample of 300 companies, 31.33% applied to IN and 16% applied to the RCN. There are also large differences between the two groups regarding the number of applicants. Of the incubated companies, 53.33% applied for IN grants at least once, while only 11.33% of the control companies did the same. For grants from RCN, 27.33% of the incubated companies applied, compared to 4.67% of the control companies.

6.3.1 Innovation Norway

Of the companies that have applied to IN, 85.11% are in the incubated group, and 14.89% are in the control group. Our findings suggest that the treated group is on average ~5.7 times more likely to apply to any type of funding from the IN's grants available used in this analysis.

The number of applications for each of the groups is also relevant, as the treated companies more often apply more times on average compared to the control group. Table 19 presents the number of applications submitted and approved for each of the two groups. These findings indicate that of the 80 incubated companies and 14 control companies applying, each company submits on average 1.73 and 1.64 applications, respectively. The table also

Table 19: Number of accepted and total applications, IN

Firm	Applications to IN
Treated Accepted	129
Num Applications	138
Control Accepted	19
Num Applications	23
Total	161

indicates that 93.48% and 82.61% of the applications submitted by the incubated and control companies, respectively, are approved.

Looking at the number of companies applying for IN grants from years zero to year four, we can see from Table 20, that most companies apply in their first three years. This suggests the motivation to apply for IN's grants while they are in the prioritised group of companies, that is, younger than three years of age.

Table 20: Number of companies applied to IN, per year after entry

Firm	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
Treated Companies	54	35	19	4	6	1
Control Companies	8	3	5	2	1	0
Total	62	38	24	6	7	1

Furthermore, the application process to IN does not limit companies from applying for any other grant types in the same year, as shown in Table 21. In the first three years there are more applications than there are companies, indicating that both the control and treated companies apply for several types of grants in the same year. The applications include the following grant types: market clarification, commercialisation, commercialisation – phase 2, innovation contracts, bioeconomy projects, environmental technology, bioeconomy projects, and ecosystem grants.

Table 21: Number of accepted and total applications, IN, per year after entry

Firm	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
Treated accepted	63	37	20	4	4	1
Num applications	67	37	23	4	6	1
Control accepted	10	2	5	1	1	0
Num applications	11	3	5	2	2	0
Total applications	74	40	28	6	8	1

To analyse the difference between the two groups, we must evaluate which models we can use. To analyse the differences between the two groups, we must evaluate which models we can use. To use a t-test we must assume normally distributed values in the dataset being analysed. The dependent variables we have get the null hypothesis of normally distributed data rejected using a Shapiro–Wilk normality test, and is less than the significance level of

0.05. This implies that the distribution of the data is significantly different from the normal distribution. We can therefore not assume normality, and thus not use F-test nor the t-test. When the data is not normally distributed, one can use the non-parametric two-sample Wilcoxon rank test, which produces a p-value of 0.4525 for the IN dataset. The p-value indicates that there is no significant statistical difference in the number of companies receiving the grant between the treated and the control group.

Regression analyses

We compute two different regression analyses: one based on whether the company has received an IN grant in one year, and one that measures the *hit rate*, defined as the number of accepted applications in a year divided by the total number of applications sent to IN during that year.

That is, company i sends y applications in time period x, where y equals the number of accepted applications w plus the number of rejected applications q. The hit rate z is thus defined as w divided by y, and is the percentage of accepted applications made by the treated and control group.

Our first regression is performed on the dependent variable, Acceptance of Application to IN, year; which has the value of either 0 or 1 dependent on whether a company has their application(s) rejected or accepted in a year. We have also obtained unbiased standard errors of the MLR coefficients, where we include the extra coefficients in the regression analysis.

It seems likely that participation in an incubator is not the only variable affecting whether a company receives grants from IN. Both observable and unobservable variables could affect the probability of a company receiving public grants. Thus, we test for extra applications, equity, total income, number of employees, and debt as control variables. However, testing for accounting numbers from the same year that a company receives the grants is not helpful, as the grants would likely increase all or some of these variables in the year of receiving grants. Thus, the *lagged value*, i.e. the value in the last application year before treatment, should be used. However, as a large share of the companies receive their grants in the first year, using lagged values will result in missing numbers in 46% of the cases for the IN-dataset, as shown in Table 22. We also include a lagged value of RCN treatment in the regression.

In the regression output in Table 22, we can see the two different regressions: one including the lagged values of the control variables, and the other where only participation in an incubator is used. From the first regression, the participation coefficient Treated Incubator shows a non-statistically significant effect of 4.6 percentage points increase in application acceptance.

Table 22: Regression result, Acceptance of Application to IN

	Dependent variable:	
	Acceptance of Application to IN, Year	Standard Errors MLR
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	0.046	0.118
	(0.061)	(0.130)
Extra applications IN		−0.047
		(0.127)
Treated RCN Grant, lag		0.120**
		(0.053)
Equity 100k, lag		0.001
		(0.001)
Debt 100k, lag		−0.0004
		(0.002)
Number of employees, lag		−0.007
		(0.010)
Total income 100k, lag		−0.0001
		(0.001)
CEO Female		0.020
		(0.077)
Constant	0.895***	0.812***
	(0.057)	(0.142)
Observations	138	
R^2	0.004	
Adjusted R^2	−0.003	
Residual Std. Error	0.248 (df = 136)	
F Statistic	0.574 (df = 1; 136)	
Note:		*p<0.1; **p<0.05; ***p<0.01

When including the observable covariates, we notice that the coefficient Treated by RCN Grant, lag, is statistically significant on the 5% level. This may imply that a company being treated by RCN in their last application year, experiences increased acceptance of an application from IN by 12 percentage points. Our findings also suggest that extra applications to IN negatively affect whether a company gets treated in a year, by 4.7 percentage points. This is, however, not statistically significant. The other lagged values do not influence the acceptance of an application by any large numbers. Considering the coefficient Constant*,* this means that on average 81.2% of the control companies receive a grant from IN, not including the other coefficients. The coefficient female CEO also increases acceptance by 2 percentage points but is not statistically significant. We can also see that the Treated Incubator coefficient shows an increase of 11.8 percentage points in application acceptance during a year with the extra covariates.

The second regression in Table 23, is based on the percentage of applications sent by a company to IN in a specific year that was accepted. A hit rate is a number between 0 or 1, mostly one of these, as the application sent by most companies are for one grant at a time. As shown in Table 23, the first OLS regression again shows that the treated incubator coefficient increases the dependent variable hit rate by 9.5 percentage points.

Table 23: Regression result, Hit Rate of Application to IN

	Dependent variable:	
	Hit Rate of Application to IN, Year	Standard Errors MLR
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	0.095	0.167
	(0.063)	(0.128)
Extra applications IN		−0.102
		(0.136)
Treated RCN Grant, lag		0.138**
		(0.056)
Equity 100k, lag		0.001
		(0.001)
Debt 100k, lag		0.0001
		(0.002)
Number of employees, lag		−0.008
		(0.010)
Total income 100k, lag		−0.0004
		(0.001)
CEO Female		0.034
		(0.076)
Constant	0.842***	0.763***
	(0.058)	(0.139)
Observations	138	
R^2	0.016	
Adjusted R^2	0.009	
Residual Std. Error	0.254 (df = 136)	
F Statistic	2.278 (df = 1; 136)	
Note:		*p<0.1; **p<0.05; ***p<0.01

When including the observable covariates, we notice that the coefficient Treated by RCN, year, is again statistically significant at the 5% level when measuring the hit rate of accepted applications. We can also notice that the number of extra applications to IN decreases the hit rate by 10.2 percentage points. The Treated incubator coefficient indicates a positive effect of 16.7 percentage points on the hit rate, suggesting that the treated companies have applications accepted more often than the control companies. The lagged accounting values have no significant effect on the difference between the two groups, as seen in the regression Table 23.

The coefficient CEO female, which has a binary value of 1 if the CEO is a female and 0 if it is a male, indicates that companies with female leaders experience increased acceptance of their applications to IN by 3.4 percentage points, although the coefficient is not statistically significant.

6.3.2 The Research Council of Norway

Of the companies that have applied to RCN, 85.42% are in the treated group and 14.58% are in the control group. Our findings show that the treated group applies on average ~5.9 times more to any type of funding from the RCN grants available in this analysis. The number of applications to RCN reaches a total of 88. To the available RCN grants, the treated group submits on average 1.95 applications, while the same number for the control group is 1.14. This implies that of the companies who apply to RCN, the treated group sends 1.7 times more applications per company than the control group. Table 24 also indicates that 53.75% and 62.5% of applications sent from the treated and control group, respectively, are accepted.

Table 24: Number of accepted and total applications, RCN

Firm	Applications to RCN
Treated Accepted	43
Num Applications	80
Control Accepted	5
Num Applications	8
Total	88

Furthermore, looking at the companies who applied (Table 25), it is evident that these companies are more active in the first three years after the entry year. This may be explained by the possibility to obtain funding for their product in the early development phase of the product. The RCN awards grants to innovative projects and organisations that need research access to develop their products.

Table 25: Number of companies applied to RCN, per year after entry

Firm	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
Treated Companies	19	18	16	5	5	2
Control Companies	1	1	3	0	2	1
Total	20	19	19	5	7	3

From Table 26 we can see that applications sent in the first three years are accepted ~50% of the time. This may indicate that the RCN are more rigorous in their evaluations, or expect more of the companies in comparison to IN. While RCN and IN both give out grants to innovative companies, RCN focuses more on research-based companies. The total number of applications per company per year is also larger than 1, implying that companies may apply for more than one grant per year. RCN has different types of applications, which include grants for research projects, innovation projects, competence and collaboration projects, coordination and support activities, and commercialisation projects, and which are mostly for companies with a research-focused background.

Table 26: Number of accepted and total applications, RCN, per year after entry

Firm	Year 0	Year 1	Year 2	Year 3	Year 4	Year 5
Treated accepted	12	9	12	3	5	2
Num applications	25	22	20	5	6	2
Control accepted	1	2	3	0	0	0
Num applications	1	3	3	0	2	1
Total applications	26	25	23	5	8	3

To be able to run a regression on the differences between the two groups, we must evaluate the models that can be used. The dependent variables we have, gets the null hypothesis of normally distributed data rejected using a Shapiro−Wilk normality test. We find that we must reject the null hypothesis that there is a normal distribution of the values. When the data is not normally distributed, we can use the non-parametric two-sample Wilcoxon rank test, which produces a p-value of 0.7728, for the RCN dataset. The p-value indicates that there are no significant differences in the number of companies receiving a RCN grant, when comparing between the control and treated groups.

Regression analyses

We compute two different regression analyses: one based on whether a company received a grant in a given year, and one where we measure the hit rate of applications sent to RCN in a year.

The first regression shows the regression on the dependent variable, Acceptance of Application to RCN, year, which has the value of 0 or 1 in the year of application for an RCN grant. That is, if a company has at least one application approved in that year, the value becomes 1.

From Table 27 we see that the Treated Incubator coefficient does not have any significant effect on whether a company has an application accepted in a given year. However, it indicates that the treated companies are on average 5.6 percentage points less likely to have an application accepted compared to the control group average, as indicated by the first regression Constant coefficient of 62.5.

Table 27: Regression result, Acceptance of Application to RCN

	Dependent variable:	
	Acceptance of Application to RCN, Year	Standard Errors MLR
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	−0.056 (0.188)	−0.105 (0.220)
Extra applications RCN		0.367*** (0.096)
Treated IN Grant, lag		0.080 (0.154)
Equity 100k, lag		−0.001 (0.001)
Debt 100k, lag		0.002* (0.001)
Number of employees, lag		−0.003 (0.009)
Total income 100k, lag		0.002 (0.002)
CEO Female		0.068 (0.207)
Constant	0.625*** (0.177)	0.507** (0.213)
Observations	73	
R^2	0.001	
Adjusted R^2	−0.013	
Residual Std. Error	0.501 (df = 71)	
F Statistic	0.088 (df = 1; 71)	
Note:		*p<0.1; **p<0.05; ***p<0.01

When including the observable covariates, namely the number of extra applications, lagged value of Treated IN grant, capital, debt, number of employees, total income, and whether the CEO is female, we see that the number of applications significantly increases the chance of getting a grant. This may be interpreted as follows: if a company sends more than one application, there is an 89.8% chance of having one of them accepted in the year of application. Our findings also suggest that most of the lagged values do not have any significant effect, also changing the values minimally. However, we can see that the lagged debt value per 100 000 NOK is significant at the 10% significance level. The 0.2 percentage point increase in acceptance of applications per 100 000 NOK debt may occur because companies with higher debt apply for grants to help obtain better liquidity in terms of their future growth plans. The lagged value of being treated by IN in the previous applications year has a coefficient of 0.08, indicating an increased application acceptance of 8 percentage points, although this is not statistically significant. The coefficient CEO female indicates that companies with female leaders experience increased acceptance of their applications by 6.8 percentage points. Again, this variable is not statistically significant.

The second regression analysis shows the dependent variable, hit rate. Here we check whether the control group or the treated group experiences a higher acceptance rate of their applications to RCN. The hit rate can take multiple values from zero to one, e.g. 0, ¼, ⅓, ½, or 1, depending on the number of applications accepted in a year divided by the total number of applications. From Table 28 we can again see that the Treated Incubator coefficient does not have any statistically significant effect on the outcome of hit rates between the two groups. However, it indicates that by being an incubated company, there is a decrease of 11.7 percentage points on the dependent variable. The constant of 62.5 is the average hit rate for the control companies, from the first regression on hit rate.

Table 28: Regression result, Hit Rate of Application to RCN

	Dependent variable:	
	Hit Rate of Application to RCN, Year	Standard Errors MLR
	OLS	*coefficient test*
	(1)	(2)
Treated Incubator	−0.117	−0.104
	(0.179)	(0.225)
Extra applications IN		0.154
		(0.126)
Treated RCN Grant, lag		0.092
		(0.157)
Equity 100k, lag		−0.001
		(0.001)
Debt 100k, lag		0.002
		(0.001)
Number of employees, lag		−0.004
		(0.009)
Total income 100k, lag		0.002
		(0.002)
CEO Female		0.041
		(0.213)
Constant	0.625***	0.497**
	(0.169)	(0.218)
Observations	73	
R^2	0.006	
Adjusted R^2	−0.008	
Residual Std. Error	0.477 (df = 71)	
F Statistic	0.430 (df = 1; 71)	
Note:		*p<0.1; **p<0.05; ***p<0.01

When including the observable variables we see that there are no statistically significant coefficients that influence the hit rate. Contrary to our first regression, concerning whether the company is treated in a year, the extra number of applications is not statistically significant. This is understandable, because hit rate is the number of accepted applications divided by the number of applications. It does however have a positive sign, with an increase of 15.4 percentage points on the hit rate, indicating that a higher number of applications increases the acceptance rate in comparison to only applying once a year. The lagged Treated IN Grant, year value also shows that the hit rate acceptance of applications increased by 9.2 percentage points if the company had an application approved by IN the last time it submitted one.

6.3.3 Remark on public subsidies

One possible explanation of our findings that companies treated by RCN in their previous application year, receive significantly more IN grants, could be because they view support from others as a sign that someone else believes in the idea of the company, and thus may act as a risk-decreasing factor.

Furthermore, from our own experience with both IN, RCN, and a Bergen-based incubator, the different institutions encourage applicants to list participation and grants from other institutions, which show that others believe in the idea and the team behind it. This is an unobservable effect that we cannot observe in our data, but as the IN dataset suggests, it is confirmed by a statistical significance level of 5%.

Another possible explanation for the differences in the number of applications sent to public subsidy grants, is that the incubated companies are encouraged to apply and receive guidance with writing applications to IN and RCN from advisors and other members of the incubators. In this way, they may gain an extra advantage over non-incubated companies.

6.4 Limitations

While this thesis can estimate the difference-in-differences between the observable covariates in the periods after treatment, there are some limitations that need to be addressed.

One such limitation is that we cannot observe the information that the treated companies, together with their respective incubators have in terms of plans for future growth. These might be the unobservable covariates not present in the accounting numbers, only available to the companies themselves.

Our analysis will also be affected by a selection bias, in terms of which companies were incubated and which companies were not. We do not observe if the control companies applied or chose not to apply for the incubation programs. The control companies could for instance have chosen no to apply for an incubator program, because they did not see the program as worth participating in. Nor do we observe why some companies are chosen to participate in an incubator program and some companies are rejected. The analysis will also be affected by selection bias, when we only choose to include the incubated companies which have start years available from their respective incubator.

As mentioned in section 5.2, in the sample of incubated companies in Oslo, 11,76% of the companies were registered in another region. This might also imply that a fraction of the companies used in the control group could be incubated in another city, resulting in a biased comparison between two incubated companies.

Notably, in the accounting data gathered from SNF there is no information about why a company exit, disappear, or is not visible. There could be many reasons as to why these companies are not available; some of the incubated companies are sole proprietorships and do not report their financials to the Government, while others may not survive the first year or may sell their idea or patent to another company, thus dissolving the incubated one. Reorganisation, bankruptcy, or shutdowns due to buy-outs can also change the organisation number.

Furthermore, the analysis used in the growth sample only examines the incubated companies over a continuous period of three years. This may raise the question of how long it takes before participation in an incubator has an effect. Another question is to ask whether our staggered treatment approach is viable, in terms of the actual length of stay in an incubator versus the natural development of a company as time passes. As some of the incubators provided length of stay and some did not, we cannot conclude with a factual "normal" treatment length.

Another limitation of this thesis is the fact that one company may experience repeated treatments, either by going into one incubator and changing to a different one later on, or having repeated stays in the same incubator. There may also be incubated companies that are not listed on the incubator register, due to poor record-holding or generally little knowledge of who the incubator's inhabitants are. Lastly, other incubators may exist that we are not aware of and are therefore in the available control sample, which could bias our results.

7. Conclusion

In this thesis, we investigate the effect of participating in an incubation program in Oslo between 2011 and 2016. This effect is measured on three main areas: performance, survival rates, and access to public subsidies. Performance is measured on four variables: sales revenues, value creation, operating profit, and number of employees. The effects are measured by constructing a representative control group consisting of companies with similar company characteristics which have not been incubated in Oslo.

We find few significant effects that support our hypothesis of incubated companies performing better than non-incubated companies. The only variable in which incubated companies perform significantly better is in terms of number of employees. Here we find that the average treatment effect on the incubated companies, results in 1.137 more employees. We also find that the incubated companies perform significantly worse in terms of operating profits compared to its unobserved counterfactual outcome. The DiD coefficient for operating profit estimates that the incubated group exhibits decreased operating profit by 263 360 NOK on average in the two-year period after entering the incubator.

In the analysis of survival rates, we found that the incubated group had a higher survival rate in the first, second, and fourth years after matching. In the third year, the control group had a higher survival rate, while the survival rates were equal during the fifth year. The survival rates are calculated as the weighted average of different periods, so the small differences between the survival rates are ambiguous and complicate drawing any clear conclusions.

In terms of public subsidies, we find that the incubated companies submit, on average, seven times more applications for public subsidies. However, we find no evidence of a significant difference between groups when measuring the number of approved applications or the hit rate.

8. References

Antweiler, W. (2001). Nested random effects estimation in unbalanced panel data. *Journal of Econometrics, 101*, 295–313. https://doi.org/10.1016/S0304-4076(00)00086-5

Arellano, M. (1987). Computing Robust Standard Errors for Within-Group Estimators. *Oxford Bulletin of Economics and Statistic 49*(4), 431-434. Retrieved from: https://ideas.repec.org/a/bla/obuest/v49y1987i4p431-34.html

Arkwright X. (n.d.). What we look for. Retrieved from: https://www.arkwrightx.no/what-we-look-for

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data* (5th edition). Chichester, England: John Wiley & Sons.

Baltar, V. T., Sousa, C., & Westphal, M. F. (2014). Mahalanobis' distance and propensity score to construct a controlled matched group in a Brazilian study of health promotion and social determinants. *Revista Brasileira de Epidemiologia 17*(3). 668-679. http://dx.doi.org/10.1590/1809-4503201400030008

Berner, E.,. Mjøs, A. & Olving, M. (2016) Regnskapsboka - Dokumentasjon og kvalitetssikring av SNFs og NHHs database med regnskaps- og foretaksinformasjon for norske selskaper. Retrieved from: https://snf.no/Files/Filer/Publications/A10_16.pdf

Bertrand, M., Duflo, E., & Mullainathan, M. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249-275. https://doi.org/10.1162/003355304772839588

Buigues, P. A., & Sekkat, K. (2011). Public Subsidies to Business: An International Comparison. *Journal of Industry, Competition and Trade*, *11*(1), 1–24. https://doi.org/10.1007/s10842-010-0074-1

Callaway, B., & Sant'Anna, P. H. C. (2019). *Difference-in-Differences with Multiple Time Periods.* Retrieved from: https://ssrn.com/abstract=3148250

Cameron, A. & Miller, D. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources 50*. 317-372. https://doi.org/10.3368/jhr.50.2.317.

Chen, C. J. (2009). Technology commercialization, incubator and venture capital, and new venture performance. *Journal of Business Research*, *62*(1), 93–103. https://doi.org/10.1016/j.jbusres.2008.01.003

Coleman, A. (2018, 05. October). Flourishing Among The Fjords: Norway's Dynamic Startup Scene. *Forbes*. Retrieved from: https://www.forbes.com/sites/alisoncoleman/2018/10/05/flourishing-among-the-fjords-norways-dynamic-startup-scene/?sh=2d1beaac5efb

Colombo M. G. & Delmastro, M. (2002). How effective are technology incubators? Evidence from Italy. *Research on Professional Responsibility and Ethics in Accounting*, *14*, 57–78. https://doi.org/10.1108/S1574-0765(2010)0000014007

Davis, P. (2002). Estimating multi-way error components models with unbalanced panel data structure. Journal of Econometrics, 106(1). 67– 95. https://doi.org/10.1016/S0304-4076(01)00087-2

Dettmann, E., Giebler, A., Weyh, A. (2019). flexpaneldid: A Stata command for causal analysis with varying treatment time and duration, IWH Discussion Papers, No. 5/2019. Retrieved from: http://nbn-resolving.de/urn:nbn:de:gbv:3:2-104857

Digitaliseringsdirektoratet (2020, 20. November). Søknader til og innvilgede prosjekter fra Forskningsrådet, unpublished raw data, Retrieved from: https://data.norge.no/datasets/d23bbbfa-fdad-4dab-b31b-4daadbfa3299

European Commission (2002). *Benchmarking of Business Incubators*. Retrieved from: https://op.europa.eu/en/publication-detail/-/publication/5f01aafc-ef62-457d-9316-c85e7fc2509e

Ferguson, R. & Olofsson, C. (2004). Science Parks and the Development of NTBFs—Location, Survival and Growth. *The Journal of Technology Transfer*, *29*(1), 5–17. https://doi.org/10.1023/b:jott.0000011178.44095.cd

Fjærli, E., Iancu, D., & Raknerud, A. (2018). *Effekten av Sivas virkemidler på vekst og verdiskaping*. (Statistics Norway Reports 2018/17). retrieved from: https://www.ssb.no/virksomheter-foretak-og-regnskap/artikler-og-publikasjoner/effekten-av-sivas-virkemidler-pa-vekst-og-verdiskaping

Gertler, P., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2011). *Impact Evaluation in Practice.* https://doi.org/10.1596/978-0-8213-8541-8

Grimsby, G., Grünfeld, L. A., & Jakobsen, E. W. (2009). *Grunnfjell og vekstmotorer i Norsk næringsliv.* (13/2009). Retrieved from: https://www.menon.no/wp-content/uploads/26menonpubl13200999smb.pdf

Hackett, S. M., & Dilts, D. M. (2004). A Systematic Review of Business Incubation Research. *The Journal of Technology Transfer*, *29*(1), 55–82. https://doi.org/10.1023/b:jott.0000011181.11952.0f

Heckman, J., Ichimura, H., Smith, J. & Todd, P. (1998). Characterizing Selection Bias Using Experimental Data. *Econometrica 66*(5). 1017–1098. https://doi.org/10.2307/2999630

Heckman, J. J., LaLonde, R. J. & Smith, J. A. (1999). The Economics and Econometrics of Active Labor Market Programs. I O. C. Ashenfelter & D. Card (Red.) *Handbook of Labor Economics.* (p. 1865-2097). https://doi.org/10.1016/S1573-4463(99)03012-6.

Holst, L. A. (2019, 08. January). Konkurser: Høyeste konkurstall på 25 år. Retrieved from: https://blogg.bisnode.no/konkurser-hoyeste-konkurstall-pa-25-ar

Høegh-Krohn, J. (2017, 28. April). Kapital i Venturemarkedet. Retrieved from: https://argentum.no/nb/2017/04/28/kapital-i-venturemarkedet/

Iacus, S., King, G., & Porro, G. (n.d.). CEM: Coarsened Excact Matching Software. Retrieved from: https://gking.harvard.edu/cem

Innovation Norway (2020a, 03. February), Our mission, retrieved from: https://www.innovasjonnorge.no/en/start-page/about/our-mission/

Innovation Norway (2020b, 25. May), Startups, retrieved from: https://www.innovasjonnorge.no/en/start-page/our-services/startups/

Innovation Norway (2020c, 04. December), Hva gjør vi?, retrieved from: https://www.innovasjonnorge.no/no/om/hva-gjor-vi/kort-om-oss/

Innovation Norway (2020d, 04. September), Hvem har fått finansiering fra oss?, retrieved from: https://www.innovasjonnorge.no/no/om/hvem-har-fatt-finansiering-fra-oss2/

Innovation Norway (2020e, 11. December), Tilskudd til kommersialisering, retrieved from: https://www.innovasjonnorge.no/no/tjenester/oppstart-av-bedrift/oppstartfinansiering/kommersialiseringstilskudd/

Keuschnigg, C., & Nielsen, S. B. (2003). Tax policy, venture capital, and entrepreneurship. *Journal of Public Economics*, *87*(1), 175–203. https://doi.org/10.1016/S0047-2727(01)00170-0

Lukeš, M., Longo, M. C., & Zouhar, J. (2018). Do business incubators really enhance entrepreneurial growth? Evidence from a large sample of innovative Italian start-ups. *Technovation*, *82–83*(July 2018), 25–34. https://doi.org/10.1016/j.technovation.2018.07.008

McLachlan, G. J. (1999). Mahalanobis Distance. *Reson 4*, 20–26. https://doi.org/10.1007/BF02834632

Ministry of Trade, Industry and Fisheries. (2015). Gode ideer - fremtidens arbeidsplasser. Retrieved from: https://www.regjeringen.no/contentassets/05f1305cb2a94a379ff48c2f2c60d688/grunderplan_2015.pdf

Nilssen, S. S. (2020, 9. December). Tror corona-konkursbølge snart er på vei. *Finansavisen.* Retrieved from: https://finansavisen.no/nyheter/naeringsliv/2020/12/09/7596450/konkursekspert-tror-corona-konkursbolge-snart-er-pa-vei

Petersen, M. L., Porter, K. E., Gruber, S., Wang, Y., & van der Laan, M. J. (2012). Diagnosing and responding to violations in the positivity assumption. Statistical Methods in Medical Research, 21(1), 31–54. https://doi.org/10.1177/0962280210386207

Rippolone, J. E., Huybrechts, K. F., Rothman, K. J., Ferguson, R. E. & Franklin, J. M. (2019). Evaluating the Utility of Coarsened Exact Matching for Pharmacoepidemiology Using Real and Simulated Claims Data. *American Journal of Epidemiology 189*(6). 613-622. https://doi.org/10.1093/aje/kwz268

Schwartz, M. (2013). A control group study of incubators' impact to promote firm survival. *Journal of Technology Transfer*, *38*(3), 302–331. https://doi.org/10.1007/s10961-012-9254-y

Siva (n.d.a), Om Siva. Retrieved from: https://siva.no/om-siva/

Siva (n.d.b), Siva i tall. Retrieved from: https://siva.no/siva-i-tall/

Siva (n.d.c), Siva-strukturen. Retrieved from: https://siva.no/om-siva/siva-strukturen/?kategori=Inkubator#soket

StartupLab (n.d.). Accelerating top tier startups. Retrieved from: https://startuplab.no/#lab

Statistics Norway (2020a, 30. September). Nyetablerte foretaks overlevelse og vekst. Retrieved from: https://www.ssb.no/fordem

Statistics Norway (2020b, 5. February). Gjennomsnittlig månedslønn for kvinner og menn i ulike sektorer. Retrieved from https://www.ssb.no/arbeid-og-lonn/statistikker/lonnansatt

Stubberud, H. A. (2016). *Business Incubators and Entrepreneurial Performance: The Influence of Network Value and Absorptive Capacity.* (Doctoral dissertation, Norwegian School of Economics). Retrieved from: https://openaccess.nhh.no/nhh-xmlui/handle/11250/2412398

Tandsæther-Andersen, B. (2017, 20.12). Kartet som viser Oslos inkubator-boom: Noen vil vinne, andre forsvinne, men hva passer for gründeren? *Shifter*. Retrieved from: https://shifter.no/startup/kartet-som-viser-oslos-inkubator-boom-noen-vil-vinne-andre-forsvinne-men-hva-passer-for-grnderen/106355

The Research Council of Norway (n.d.), Tasks and organisation, retrieved from: https://www.forskningsradet.no/en/about-the-research-council/Tasks-and-organisation/

The Research Council of Norway (2019a, 10. April), What does the Research Council do?, retrieved from: https://www.forskningsradet.no/en/about-the-research-council/Tasks-and-organisation/what-does-the-research-council-do/

The Research Council of Norway (2019b, 05. June), Få skattefradrag via skatteFUNN, retrieved from: https://www.forskningsradet.no/sok-om-finansiering/hvem-kan-soke-om-finansiering/naringsliv/ordninger-for-okonomisk-stotte/skattefradrag/

The Research Council of Norway (2019c, 06. March), Innovasjonsprosjekt i næringslivet, retrieved from: https://www.forskningsradet.no/sok-om-finansiering/hvem-kan-soke-om-finansiering/naringsliv/innovasjonsprosjekt-naringsliv/

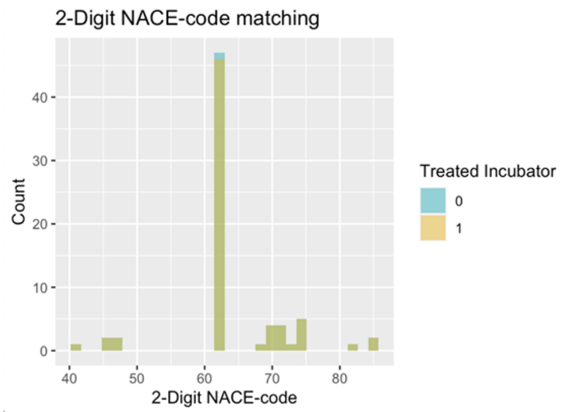
Tobiassen, M. (2015, 21. November). Godt nytt for oss, skikkelig dårlig nytt for as Norge. *Dagens Næringsliv.* Retrieved from: https://www.dn.no/marked/fond/grunder/-godt-nytt-for-oss-skikkelig-darlig-nytt-for-as-norge/1-1-5514088

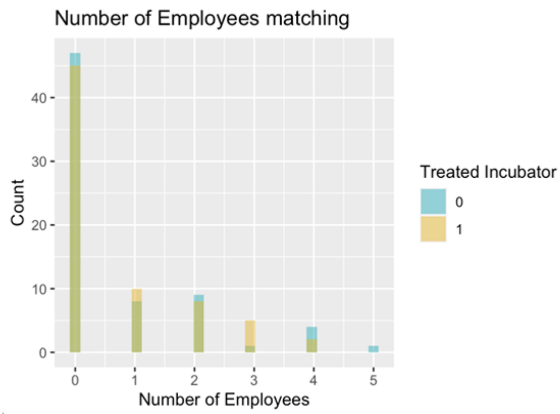
Wennekers, S., & Thurik, R. (1999). Linking Entrepreneurship and Economic Growth. *Small Business Economics*, *13*(1), 27–55.

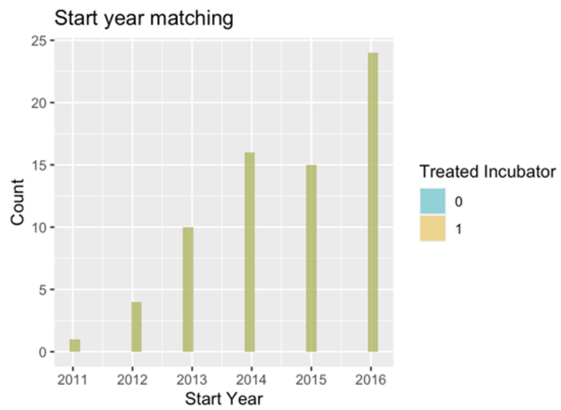
Øyvann, S. (2017, 31. May). Forhåndssalg for 100 millioner kroner. *Computerworld.* Retrieved from: https://www.cw.no/artikkel/innovasjon/forhandssalg-100-millioner-kroner

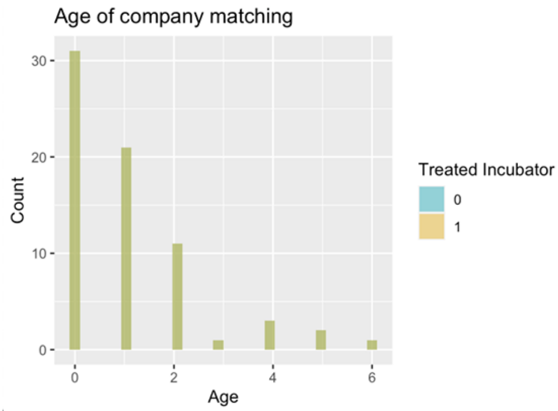
Appendix

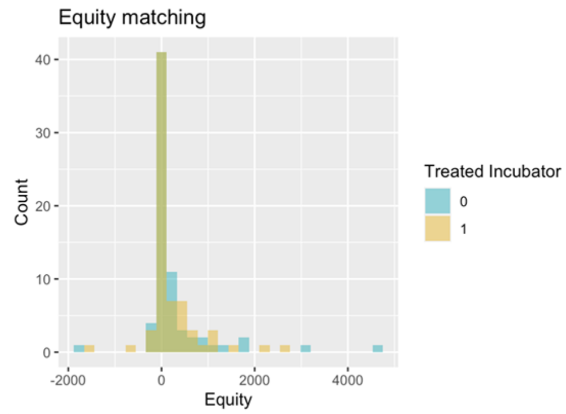
Appendix A

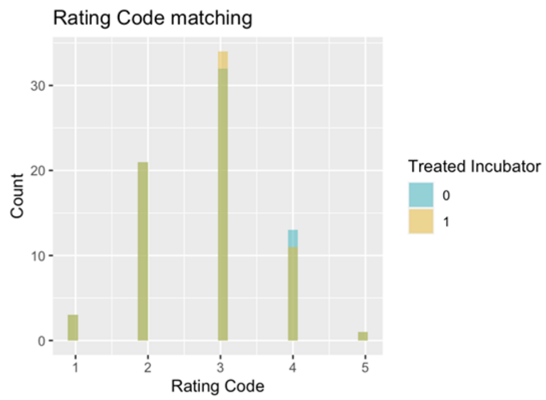


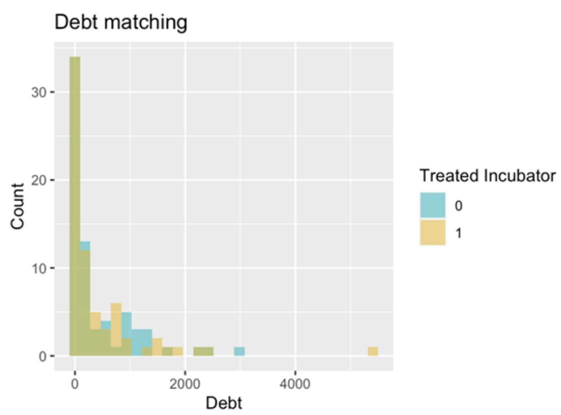


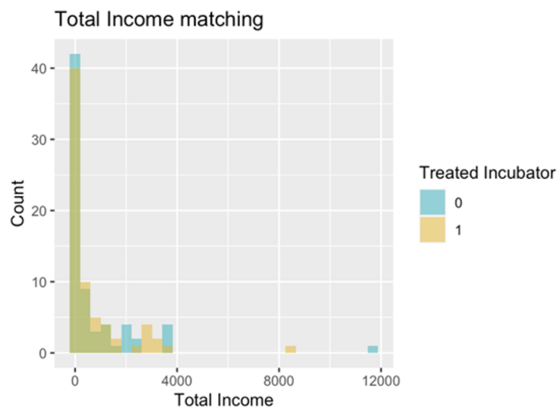




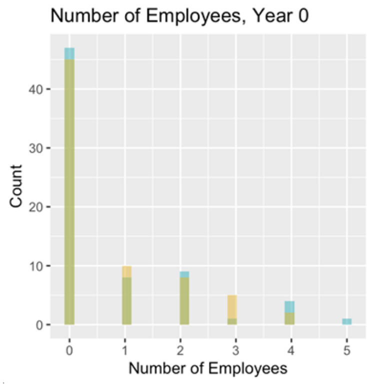


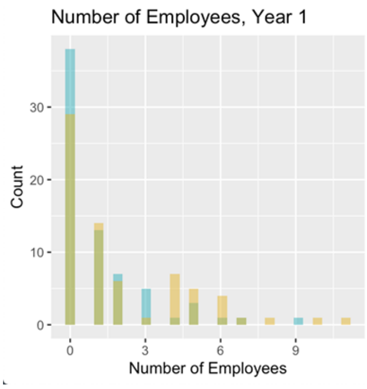


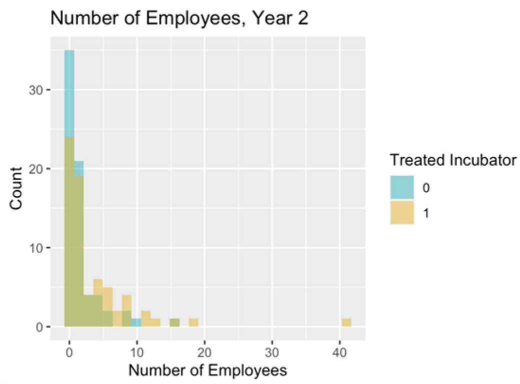




Appendix B


Number of Employees, Year 0
Count
0
10
20
30
40
0
1
2
3
4
5
Number of Employees


Number of Employees, Year 1
Count
0
10
20
30
0
3
6
9
Number of Employees


Number of Employees, Year 2
Count
0
10
20
30
0
10
20
30
40
Number of Employees
Treated Incubator
0
1