Is new always better?

The effect of technological diversity change on the performance of climate-focused start-ups

Master thesis | Master Innovation Sciences
Copernicus Institute of Sustainable Development
Department of Innovation, Environmental and Energy Sciences
Faculty of Geosciences | Utrecht University
April 4th 2018

Author:
Jip Leendertse BSc
Utrecht University
j.leendertse@uu.nl
Student number: 3848043

Supervisors:
dr. F.J. van Rijnsoever
C.P. Eveleens MSc

Second reader:
dr. M.M.H. Chappin

Universiteit Utrecht

Climate-KIC

Abstract

Climate-focused start-ups often aim to both exploit a market opportunity and to reduce the impact of climate change. The performance of these start-ups thus consists of a business as well as a climate dimension. This study is the first to include both performance dimensions, as the existing literature has so far only considered business performance. The aim of this study is to provide insight into how the technological position of a climate-focused start-up influences it's business and climate performance. The *technological position* is measured as the degree of technological diversity change in the system caused by the introduction of a start-up's technology. This study performs quantitative analyses using a sample of 197 start-ups which participate in Europe's largest climate innovation accelerator, the Climate-KIC accelerator. Regarding *business performance* the results did not support the hypothesis that start-ups which open up new technological trajectories, and create technological diversity, have significantly lower business performance. Regarding *climate performance*, this study finds that technological diversity creation has a positive influence on a start-up's potential to reduce CO_2e emissions. These findings confirm the expectation that start-ups which create diversity, have a higher technological potential and subsequently a higher potential climate impact. The results of this study thus reveal an interesting dynamic between the three societal functions of climate-focused start-ups – (1) climate mitigation, (2) economic development, and (3) stimulating technological change by introducing technological diversity. Start-ups that create technological diversity have a higher potential to reduce CO_2e emissions, while there is no significant influence on business performance. Furthermore, start-ups with a software technology are found to be more likely to have higher business performance, but they also have less potential to reduce CO_2e emissions. This research thus provides support for the notion that the business and climate dimensions of performance are fundamentally different from each other. The goals of climate and business performance do not align, and while business performance is necessary to translate potential climate impact into *realized* climate performance, it is not sufficient. Policy makers, incubators and accelerators therefore face a challenge in balancing these two performance dimensions.

Acknowledgements

First, I want to thank my supervisors Frank van Rijnsoever and Chris Eveleens for providing me with their critical feedback and for challenging me to tell more in less words. Second, my sincere appreciation goes to Climate-KIC and the ECCI for granting me access to their program and the data used in this research. In this my special thanks go to Erik Faassen for all his efforts, which eventually made it possible to overcome the many challenges of the data collection process and to Cassi Welling for hosting me on my research visit to Berlin. Furthermore, I want to thank Ellen, Florian, Inge, John, Maria, Peter, Sam, and Tijs for their feedback. Finally, I want to thank my parents and girlfriend for supporting me through the ups and downs of the whole process which makes me proud to present this finished thesis!

Table of content

1. Introduction

To combat climate change both researchers and policy makers emphasize the importance of transitioning to a more environmentally sustainable economy (Chouinard et al., 2011; Dietz & O'Neill, 2013; Office of the Press Secretary, 2015). This transition will strongly be driven by technological change (Bjornali & Ellingsen, 2014; Gibbs, 2006). Technological change usually takes place along ordered and selective patterns, which are called technological trajectories. These trajectories consist of incremental improvements which are the result of increased experience with a technology (Dosi & Nelson, 2009; Verspagen, 2007). Novel technologies, which do not fit in existing trajectories, can open up new opportunities for improvement and cause new trajectories to emerge (Zhong & Verspagen, 2016). These novel technologies create *technological diversity*, they increase the amount and variety of technological components in the system (Stirling, 2007).

Start-ups play a key role in technological change by accessing new technological knowledge and turning this into commercial applications (Acs & Audretsch, 2005; Ács & Varga, 2005; Degroof & Roberts, 2004; European Commission, 1998; Spencer & Kirchhoff, 2006). As a result, start-ups are found to create more technological diversity than other firms (Almeida & Kogut, 1997; Stirling, 2007). The creation of technological diversity is beneficial to society because of three reasons. First, it helps to prevent lock-in on a sub-optimal technology. Second, it increases the flexibility to react to a changing environment (Van den Bergh, 2008). Finally, it also increases the chance of connecting technologies to create recombinant innovations (Van den Bergh, 2008; Van Rijnsoever et al., 2015). The start-ups' role in creating technological diversity is of particular importance for 'climate-focused' start-ups, because a higher diversity of technologies is expected to be necessary for the transition to a more environmentally sustainable economy (Bjornali & Ellingsen, 2014; Gerlach, 2003; Gibbs, 2006; Stirling, 2010).

Besides this societal contribution, the technological diversity change (both the creation and reduction of diversity are possible) caused by climate-focused start-ups is also expected to be of crucial importance for the start-up itself. This is the case because a start-up's technology, and the resulting technological position, is considered to strongly influence the start-up's competitive advantage, and its subsequent business performance (Aharonson & Schilling, 2016; Debackere et al., 1999; Zahra, 1996).

However, the relation between *the degree of technological diversity change caused by a climate-focused start-up* and it's *performance* has not yet been tested in practice. The existing literature on the technological trajectories of low carbon technologies has so far mainly focused on the development of technologies, such as hydrogen and solar power, at the macro-level (Anandarajah & McDowall, 2015; Nemet, 2006). The few studies which are conducted at the micro-level, look only at the influence of patent novelty on the economic performance of these patents (Fontana et al., 2009; Harhoff et al., 1999; Verhoeven et al., 2016; Verspagen, 2007). However, patents are not a reliable indicator for start-up's, because they often do not file for them due to the large costs associated with patenting (Graham & Sichelham, 2008; Helmers & Rogers, 2011). The existing technological trajectory literature therefore does not provide insights into how the level of technological diversity change caused by a climate-focused start-up influences the start-up's performance. In this study I aim to fill this research gap.

Soetanto & Jack (2016) do study a related phenomenon by researching the effect of a start-up's innovation strategy, defined as either technology exploitation or technology exploration, on its performance. However, they only look at the business dimension of performance (Soetanto & Jack, 2016). They fail to include the environmental impact of the start-ups, even though this is also a critical performance measure from a societal perspective (Bjornali & Ellingsen, 2014; Calel & Dechezlepretre, 2013). The climate impact should therefore be included as a dimension of firm performance (Elkington, 1994; Meyskens & Carsrud, 2013; Slaper & Hall, 2011). However, no research has yet included the climate dimension when studying start-ups (Bjornali & Ellingsen, 2014; Dean & McMullen, 2007; Meyskens & Carsrud, 2013).

In this study I aim to fill both the aforementioned research gaps by addressing the following research question:

What is the influence of the degree of technological diversity change caused by climate-focused start-ups' on their business and climate performance?

I quantitatively test the influence of technological diversity change on performance using start-ups who participate in the Climate-KIC accelerator program in the Netherlands and the DACH and Nordics regions[1]. The Climate-KIC accelerator program is a EU-funded program which provides support services to climate-focused start-ups in fourteen European countries.

This study has two key theoretical contributions. First, I apply arguments from the technological trajectory literature to the literature on start-up business performance combining these two literature strands. Second, by including the climate dimension of start-up performance this study takes a first step towards a more holistic evaluation of performance, which includes their societal contributions as well as their business performance. From a managerial perspective, this study helps managers of incubators and accelerators to better understand the influence of start-ups' technologies on the performance of the start-ups. As such, these managers can make better decisions in selecting start-ups for their program. For policy makers, the use of climate impact as a performance dimension helps them to better evaluate the societal contribution of these climate-focused start-ups.

This thesis continues with a theory section, which discusses the literature to derive hypotheses about the influence of technological diversity change on the dependent variables. The third section describes the methodology which is used to perform the empirical research as well as an explanation on how the different variables are constructed. This is followed by a chapter which contains the results. Finally, the discussion contains a reflection on the research itself, as well as its outcomes.

[1] DACH*: Germany, Austria, and Switzerland.*
Nordics: *Denmark, Norway, Sweden and Finland*

2. Theory

The literature on technological trajectories is one of the dominant theories in the innovation sciences (Dosi & Nelson, 2009, 2013). The theory is built on the notion that a process of variation and selection causes technological development to occur along trajectories. Technologies which fit in existing trajectories perform better because they benefit from the accumulated experience in the trajectory (Anandarajah & McDowall, 2015; Ibenholt, 2002; Nemet, 2006). On the other hand, technologies that open up new trajectories face a larger risk that they will not become technologically and economically successful (Marra et al., 2003; Zhong & Verspagen, 2016). So far the technological trajectory literature is mainly used to study particular technologies, such as hydrogen and solar power, at the macro-level (Rogner, 1998; Witajewski-Baltvilks et al., 2015; Yu et al., 2011). In this thesis these theoretical concepts are used to derive hypotheses about start-ups at the micro-level.

2.1. Start-ups and their performance

Start-ups are usually defined as small and young entrepreneurial ventures which are in the process of exploring a technology to develop their business (Bjornali & Ellingsen, 2014; Fontes & Coombs, 2001; Klotz et al., 2013). This study focusses on one particular sub-group of start-ups, *climate-focused start-ups*. These start-ups develop and commercialize technological knowledge that is beneficial for the environment (Bjornali & Ellingsen, 2014; Meyskens & Carsrud, 2013). Based on Bjornali & Ellingsen's (2014) definition of clean tech start-ups, I define a climate-focused start-up as: '*an entrepreneurial venture which significantly reduces greenhouse gas emissions by exploiting technological knowledge*'.

The performance of a start-up is defined as whether the start-up achieves its desired purpose (Wright & Stigliani, 2012). Entrepreneurs have multiple purposes and start-up performance is therefore best conceptualized as a multidimensional construct (Eveleens et al., 2016; Zahra, 1996). The aim of climate-focused start-ups is often to both exploit a market opportunity and to aid in developing a new technology which can help reduce the impact of climate change (Bjornali & Ellingsen, 2014; Parrish, 2010). Therefore, the business and climate performance constitute two different dimensions of the performance of climate-focused start-ups (Bennett, 1991).

2.1.1. Business performance

Start-up business performance has been conceptualized through a wide range of different concepts (Eveleens et al., 2016). However, these different concepts hardly correlate and all come with certain limitations (Eveleens et al., 2016; Murphy et al., 1996; Witt, 2004). Measuring and conceptualising the business performance of start-ups is thus not a straightforward process and the choice of performance measure considerably influences the results of a research (Eveleens et al., 2016; Song et al., 2008; Witt, 2004). Therefore, a combination of performance measures should be used to conceptualise start-up business performance (Eveleens et al., 2016; Song et al., 2008).

Looking at the network incubation literature, Eveleens et al. (2016) found that *firm size* and *investments* are among the most frequently used dimensions of business performance. This paper therefore uses these dimensions, which represent different aspect of business performance, to conceptualize business performance.

The first dimension, *firm size,* can be conceptualized along two aspects. First, the contribution of a start-up to the economy depends on its performance on the market. This is indicated by its *revenues*, a performance measure which is more relevant in the later stages of a start-up, as very early stage start-ups often have not yet recorded any sales (Groenewegen & De Langen, 2012; Rothaermel & Thursby, 2005; Sullivan & Marvel, 2011). Second, the *number of employees* as a dimension of firm size represents the contribution of a start-up to the labour market (Groenewegen & De Langen, 2012; Sullivan & Marvel, 2011).

As the climate-focused sector is very capital intensive, start-ups need external funding (Bjornali & Ellingsen, 2014). Obtaining funding is therefore often a main goal in the early phase of start-ups (Bjornali & Ellingsen, 2014; Eyraud et al., 2013). The obtainment of such external funding also represents the expectations of investors for the start-up's long term performance, and as such bestows legitimacy upon a start-up (Rothaermel & Thursby, 2005). Thus, the ability of a start-up to obtain *investments* represents the second dimension of its business performance (Rothaermel & Thursby, 2005).

2.1.2. Climate performance

Because no research has been performed on the climate performance of start-ups, I turn to the literature on large firms to define the concept (Bjornali & Ellingsen, 2014; Meyskens & Carsrud, 2013). Here, the existing scholarly efforts have researched the climate impact of environmental initiatives by looking at their reduction in CO_2 equivalent (CO_2e) emissions (Cohen & Winn, 2007; Gohar & Shine, 2007; Meyskens & Carsrud, 2013). Therefore, I define climate performance as the reduction in CO_2e emissions caused by a start-up's technology in comparison with the conventional alternative (Bjornali & Ellingsen, 2014; Rasmussen et al., 2012).

Due to the small size of start-ups in the first years of their business, their emission reductions will inherently also be small during these years (Hyytinen et al., 2015). As such, it is more relevant to look at the potential of their technology to reduce carbon emissions (Bjornali & Ellingsen, 2014). Therefore, the *potential CO_2e emission reductions* (long term) are considered in this research.

2.2. Technological diversity change

Diversity can be defined as *"the evenness in a distribution of elements among a number of categories in a system"* (Van Rijnsoever et al., 2015, p. 1096). According to Stirling (2007) diversity consists of the three elements variety, balance and disparity. *Variety* is "the number of categories into which system elements are placed", *balance* is the number of elements in each category, and *disparity* refers to how different the elements are (Stirling, 2007, p. 709). A system with less similarity between the components of technologies has more diversity than a system with greater similarity (Aharonson & Schilling, 2016). The technological diversity of a system changes with the introduction of a new technology (Páez-Avilés et al., 2016).

In this study, technological diversity change is a start-up characteristic (micro-level construct) which resembles how the technological diversity of the system (macro-level construct) changes due to the introduction of that particular start-up's technology (Murmann & Frenken, 2006; Páez-Avilés et al., 2016; Van Rijnsoever et al., 2015). To analyse the technologies of climate-focused start-ups, the *technological system* consists of technologies that contribute to climate change mitigation (USPTO, 2018). In the next two sections the mechanisms, through which the degree of technological diversity change is expected to influence the business and climate performance of start-ups, are explained.

2.2.1. Business performance

Technologies that open up new trajectories, and create technological diversity, are associated with higher risks. These higher risks often reduce their adoption rates and business performance (Fleming, 2001; Marra et al., 2003; Verhoeven et al., 2016). Similarly, Hyytinen et al. (2015) find, in an empirical study, that more innovative start-ups are likely to encounter a greater liability of novelty, which makes them less likely to achieve high business performance (Hyytinen et al., 2015).

Technologies that reduce diversity, on the other hand, build closely on technologies in existing technological trajectories. As such, they can be expected to benefit from economies of scale and learning effects obtained through experience with these other technologies (Yu et al., 2011; Zhong & Verspagen, 2016). This makes these technologies more competitive on the market (Anandarajah & McDowall, 2015; Rogner, 1998). For start-ups, Soetanto & Jack (2016) find that a strategy of *optimizing existing technologies* is more likely to lead to successful businesses than a strategy of *discovering new knowledge,* and thus creating technological diversity. Malerba (2009) associates the strategy of *optimizing existing technologies* with reducing diversity and this leads to the following hypothesis:

Hypothesis 1: Start-ups that increase technological diversity are expected to have a lower business performance

2.2.2. Climate performance

A start-up's climate performance is determined by the ability of its technology to reduce CO_2e emissions (Zhang et al., 2013). The start-up's potential to reduce CO_2e emissions is thus dependent on the technological potential of the start-up's technology.

Technologies that create technological diversity are still at the beginning of technological trajectories (Verhoeven et al., 2016; Zhong & Verspagen, 2016). These technologies are more likely to be breakthrough innovations with a high technological potential (Aharonson & Schilling, 2016; Fleming, 2001). This is the case because, for these technologies, learning effects have not yet occurred and they therefore have more potential to improve when they mature (Nemet, 2006; Rogner, 1998; Yu et al., 2011). Technologies that create technological diversity are therefore expected to have a larger potential to reduce CO_2e emissions (Aghion et al., 2012; Aghion et al., 2014; Bjornali & Ellingsen, 2014; Nemet, 2009). As such, I arrive at the following hypothesis:

Hypothesis 2: Start-ups that increase technological diversity are expected to have a higher potential climate performance

3. Methods

3.1. Research design and data collection

In this study I use quantitative analyses to test the aforementioned hypotheses. This is done through a cross-sectional research design where the independent and control variables are established prior to the dependent variables. This time difference between the variables makes it possible to derive causal inferences, which matches the aim of this study (Bryman, 2012).

The research sample consists of start-ups who are or were part of the Climate-KIC accelerator program in the period 2012-2016. The Climate-KIC accelerator is the largest climate innovation accelerator in Europe (Climate-KIC, 2018). The start-ups in this program are especially suited for this research because the program only selects young entrepreneurial ventures with a positive climate impact, which meets my definition of climate-focused start-ups (Climate-KIC, 2017). Furthermore, the accelerator program is highly similar across the different countries in which it is run (Climate-KIC, 2017). This means that the selected start-ups receive similar resources in the form of network opportunities, finance, and training and as such can be compared on a more even basis.

The data is collected for start-ups who participated in the three largest Climate-KIC accelerator programs, the Netherlands, the DACH region (Germany, Austria, Switzerland) and the Nordics (Denmark, Norway, Sweden, Finland). The chosen sample also means that the start-ups are active in multiple (North-Western) European countries, which makes it possible to generalize to the population of European climate-focused start-ups. The data on these start-ups is collected from a combination of three types of sources: (1) The Climate-KIC evaluation surveys are used as the data source for the business performance variables. (2) The application forms to the accelerator are text-mined to collect the information for the independent and climate dependent variables as well as for the control variables. (3) A combination of public sources (such as the Chamber of Commerce) are used to fill in missing information for the control variables.

I managed to collect application forms on a total of 870 start-ups who applied to the accelerator program. 303 of these start-ups were accepted into the program. From this group, there is performance data for 197 of the start-ups. So the response rate for the performance survey was 65%. These 197 start-ups constitute the research sample for which the complete set of variables is collected.

3.2. Operationalisation

3.2.1. Dependent variable: Business performance

The data on the *business performance variables* originates from the evaluation survey which Climate-KIC conducted about the performance of the start-ups in 2014, 2015 and 2016. Business performance is operationalised along two dimensions: firm size (*revenues* and *number of employees*) and *investments*. Using the combination of the two dimensions to operationalise business performance enhances the validity of this dependent variable (Eveleens et al., 2016; Song et al., 2008; Zahra, 1996).

The first aspect of *firm size* concerns the *number of employees* who are employed by the start-up (Eveleens et al., 2016; Groenewegen & De Langen, 2012; Peña, 2004). This is operationalized as a count of the number of people who worked for the start-up at the time of the performance survey. The second dimension of *firm size* are the *revenues*, which is the absolute, cumulative amount of turnover created by the company in the year of the performance survey (Groenewegen & De Langen, 2012; Rothaermel & Thursby, 2005; Sullivan & Marvel, 2011). The data from the performance survey is of an ordinal nature and the revenues are therefore operationalised on a four-level ordinal scale, the exact levels of which are shown in the operationalisation table in Appendix A.

The *investments* are operationalised as the absolute, cumulative amount of external investments made into the company between the start-ups foundation and the moment of the performance survey (Rothaermel & Thursby, 2005). This variable is also measured on a four-level ordinal scale due to the nature of the available data (for levels see Appendix A).

3.2.2. Dependent variable: Climate performance

As part of their application to the Climate-KIC accelerator the start-ups are asked to provide a description of how their business will contribute to a more environmentally sustainable economy and help reduce the emission of greenhouse gases. In this study, these descriptions are used to assess the *potential* climate impact of these start-ups. This assessment takes place in the form of an expert coding by the author (Hallgren, 2012).

To assess the start-up's *potential to reduce CO_2e emission* I analysed the sections on the application forms, which described the *climate potential* as well as the *business idea*. Based on these qualitative descriptions I reviewed each start-up's potential to reduce CO_2e emission if their business idea becomes successful. I then scored this potential on a 5-point scale, in which a one stood for a very low potential and a five for a very high potential.

As such, a subjective measure for climate performance is used. Subjective performance measures are used more often in research on start-up performance, but are at risk of psychological biases, which influences their reliability (Bryman, 2012; Eveleens et al., 2016; Richard et al., 2009). To increase the reliability of the measure I therefore verify the author assessments with those of a group of experts. As part of the application process for Climate-KIC, a panel of industry experts rates the *potential climate impact* of start-ups on a five-point scale[2]. The expert scores themselves could not be used as the climate performance measure, because panel reviews were only available for 127 out of the 197 start-ups. However, the inter-rater-reliability (IRR) between the expert and author scores can be used to verify the reliability of the author scores (Hallgren, 2012). The IRR provides a way of quantifying the degree of agreement between two or more coders who make independent ratings about the features of a set of subjects.

The IRR is calculated through one-way, single-measure Inter Class Correlations (ICC). This is the appropriate method to generalize the results from a subset (127 start-ups) to the full dataset (197) when the ratings are on an ordinal scale (Hallgren, 2012). The ICC is calculated using the *irr* package in 'R' software (Gamer et al., 2012; R Core Team, 2017).

[2] In the Nordics a 1-6 scale was used. The results were robust when either adapting this scale to 1-5 by coding the 6 as a 5 or by comparing with the original 1-6 scores.

When looking at the scores of the individual panel members two things stand out. First, there are a number of panel members who only provided climate impact evaluations for a single start-up. And second, there is a group of panel members whose assessments show very poor correlations (below 0.2) with the *mean* score of the panel members. When these two groups of panel members are removed from consideration, data on about 50 panel members who have scored 122 start-ups for a total of 509 pairs of unique panel member/start-up climate assessments remains. Because the panel member data is relatively sparse, a single score for the panel member group is constructed by calculating the mean. The ICC values between the author scores and panel member mean is 0.627 showing a good IRR and thus proving that the climate impact assessment of the author is a reliable measure (Cicchetti, 1994; Hallgren, 2012). More detailed information on the construction and robustness of this variable can be found in Appendix B.

3.2.3. Independent variable: Technological diversity change

The technological diversity change caused by a start-up is measured as the difference in diversity in the technological system with or without the start-up's technology in the system (Páez-Avilés et al., 2016). As such, measuring the diversity change caused by a start-up requires mapping the technological system and determining the position of each start-up within this system. The technological system in this study consists of technological innovations, commercialized technologies, rather than technological inventions. Previous studies have shown that text-mining is a particularly well suited approach to map technological systems because it can be used to accurately assess a technology's complex features and identify patterns between different technologies (Aharonson & Schilling, 2016; Arts et al., 2013; Blei, 2011; Páez-Avilés et al., 2016).

Usually, patent data is used as the input for text-mining models in technology studies (Aharonson & Schilling, 2016; Kaplan & Vakili, 2013). However, patent data is only suited to assess firms for which patents are a reliable indicator of their technological capabilities, which is not the case for start-ups (Aharonson & Schilling, 2016; Graham & Sichelham, 2008; Helmers & Rogers, 2011) I instead use the start-up's application forms to the Climate-KIC accelerator as the input to the text-mining models. These application forms are well suited for this purpose because they contain a section which consists of a description of the start-up's technology. Furthermore, the descriptions are very similar in length to the abstracts which are often used as the input in text-mining models (Grün & Hornik, 2011; Páez-Avilés et al., 2016; Zhao et al., 2015). The application forms thus provide a unique database of detailed technological descriptions for start-ups.

The technological descriptions of all 303 start-ups accepted into the Climate-KIC accelerator program in the three regions are used to form the technological field. The choice to use these start-ups as the delimitation of the technological field is made because of two reasons. First, start-ups usually focus on developing a single technology and as such the data is not 'polluted' through the combination of multiple technologies for one company (Hyytinen et al., 2015). Secondly, text-mining models are particularly well suited to compare and analyse different documents of the same type (Grün & Hornik, 2011; Hotho et al., 2005). This makes it practically unsuitable to include other firms and descriptions into the text-mining models. Including a broader range of companies by using website texts instead of the application forms proved to be unfeasible due to the limited availability of technological information on these websites. Defining the technological system exclusively through start-up technologies is nevertheless a limitation of this study and the results should therefore be interpreted with the knowledge that the technological field consisted only of start-ups.

To operationalize diversity I follow the approach of other studies in the field who use the Shannon-Weaver entropy index, which contains *variety* and *balance* (Páez-Avilés et al., 2016; Shannon, 1948; Van Rijnsoever et al., 2015). This entropy index is chosen instead of the Stirling diversity because the third dimension of diversity change, *disparity*, depends on subjective interpretations and is therefore very difficult to conceptualize in technology studies (Van Rijnsoever et al., 2015). The entropy value is then used to measure technological diversity (Huang & Chen, 2010):

$$H = -\sum_{i=1}^{R} p_i \, ln p_i \tag{1}$$

Here, H is the entropy value for diversity and p is the proportion of start-ups with a specific topic (i) (Páez-Avilés et al., 2016). The diversity change in the system caused by a start-up (ΔH) can be calculated through the difference between the entropy of the population of clean-tech start-ups (H_1) and a hypothetical population in which that particular start-up does not exist (H_0).

$$\Delta H = H_1 - H_0 \tag{2}$$

To calculate the entropy values, the latent Dirichlet allocation probabilistic topic model (LDA) is used. This is a text-mining approach which analyses the words of documents to discover the themes that run through the documents and the connections between these themes (Blei, 2011). The themes which are identified through the LDA are represented through latent topics, each of which is a set of probability distributions over the words, while each document is a set of probability distribution over the topics (Lee et al., 2012; Steyvers & Griffiths, 2007). The latent topics form the *categories* (for variety) and the start-up technologies the *elements* (for balance) which make up the entropy formula (Eq. 1).

There are a number of data transformation steps necessary to perform the LDA, these are described in Appendix C (Feinerer, 2017; Meyer et al., 2008). Similar to other studies I use the VEM algorithm to perform the LDA (Grün & Hornik, 2011). When performing an LDA the Dirichlet parameter (α) and the number of topics (k) need to be set a priori (Blei & Lafferty, 2009; Grün & Hornik, 2011). With the VEM algorithm α can be estimated or set to a fixed value of 50/k (Grün & Hornik, 2011). The choice between an estimated and fixed value for α is made by comparing the perplexities (Grün & Hornik, 2011; Su & Liao, 2013; Teh et al., 2005). The perplexity is a measure that represents how well a probability model is able to predict a sample. This is done through 10-fold cross validation. The data is split in 10-folds, which are each held-out from model calculations once, while the other nine folds are used to train the model. The held-out set is then used to test how well the trained model predicts the word distributions in the documents of the test-fold (Grün & Hornik, 2011). The perplexities of the models with the *fixed α* are consistently lower than the models with the *estimated α* and as such the fixed α approach is used (see Appendix C).

In order to determine the appropriate amount of topics, I estimate multiple models. The most appropriate model is the first number of topics whose rate of perplexity change (RPC) is smaller than the following number of topics (Zhao et al., 2015). This is the case for the model with 15 topics (see Appendix C). The 10 most frequent words for five of these topics are shown in table 1 while the complete topic overview is depicted in Appendix C.

Table 1: The ten most frequent terms for five of the topics resulting from the technologies of 303 climate-focused start-ups.

Topic 2	Topic 4	Topic 5	Topic 6	Topic 12
food	charg	pile	clean	heat
farmer	vehicl	foundat	water	wast
crop	grid	soil	ship	wood
chemic	station	tip	fuel	dri
water	box	crane	contain	fuel
fertil	car	partnership	robot	hous
agricultur	street	vibrat	captur	glass
soil	meter	introduc	surfac	modheat
farm	raft	invent	cleaner	forest
yield	driver	lift	load	combust

The results of the LDA are these 15 topics (clusters of words) and for each document a distribution over the topics. The technologies that exhibit similar distribution over the topics are then clustered together (Aggarwal & Zhai, 2013). Here, each document represents a start-up's technology and the network of clustered documents represents the technological system. The topic distribution and probabilities are then used to calculate the diversity change caused by each start-up through the aforementioned Shannon-Weaver entropy index (Eq. 2). The histogram of technological diversity change (figure 1) displays a positive kurtosis and skewness but is otherwise relatively normally distributed. However, as the diversity change values are very small they are multiplied by 1000 in order to construct more normalized coefficients.

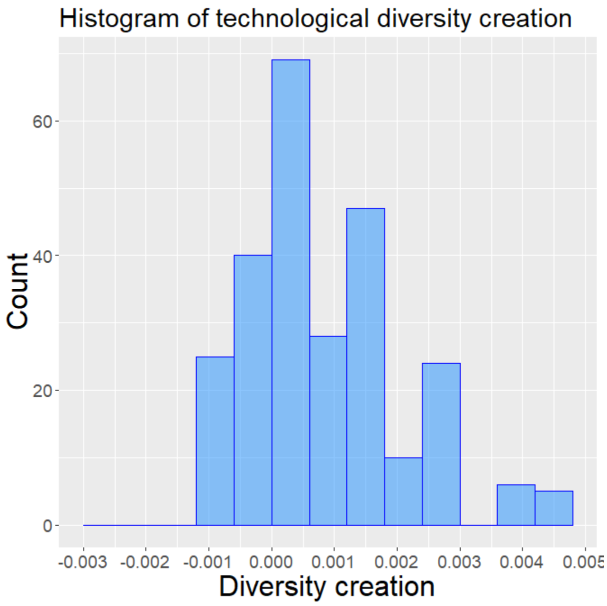


Figure 1: Histogram of the technological diversity change values of the start-ups.

To test the reliability of the construction of this variables I performed two robustness checks. First, an alternative algorithm, in the form of Gibbs sampling, is used to run the LDA and calculate the diversity values (Blei, 2011; Srivastava & Shami, 2009). Second, the technological diversity change variable is also constructed based on the 197 start-ups included in this research. This was done to check for the influence of the size of the technological field on the results. In both cases the results of the regression models proved to be robust (see Appendix E).

3.2.4. Control variables

I use a number of control variables which were identified to influence start-up performance in the existing literature.

The start-ups in this research vary in their foundation years, the year they entered the accelerator and the moment at which they filled in the evaluation survey. In order to account for the varying time dimension in this study I control for the *age of the start-ups*, calculated as the number of years between foundation and the time of the performance survey (Ortín-Ángel & Vendrell-Herrero, 2014). *Start-up age* is also included because previous research finds a positive significant relation between the age of a start-up and its performance (Soetanto & Jack, 2016, 2013; Song et al., 2008).

Second, Soetanto & Jack (2013) also find a positive significant relation between spin-off size and performance, and bigger teams are better able to mobilize resources than smaller teams (Klepper, 2001; Leonard & Sensiper, 1998). As such, the *initial founding team size* is used as a control variable. This is operationalized as a count variable representing the number of founders at the start-up's moment of foundation.

Furthermore, one of the most common explanations on start-up performance concerns the founding team's level of experience (Shepherd & Wiklund, 2006). The available human capital in a start-up is positively related to its business performance (Unger et al., 2011). The human capital can be divided into two categories, general- and specific human capital (Becker, 1964; Rauch & Rijsdijk, 2013). General human capital is not directly related to a certain job and consists of work experience and education (Rauch & Rijsdijk, 2013). Founders with more working experience are likely to have better judgement and more knowledge which will benefit them when founding a new start-up (Colombo & Grilli, 2010). As such, a count variable representing the cumulative *number of years working experience* of the founding team is used (Colombo & Grilli, 2010; Rauch & Rijsdijk, 2013).

Specific human capital is directly related to the start-ups activities and consists of industry-specific, start-up, and management experience (Brüderl et al., 1992; Rauch & Rijsdijk, 2013). The existent literature suggests that entrepreneurs who have more *industry experience* found better-performing ventures (Toft-Kehler, Wennberg, & Kim, 2014). Prior industry experience is a determinant for the performance of start-ups because the experience can be applied in the start-up, which can lead to higher quality products (Dahl & Reichstein, 2007; Delmar, 2006; Klepper, 2001). The industry experience is operationalised as a binary measure that represents whether any founder has previous working experience in an industry relevant to the start-up. This variable is author coded based on a combined review of the resume of all founders as well as the activities of the start-up.

In addition, founders with previous *managerial experience* are seen as better able to exploit opportunities and as such have a positive influence on the performance of a start-up (Dencker & Gruber, 2015; Toft-Kehler et al., 2014). I thus include a binary variable indicating whether the founding team has previous management experience.

Finally, regarding specific human capital, a lot of the relevant knowledge about creating a new company is learned-by-doing and as such, previous experience as a start-up founder is an important variable that influences start-up performance (Cassar, 2014; Delmar & Shane, 2006; Shane & Khurana, 2003). Delmar & Shane (2006). Because Delmar & Shane (2006) find that the difference between any and no *start-up experience* is the driver for differing start-up survival rates a binary indicator is used to indicate whether any founder had previous experience as a start-up founder (Cassar, 2014).

Previous studies also find that the market environment and the type of industry influence start-up performance (Schwartz & Hornych, 2010; Song et al., 2008; Wright & Stigliani, 2012). During conversations with experts from the accelerator they described that the *type of customer* formed an important element of the market environment. Therefore, a binary control variable that represents whether the start-up sells its products to businesses (B2B) or consumers (B2C) is also included in this research. In addition, Chatterjee & Hambrick (2007) find that the *type of product,* whether the start-up produced a hardware or a software technology, influences start-up performance. I therefore included a binary variable, representing whether the start-up offered a software or a hardware product. Start-ups with a software-hardware combination are coded as hardware start-ups.

Furthermore, I control for *gender differences* by including the percentage of males in the initial founding team as a control variable (Chowdhury, 2005; Verheul & Thurik, 2001). Finally, there are small differences between the accelerator program and they are located in countries with different institutional contexts and cultures (Climate-KIC, 2017). As such a categorical control variable which represents the *accelerator region* is also used.

3.3. Descriptive statistics and correlations

For 13 out of the 197 start-ups the performance data only indicates start-ups which had seized to exist at the moment of the survey. The sample size for the business performance variables is thus 184[3]. Table 2 shows the number of observations, mean, standard deviation and the Spearman correlations for each of the variables. Notable are the high correlations between the two measures of *firm size,* revenues and employees, with a correlation coefficient of 0.498 and between the total years of working experience and the presence of management experience with a correlation coefficient of 0.65. More detailed descriptive statistics for all variables, in the forms of bar diagrams and histograms, are provided in Appendix D. Three of the control variables, *accelerator region*, *type of product* and *type of customer*, are nominal variables and as such it is not possible to calculate the aforementioned descriptive statistics for these variables. It can be observed that most start-ups are located in the *accelerator region* the Netherlands, the *type of product* is mostly

[3] For the revenue and employee variable it is possible to include the non-surviving start-ups in the regression analyses as having zero revenues and employees. A robustness check which included the non-surviving start-ups for these variables had very similar results to the outcomes presented in this study.

hardware, and the *type of customer* are mostly consumers (see Appendix D). These three categories are used as the baseline categories in the analyses.

3.4. Imputation

For a number of variables there is missing data as some start-ups did not provide information on the respective variable (Figure 2). In order to include these start-ups in the analyses I use the *mice* package in R, which applies multivariate imputation by chained equations, to impute the missing variables (Buuren & Groothuis-Oudshoorn, 2011). The imputation through mice consists of a three step process. The first step, *imputation,* consists of multiple (ten) rounds in which the missing values are replaced by plausible data values. These plausible values are drawn from a distribution specifically modelled for each missing entry. This results in ten complete datasets in which the missing values have been replaced by plausible values. Secondly, in the *analysis* phase, the appropriate regression model is run to calculate the regression coefficients (Q) for each of the datasets. Third, in the *pool* stage, the method outlined in Rubin (1987, pp. 76–77) is used to calculate the mean over the regression coefficients $\hat{Q}^{(1)}, \ldots, \hat{Q}^{(10)}$ as well as the sum of the within- and between-imputation variance (Buuren & Groothuis-Oudshoorn, 2011). These are then reported as the coefficients for the model.

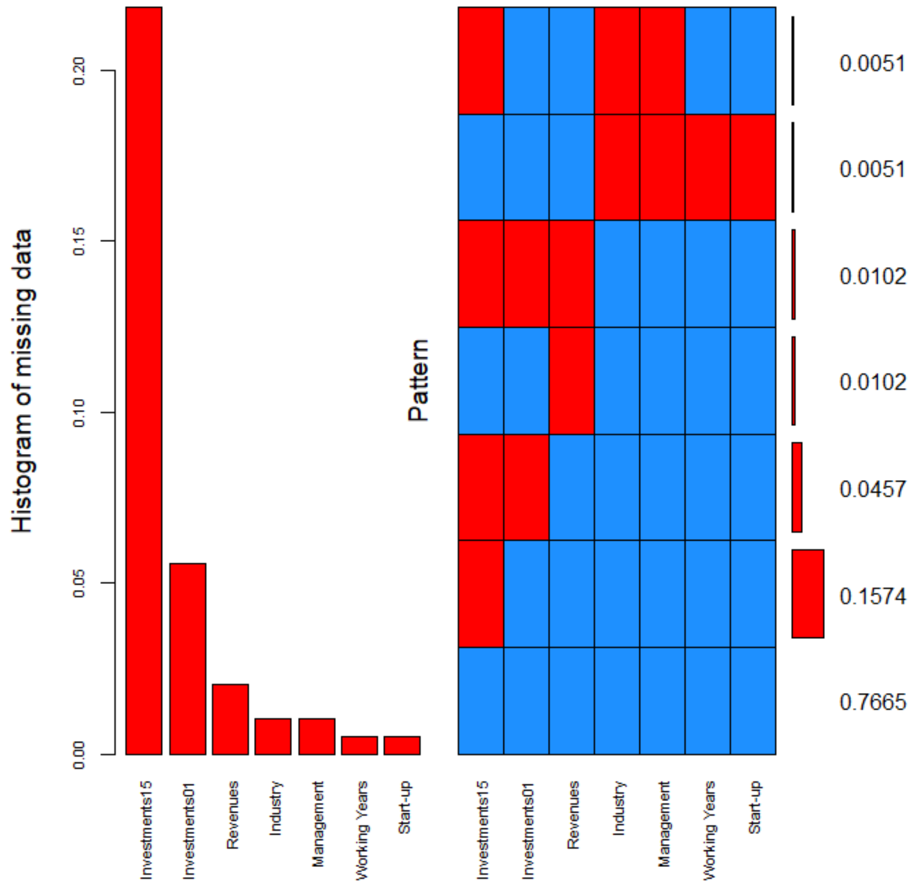


Figure 2: Overview of the percentage of missing variables and their combinations.

The number of missing observations for the ordinal scale investments is noticeably higher because this variable was only collected in the 2016 performance survey, while the 2014 and 2015 surveys used a binary scale. I therefore use the binary measure to test the robustness of the analyses with investments as the dependent variable.

Table 2: Descriptive statistics and correlations

	#	n	Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	13
Number of employees	**1**	184	5.565	5.01	1												
Revenues	**2**	180	1.256	1.18	0.498	1											
Cumulative investments (binary)	**3**	173	0.301	0.46	0.404	0.242	1										
Cumulative investments (ordinal)	**4**	141	1.688	1.10	0.434	0.238	0.965	1									
CO_2 impact score	**5**	197	2.873	1.05	0.048	-0.057	0.139	0.165	1								
Technological diversity change	**6**	197	0.291	1.93	0.057	-0.001	0.107	0.123	0.230	1							
Age	**7**	197	2.801	1.87	0.177	0.357	0.421	0.456	-0.056	0.036	1						
Initial number of founders	**8**	197	2.508	0.95	0.236	-0.063	0.049	-0.001	-0.098	0.026	-0.080	1					
Years working experience	**9**	196	19.296	19.03	-0.049	-0.108	0.035	-0.022	0.074	-0.103	0.040	0.183	1				
Industry experienceYN	**10**	195	0.585	0.49	-0.057	-0.032	-0.112	-0.161	0.075	-0.031	-0.062	0.085	0.496	1			
Management experienceYN	**11**	195	0.426	0.50	0.028	-0.014	0.004	0.003	0.116	-0.016	0.064	-0.058	0.652	0.263	1		
Startup experienceYN	**12**	196	0.469	0.50	0.024	-0.020	0.122	0.100	0.089	0.037	0.054	0.075	0.464	0.121	0.359	1	
% of males	**13**	197	0.889	0.23	0.104	0.154	0.176	0.166	0.035	-0.004	0.130	-0.120	-0.101	0.030	0.082	0.022	1

3.5. Data analysis

To test the hypotheses on the relation between the independent and dependent variables I perform multiple regression analyses using 'R'. The *number of employees*, is a count variable and as such a Poisson or negative binomial model is the appropriate regression model. The Poisson model is not the right fit for the data because the overdispersion is higher than the threshold of 2 (3.18) and significant (see Appendix E). Therefore, a negative binomial model is used for this variable. The *revenues*, *investments* and *climate performance* variables are all ordinal in nature and therefore an Ordinal Logit Model (OLM) is used for these dependent variables. To perform the robustness test with the binary investment variable a Binary Logit Model (BLM) is the appropriate model.

For each dependent variable two models are constructed, one including only the control variables and one including the independent variable, technological diversity change. I use the *lrtest* function in R to compare the two models and the McFadden Pseudo R^2 to report the performance of the respective models (Hoetker, 2007; Jackman, 2017; Zeileis & Hothorn, 2002). The McFadden Pseudo R^2 values of the models range from 0.04 to 0.14 indicating acceptable to good model fits (McFadden, 1974).

For each of the analyses I verify whether the appropriate assumptions hold (see Appendix D and E). The assumption of multicollinearity is tested by looking at the Spearman's correlations, which show no particularly worrisome correlations. I also calculate the VIF scores, which are all below 2 and as such there is no issue with multicollinearity of the independent and control variables (Field et al., 2012). Furthermore, scatterplots of the variables and the residuals show that the residuals are homoscedastic. Finally, there are no observations with a Cook's Distance larger than 1 and thus no outliers of particular concern (Cook & Weisberg, 1982; Field et al., 2012).

For the OLM analyses I also test whether the parallel regression assumption holds. This is done by comparing the full ordinal logit model with constrained models. In the constrained models one independent variable is taken out of the full model and allowed to vary with the outcome categories (Ari & Yildiz, 2014). Because the outcomes are not significant the parallel regression assumption is not violated .

4. Results

4.1. Business performance

This section describes the results of the regression analyses with *firm size*, which is operationalized as the *number of employees* and the *revenues*, and *investments* as the dependent variables. The results of the regression models for these dependent variables are shown in Table 3.

Table 3: Results of the Negative Binomial Models with the number of employees and the OLM models with revenues and investments as the dependent variable.

	Control employee	Full employee	Control revenues	Full revenues	Control invest	Full invest
Intercept	0.250 (0.381)	0.124 *(0.400)*	-	-	-	-
Technological Diversity	-	0.015 *(0.603)*	-	-0.030 *(0.701)*	-	0.055 *(0.616)*
Age	0.123 *(0.000)****	0.123 *(0.000)****	0.470 *(0.000)****	0.472 *(0.000)****	0.571 *(0.000)****	0.513 *(0.000)****
Initial founding team size	0.200 *(0.001)****	0.199 *(0.001)****	-0.069 *(0.668)*	-0.062 *(0.700)*	0.217 *(0.317)*	0.209 *(0.339)*
Total years of working experience	-0.006 *(0.138)*	-0.006 *(0.146)*	-0.025 *(0.036)**	-0.026 *(0.034)**	0.000 *(0.991)*	0.001 *(0.932)*
Industry Experience	-0.239 *(0.048)**	-0.236 *(0.051)*	0.367 *(0.273)*	0.361 *(0.281)*	-0.928 *(0.036)**	-0.932 *(0.035)**
Management Experience	0.226 *(0.114)*	0.214 *(0.130)*	0.161 *(0.677)*	0.176 *(0.649)*	-0.197 *(0.712)*	-0.206 *(0.700)*
Start-up Experience	0.117 *(0.337)*	0.121 *(0.332)*	0.213 *(0.516)*	0.225 *(0.496)*	0.426 *(0.305)*	0.400 *(0.339)*
Hardware-Software	0.307 *(0.008)***	0.310 *(0.007)***	0.790 *(0.011)**	0.790 *(0.011)**	-0.384 *(0.350)*	-0.366 *(0.374)*
B2B-B2C	0.249 *(0.048)**	0.258 *(0.042)**	0.596 *(0.087)*	0.580 *(0.098)*	-0.164 *(0.715)*	-0.145 *(0.750)*
Percentage of males	0.520 *(0.033)**	0.524 *(0.032)**	0.970 *(0.154)*	0.972 *(0.154)*	2.273 *(0.034)**	2.286 *(0.033)**
Accelerator DACH	0.341 *(0.013)**	0.338 *(0.013)**	0.504 *(0.181)*	0.506 *(0.180)*	0.634 *(0.214)*	0.621 *(0.224)*
Accelerator Nordics	-0.189 *(0.206)*	-0.187 *(0.210)*	-0.152 *(0.705)*	-0.173 *(0.670)*	-0.748 *(0.172)*	-0.736 *(0.180)*
N	184	184	184	184	184	184
McFadden R^2	0.058	0.059	0.099	0.099	0.141	0.142

*Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$*

Both models with the *number of employees* as the dependent variable have McFadden values of 0.06 (table 3). The difference between the two models and the influence of the technological diversity change variable on the number of employees are not significant at the 5% level. The McFadden value for both OLM models with *revenues* as the dependent variable are 0.10 (table 3). The value for the linear model is slightly higher, however the difference between the control and the full model is not significant. The technological diversity change variable also does not have a significant influence on the revenues earned by a start-up (5% level).

The OLM with *investments* as the dependent variable both have a McFadden value of 0.14 (table 3). The difference between the control and full model is not significant and neither is the coefficient for technological diversity change (5% level). The effect of the technological diversity change variable is thus not significant for investments. The results are the same for the robustness test with the binary investment variable (see Appendix E).

The results of the business performance analyses with *firm size* and *investments* as the dependent variables therefore do not provide evidence to support hypothesis 1.

Regarding the control variables for the *firm size* models, the *start-ups age* has a positive effect on both the number of employees and revenues which is significant (0.1% level). Furthermore, a larger *initial founding team* has a positive, significant effect on the number of employees (0.1% level). Interesting is that having experience in the same industry has a negative effect on the number of employees that is significant at the 5% level. While the founding teams *total years of working experience* has a small, but significant (5% level) negative effect on the revenues. This effect is robust when the average number of working experience per team member is used instead of the total years of working experience.

Furthermore, start-ups with a *software* product perform significantly better than their *hardware* counterparts for both dimensions of firm size (5% level). Also, start-ups selling their products to *consumers* have significantly larger number of employees than their counterparts who deliver to *businesses* (5% level). Furthermore, a larger percentage of males has a positive effect on the number of employees (5% level). Finally, it is found that start-ups from the *DACH* region are significantly larger than their counterparts from the *Netherlands* and *Nordics* (1% level).

Regarding the investment model, start-up *age* has a positive effect on the cumulative investments at the 0.1% level. Furthermore, industry experience has a negative effect on the cumulative investments which is significant at the 5% level. Finally, start-ups with a higher percentage of male founders have gathered significantly more investments (5% level). The results of the robustness check with the binary variable for investments are highly similar except for the fact that in the robustness model the negative effect of industry experience is not significant anymore.

4.2. Climate performance

Table 4 shows the outcomes for the OLM with climate performance as the dependent variable. The control model has a McFadden value of 0.05 while the McFadden value for the full model is 0.06. The full model is significantly better than the control model (5% level) and the technological diversity change variable has a positive effect on the potential climate performance of the start-up, which is significant at the 5% level. The results of the *climate performance* analyses therefore provide support for hypothesis 2.

Table 4: Results of the Ordinary Logit Models with climate performance as the dependent variable.

	Control Model	**Full Model**
Technological Diversity	-	0.182 *(0.009)***
Age	-0.003 *(0.966)*	-0.019 *(0.806)*
Initial founding team size	-0.222 *(0.157)*	-0.258 *(0.110)*
Total years of working experience	-0.017 (0.145)	-0.012 *(0.286)*
Industry Experience	0.459 *(0.145)*	0.455 *(0.147)*
Management Experience	0.539 *(0.139)*	0.466 *(0.200)*
Start-up Experience	0.337 *(0.274)*	0.284 *(0.358)*
Hardware-Software	-1.170 *(0.000) ****	-1.172 *(0.000)****
B2B-B2C	-0.170 *(0.592)*	-0.149 *(0.638)*
Percentage of males	0.192 *(0.731)*	0.189 *(0.736)*
Accelerator DACH	0.700 *(0.045)**	0.710 (0.042)*
Accelerator Nordics	0.697 *(0.070)*	0.751 *(0.055)*
N	197	197
McFadden R^2	0.048	0.060

*Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$*

In both the control and the full model the control variables representing the *type of product* are significant at the 0.1% level. Start-ups with a *hardware* product have significantly higher potential to reduce CO_2e emissions than their *software* counterparts. Furthermore, start-ups from the DACH region have significantly higher climate potential than their Dutch counterparts (5% level).

5. Discussion

The goal of this study was to understand how a start-up's influence on the diversity of the technological system, influences the start-up's performance along the business and climate dimensions. To do so the following research question was formulated: *What is the influence of the degree of technological diversity change caused by climate-focused start-ups' on their business and climate performance?*

The findings show that the creation of technological diversity is not associated with a significantly lower business performance, in the form of firm size and investments. *Hypothesis 1 is therefore not accepted*. For the climate dimension of performance the results show that diversity creation by a start-up has a positive effect on the climate performance in terms of the potential to reduce CO_2e emissions. This confirms the expectation that start-ups with a more unique technology have a higher potential to reduce CO_2e emissions. *Hypothesis 2 is therefore accepted*.

5.1. Theoretical implications

This study used concepts from the macro-level technological trajectory literature to derive micro-level hypotheses about the influence of technological diversity change caused by a start-up on the start-up's performance (Yu et al., 2011). Regarding business performance the study did not find support for the hypothesis that start-ups which create technological diversity have significantly lower business performance. This study was therefore not able to confirm the arguments that: (1) technologies which open up new trajectories, and create technological diversity, are associated by customers as having higher risks which reduces their adoption (Marra et al., 2003; Verhoeven et al., 2016), and (2) technologies that reduce diversity have higher performance because they can profit from economies of scale and learning effects obtained through similar technologies (Anandarajah & McDowall, 2015; Rogner, 1998; Yu et al., 2011).

A potential explanation for this can be found in the literature on competitive advantage (Porter, 1998). A start-up's competitive advantage is strongly influenced by the level of uniqueness of its resources, including its technology (Barney, 1995; Granstrand, 1998; Mahoney & Pandian, 1992; Teece, Pisano, & Shuen, 1997). A non-unique technology, which reduces technological diversity, has less competitive advantage which could lead to a lower start-up business performance (Aharonson & Schilling, 2016; Debackere et al., 1999; Harrigan & DiGuardo, 2014; Zahra, 1996). This effect is opposite to the relation as expected from the technological trajectory literature and the mechanisms could thus have countered each other.

Regarding the climate performance, this study found that technological diversity creation has a positive influence on a start-up's potential to reduce CO_2e emissions. These findings correspond with the expectation that start-ups which create diversity have not yet profited from learning effects (Aharonson & Schilling, 2016; Nemet, 2006). These start-ups therefore have a higher technological potential which results in a higher potential climate impact (Aghion et al., 2012; Aghion et al., 2014; Bjornali & Ellingsen, 2014; Nemet, 2009).

This study uses climate, in the form of the potential to reduce CO_2e emissions, as a new dimension of start-up performance. In doing so it answered the call from previous literature to include the environmental dimension in research of business performance (Bjornali &

Ellingsen, 2014; Dean & McMullen, 2007; Elkington, 1994; Meyskens & Carsrud, 2013; Slaper & Hall, 2011). This research also provides support for the notion that the business and climate dimensions are fundamentally different from each other. First, the results show both negative and very low correlations between the climate and the business dimensions of performance indicating that they are not related. Furthermore, control variables which positively influence business performance (such as start-up age) do not have the same effect on climate performance. Interesting is also that although start-ups with a *software* product have significantly higher revenues and number of employees, they also have a significantly *lower potential to reduce* CO_2e *emissions* than start-ups with a *hardware* product.

There are a number of additional interesting outcomes concerning the control variables which are worth discussing.

First, the negative effect of industry experience on business performance, which is significant for both the number of employees and the investments of a start-up, contradicts previous literature which suggested that *industry experience* leads to better-performing ventures (Dahl & Reichstein, 2007; Delmar & Shane, 2006; Klepper & Sleeper, 2005; Toft-Kehler et al., 2014). One potential explanation for this negative influence is that, in this study, experience on a relevant topic which is obtained while working at a university is coded as having industry experience. Previous research has shown that ventures started from universities generally perform less well than other start-ups (Harrison & Leitch, 2010). Including experience obtained at universities could thus have led to the negative influence of industry experience on business performance. Also surprising, and contradictory to the existing literature (Colombo & Grilli, 2010; Rauch & Rijsdijk, 2013), is that the total amount of working experience of the founding team has a significant negative effect on the revenues dimension of business performance. A potential explanation for this could be that very experienced founders have more capital available for their business and as such are less likely to focus on acquiring early revenues for their business (Headd, 2003).

Although consistent with previous research (Chowdhury, 2005; Kanze et al., 2017; Malmström et al., 2017; Verheul & Thurik, 2001), it is noteworthy that a higher percentage of males in the initial founding team is associated with significantly higher investments and more employees. The first potential explanation is that male entrepreneurs have more confidence in their own abilities and are more ambitious which influences their performance (Verheul & Thurik, 2001). Secondly, because a majority of investors and entrepreneurs are male, investors (subconsciously) look for similar people when investing in start-ups, leading to a gender bias towards investing in male entrepreneurs (Kanze et al., 2017; Malmström et al., 2017).

Finally, the results show that start-ups in the DACH region have a higher business performance which is significant for their number of employees. A potential explanation that came up when talking to start-ups from this region is the fact that these start-ups are dealing with a larger home market. As a result they require larger teams to deal with the travel required to reach different parts of their market.

5.2. Limitations

The business performance data used in this research is obtained from performance surveys at three different moments in time. In order to correct for the differing time lines I use the start-up age, calculated as the year of the performance survey minus the start-up foundation year. Nevertheless fluctuations in the external environment between the different years, such as the content of the accelerator program, could have influenced the results.

In start-up research *profitability* is also a frequently used dimension of business performance. This variable is not included in this research for two reasons. First, because the climate-focused industry is generally very resource intensive the start-ups require additional resources, through investments, before they can achieve profitability (Bjornali & Ellingsen, 2014; Eyraud et al., 2013; Hyytinen et al., 2015). Profitability is therefore a long term goal for the start-ups included in this study and is therefore not a relevant dependent variable at this stage. Second, the annual Climate-KIC evaluation survey which is used to obtain the business performance data does not include a measure for profitability.

The climate measure in the form of the potential to reduce CO_2e emissions proved reliable using expert scores as a verification measure. Nevertheless, it would be preferable to have quantitative numbers for this measure. However, this requires individual collaboration from each start-up regarding the exact features of their business. It was therefore not possible to include such a measure in this research. This is a potential avenue for future research as Climate-KIC Scotland and the ECCI recently started writing Carbon Audit Reports, which quantify the potential reduction in CO_2e emissions for each of the start-ups in the local accelerator program. Furthermore, by combining these estimations with the revenues of the start-ups, future studies can also calculate the *realized* climate performance. This makes it possible to evaluate the actual climate impact and is a next step in evaluating the climate performance of start-ups. During this study I did investigate combining the potential climate scores and revenues to construct a variable for the actual climate performance. However, the results of this process were not deemed to be consistent and reliable and therefore not included in this thesis.

Regarding the independent variable, it is important to note that the change in technological diversity (the entropy change, Eq. 2) was calculated based on the full 303 start-ups and not only regarding the start-ups founded prior to that particular start-up. This was the case because the relatively low number of start-ups in the earlier years seriously biased the results if this was considered. The fact that the changes in technological diversity were calculated as if all start-ups were founded at a single time is a limitation of this study.

A potential concern with this research could be the limited generalizability of the sample towards start-ups not participating in incubation and accelerator programs. However, there are two arguments which ensure that this is not problematic. First, although being accelerated does influence start-up performance (van Rijnsoever et al., 2016), it is not expected to do so differently for technological diversity creating or reducing start-ups. Secondly, because the climate-focused industry is generally very resource intensive, climate-focused start-ups encounter large liabilities of newness and smallness (Bjornali & Ellingsen, 2014; Eyraud et al., 2013; Hyytinen et al., 2015). The start-ups often require additional resources to overcome these liabilities which they acquire by entering accelerators and incubators (Klofsten et al., 2016; Shane & Khurana, 2003; van Rijnsoever et al., 2016; van Weele et al., 2017). They are able to do so because there is a large availability of these programs (Bank & Kanda, 2016; Bergek & Norrman, 2008; Tamásy, 2007). It can thus be expected that the majority of climate-focused start-ups use the support of an acceleration or incubation program. Therefore, the sampling frame of the study is not expected to influence the generalizability.

Finally, this research did not include a measure for the ambition levels of the entrepreneurs, which is known to influence business performance (Baum & Locke, 2004). It is possible that more ambitious entrepreneurs start more technologically diverse businesses because they believe they will be able to overcome any barriers. As such, entrepreneur ambition could be a cofounding factor that influences both the dependent and independent variable, making this a relevant avenue to explore in future research.

5.3. Practical implications

From a societal perspective climate-focused start-ups have been considered to have three main functions: (1) they play a key role in climate mitigation, (2) they are a major driver of economic development, and (3) they contribute to technological change by introducing diversity into the system (Aerts et al., 2007; Almeida & Kogut, 1997; Andrea & Roberto, 2014; Bjornali & Ellingsen, 2014; Cumming et al., 2014; Klepper & Sleeper, 2005). This research reveals an interesting dynamic between the three functions, as start-ups who create technological diversity have a higher potential to reduce CO_2e emissions but not a higher business performance. Furthermore, start-up's with a software technology are found to be more likely to have higher business performance but they also have less potential to reduce CO_2e emissions.

From the results of this study it thus becomes clear that there is a discrepancy between the climate potential and business dimensions of performance. The goals of climate and business performance do not align, and while business performance is necessary to translate the potential climate impact into *realized* climate performance, it is not sufficient. The implication for policy makers and incubator and accelerator personnel is that achieving these goals simultaneously is not a straightforward process and it is unlikely to result from the current approach. Instead these actors are presented with a tough challenge to balance the different societal impacts of start-ups.

However, this study also identifies a particular group of start-ups which could help solve this dilemma. There is a small group of start-ups in this research which performed well regarding both business performance (in the form of revenues) and climate potential (see Appendix F). These start-ups are thus able to achieve both strong business and climate performance. Understanding how these start-ups differ from the other start-ups is a future research avenue which could help policy makers and incubators and accelerators to aim for both high business and climate performance.

The fact that incubators and accelerators are currently evaluated based on the *business performance* of their start-ups this study has another important implication (Sepulveda, 2012; Tamásy, 2007). The findings in this research show that the climate-focused start-ups with a software product are significantly more likely to lead to higher business performance. The quickest way for incubator and accelerator personnel to meet their evaluation requirements is thus to select these start-ups into their programme.

5.4. Future research

This research shows that there are fundamental differences between start-ups based on their technology and type of product, thus supporting the argument that these start-ups also require different types of support when establishing their business (Soetanto & Jack, 2013). Future research could focus on these differences to better understand how the needs of different types of start-ups differ. The increased understanding of this topic could help incubators and accelerators to better support start-ups with a different nature.

Currently, most research on accelerated start-ups studies the performance of start-ups at a single point in time rather than taking a longitudinal approach, which accounts for the different growth curves of the start-ups (Kebbi & Valliere, 2007). As the Climate-KIC accelerator is annually gathering start-up performance data this data could be used to fill this research gap by studying the performance of accelerated start-ups over time.

The topic modelling approach which is used to define the independent variable has proven to be an interesting and reliable way to research technological fields and the contribution of individual technologies to these fields (Páez-Avilés et al., 2016). In this study it is applied to study how a change in technological diversity caused by a start-up influences its performance. However, this method could also be used to shed insights into another research gap. The patterns of technological development are known to differ between industries, but it is still unknown how different actors contribute to these different patterns (Breschi et al., 2000; Malerba & Orsenigo, 1996). In order to fill this research gap the topic modelling method could be used to research how start-ups, and other types of actors, have contributed to the development of particular technologies (e.g. wind or tidal power) over time, using the patent texts and applicants of these technologies.

6. References

Acs, Z. J., & Audretsch, D. B. (2005). Entrepreneurship, innovation, and technological Change, *1*(4), 149–195.

Ács, Z. J., & Varga, A. (2005). Entrepreneurship, agglomeration and technological change. *Small Business Economics*, *24*(3), 323–334. https://doi.org/10.1007/s11187-005-1998-4

Aerts, K., Matthyssens, P., & Vandenbempt, K. (2007). Critical role and screening practices of European business incubators. *Technovation*, *27*(5), 254–267. https://doi.org/10.1016/j.technovation.2006.12.002

Aggarwal, C. C., & Zhai, C. X. (2013). *Mining text data*. *Mining Text Data*. https://doi.org/10.1007/978-1-4614-3223-4

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., & Reenen, J. Van. (2012). Carbon Taxes , Path Dependency and Directed Technical Change : Evidence from the Auto Industry. In *Center for Economic performance*.

Aghion, P., Hepburn, C., Teytelboym, A., & Zenghelis, D. (2014). Path dependence, innovation and the economics of climate change. *The New Climate Change Economy*, (November). Retrieved from http://2014.newclimateeconomy.report/wp-content/uploads/2014/11/Path-dependence-and-econ-of-change.pdf

Aharonson, B. S., & Schilling, M. a. (2016). Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution. *Research Policy*, *45*(1), 81–96. https://doi.org/10.1016/j.respol.2015.08.001

Almeida, P., & Kogut, B. (1997). The Exploration of Technological Diversity and the Geographic Localization of Innovation. *Small Business Economics*, *9*(1), 21–31. https://doi.org/10.1023/A:1007995512597

Anandarajah, G., & McDowall, W. (2015). Multi-cluster technology learning in TIMES: a transport sector case study with TIAM-UCL. In *Informing Energy and Climate policies using Energy Stems Models* (pp. 261–278). Springer International Publishing.

Andrea, F., & Roberto, G. (2014). Spin-off performance in the start-up phase – a conceptual framework. *Journal of Small Business and Enterprise Development*, *21*(3), 528–544. https://doi.org/10.1108/JSBED-04-2014-0055

Annand, S. (2017). Finding Optimal Number of Clusters. Retrieved April 2, 2018, from https://www.r-bloggers.com/finding-optimal-number-of-clusters/

Ari, E., & Yildiz, Z. (2014). Parallel Lines Assumption in Ordinal Logistic Regression and Analysis Approaches. *International Interdisciplinary Journal of Scientific Research*, *1*(3), 8–23.

Arts, S., Appio, F. P., & Van Looy, B. (2013). Inventions shaping technological trajectories: Do existing patent indicators provide a comprehensive picture? *Scientometrics*, *97*(2), 397–419. https://doi.org/10.1007/s11192-013-1045-1

Bank, N., & Kanda, W. (2016). Tenant recruitment and support processes in sustainability-profiled business incubators. *Industry and Higher Education*, *30*(4), 267–277. https://doi.org/10.1177/0950422216659567

Barney, J. B. (1995). Looking inside for Competitive Advantage Looking inside for competitive advantage. *The Academy of Management Executive (1993-2005) Vol. 9, No. 4 (Nov., 1995), Pp. 49-61*, *9*(4), 49–61.

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*, *89*(4), 587–598. https://doi.org/10.1037/0021-9010.89.4.587

Becker, G. S. (1964). *Human capital*. New York.

Bennett, S. J. (1991). *Ecopreneuring : the complete guide to small business opportunities from the environmental revolution*. Wiley.

Bergek, A., & Norrman, C. (2008). Incubator best practice: A framework. *Technovation*, *28*(1–2), 20–28. https://doi.org/10.1016/j.technovation.2007.07.008

Bjornali, E. S., & Ellingsen, A. (2014). Factors Affecting the Development of Clean-tech Start-Ups: A Literature Review. *Energy Procedia*, *58*(58), 43–50. https://doi.org/10.1016/j.egypro.2014.10.407

Blei, D. M. (2011). Probabilistic Topic Models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., & Lafferty, J. D. (2009). Topic Models. *Text Mining: Classification, Clustering, and Applications*, 71–89. https://doi.org/10.1145/1143844.1143859

Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological Regimes and Schumpeterian Patterns of Innovation. *The Economic Journal*, *110*(463), 388–410. Retrieved from http://www.jstor.org/stable/2566240

Brüderl, J., Preisendörfer, P., & Ziegler, R. (1992). Survival Chances of Newly Founded Business Organizations. *American Sociological Review*, *57*(2), 227–242.

Bryman, A. (2012). *Social Research Methods,*. New york: Oxforfd university press. Retrieved from http://www.bol.com/nl/p/social-research-methods/1001004010589585/

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, *45*(3). https://doi.org/10.18637/jss.v045.i03

Calel, R., & Dechezlepretre, A. (2013). *Environmental Policy and Directed Technological Change : Evidence from the European carbon market* (Vol. 98). https://doi.org/10.1162/REST_a_00470

Cassar, G. (2014). Industry and startup experience on entrepreneur forecast performance in new firms. *Journal of Business Venturing*, *29*(1), 137–151. https://doi.org/10.1016/j.jbusvent.2012.10.002

Chatterjee, A., & Hambrick, D. C. (2007). It’s All about Me: Narcissistic Chief Executive officers and their effects on Company Strategy and Performance. *Administrative Science Quarterly*, *52*, 351–386. https://doi.org/10.2189/asqu.52.3.351

Chouinard, Y., Ellison, J., & Ridgeway, R. (2011). The Sustainable Economy. *Harvard Business Review*, *October*(October).

Chowdhury, S. (2005). Demographic diversity for building an effective entrepreneurial team: Is it important? *Journal of Business Venturing*, *20*(6), 727–746. https://doi.org/10.1016/j.jbusvent.2004.07.001

Cicchetti, D. V. (1994). Guidlines, Criteria, and Rules of Thumb for Evalauting Normed and Standardized Assessment Instruments in Psychology. *Psychological Assessment*, *6*(4), 284–290.

Climate-KIC. (2017). Climate-KIC Accelerator: where great ideas grow into great businesses.

Climate-KIC. (2018). Climate-KIC Accelerator. Retrieved March 22, 2018, from https://accelerator.climatekic.nl/

Cohen, B., & Winn, M. I. (2007). Market imperfections, opportunity and sustainable entrepreneurship. *Journal of Business Venturing*, *22*(1), 29–49. https://doi.org/10.1016/j.jbusvent.2004.12.001

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

Cook, R. D., & Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman and Hall. https://doi.org/10.2307/1269506

Cumming, D., Johan, S., & Zhang, M. (2014). The Economic Impact of Entrepreneurship: Comparing International Datasets. *Corporate Governance (Oxford)*, *22*(2), 162–178. https://doi.org/10.1111/corg.12058

Dahl, M. S., & Reichstein, T. (2007). Are You Experienced? Prior Experience and the Survival of New Organizations. *Industry & Innovation*, *14*(5), 497–511. https://doi.org/10.1080/13662710701711414

Dean, T. J., & McMullen, J. S. (2007). Toward a theory of sustainable entrepreneurship: Reducing environmental degradation through entrepreneurial action. *Journal of Business Venturing*, *22*(1), 50–76. https://doi.org/10.1016/j.jbusvent.2005.09.003

Debackere, K., Luwel, M., & Veugelers, R. (1999). Can technology lead to a competitive advantage? A case study of Flanders using european patent data. *Scientometrics*, *44*(3), 379–400. https://doi.org/10.1007/BF02458486

Degroof, J. J., & Roberts, E. B. (2004). Overcoming Weak Entrepreneurial Infrastructures for Academic Spin-Off Ventures. *The Journal of Technology Transfer*, *29*(3/4), 327–352. https://doi.org/10.1023/B:JOTT.0000034126.23592.23

Delmar, F. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, *4*(3), 215–247. https://doi.org/10.1177/1476127006066596

Delmar, F., & Shane, S. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization*, *4*(3), 215–247. https://doi.org/10.1177/1476127006066596

Dencker, J. C., & Gruber, M. (2015). The effects of opportunities and founder experience on new firm performance. *Strategic Management Journal*, *36*(2), 1035–1052.

https://doi.org/10.1002/smj.2269

Dietz, R., & O'Neill, D. (2013). *Enough Is Enough: Building a Sustainable Economy in a World of Finite Resources* (1st ed.). Abingdon: Wilsted & Taylor Publishing Services. Retrieved from https://books.google.nl/books?hl=en&lr=&id=YVTzIeBM_zUC&oi=fnd&pg=PP2&ots=GQglW48FXQ&sig=rZf8i4_CqLBtMr_BGaAYTYeRdTs#v=onepage&q&f=false

Dosi, G., & Nelson, R. R. (2009). *Technical change and industrial dynamics as evolutionary processes* (Lem Working Paper).

Dosi, G., & Nelson, R. R. (2013). The Evolution of Technologies: An Assesment of the State-of-the-Art. *Eurasian Business Review*, *3*(1), 3–46. https://doi.org/10.14208/BF03353816

Elkington, J. (1994). Towards the sustainable corporation: Win-win-win business strategies for sustainable development. *California Management Review*. Retrieved from http://cmr.ucpress.edu/content/36/2/90.abstract

European Commission. (1998). *Risk Capital: A Key to Job Creation in the European Union*. *European Commission, Brussels, Communication from the Commission to the Council and the European Parliament.*

Eveleens, C. P., van Rijnsoever, F. J., & Niesten, E. M. M. I. (2016). *How network-based incubation helps start-up performance: a systematic review against the background of management theories*. *Journal of Technology Transfer*. Springer US. https://doi.org/10.1007/s10961-016-9510-7

Eyraud, L., Clements, B., & Wane, A. (2013). Green investment: Trends and determinants. *Energy Policy*, *60*(April 2017), 852–865. https://doi.org/10.1016/j.enpol.2013.04.039

Feinerer, I. (2017). Introduction to the tm Package: Text Mining in R. *R Vignette*, 1–8. https://doi.org/10.1201/9781420068740

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics Using R*. London: SAGE Publications. https://doi.org/10.1017/CBO9781107415324.004

Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science*, *47*(1), 117–132. https://doi.org/10.1287/mnsc.47.1.117.10671

Fontana, R., Nuvolari, A., & Verspagen, B. (2009). Mapping technological trajectories as patent citation networks. An application to data communication standards. *Economics of Innovation and New Technology*, *18*(4), 311–336. https://doi.org/10.1080/10438590801969073

Fontes, M., & Coombs, R. (2001). Contribution of new technology-based firms to the strengthening of technological capabilities in intermediate economies. *Research Policy*, *30*(1), 79–97. https://doi.org/10.1016/S0048-7333(99)00095-5

Gamer, T., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various Coefficients of Interrater Reliability and Agreement*. Retrieved from http://www.r-project.org

Gerlach, A. (2003). Sustainable entrepreneurship and innovation. *Conference Proceedings of Conference Corporate Social Responsibility and Environmental Management 2003 in Leeds, UK*, (1991), 101–110.

Gibbs, D. C. (2006). Sustainability Entrepreneurs, Ecopreneurs and the Development of a Sustainable Economy. *Greener Management International*. https://doi.org/10.9774/GLEAF.3062.2006.au.00007

Gohar, L. K., & Shine, K. P. (2007). Equivalent CO2 and its use in understanding the climate effects of increased greenhouse gas concentrations. *Weather*, *62*(11), 307–311. https://doi.org/10.1002/wea.103

Graham, S. J., & Sichelham, T. (2008). Why Do Start-Ups Patent. *Berkeley Technology Law Journal*, *23*(3). https://doi.org/10.15779/Z38M111

Granstrand, O. (1998). Towards a Theory of the. *Technology-Based Firm, Research Policy*, *27*(December), 465–485. https://doi.org/10.1016/S0048-7333(98)00067-5

Groenewegen, G. J., & De Langen. (2012). Critical Success Factors of the Survival of Start-Ups with a Radical Innovation, *2*(3), 155–171.

Grün, B., & Hornik, K. (2011). topicmodels : An R Package for Fitting Topic Models. *Journal of Statistical Software*, *40*(13), 1–30. https://doi.org/10.18637/jss.v040.i13

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34. https://doi.org/10.20982/tqmp.08.1.p023

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation fequency and the value of patented invetions. *Review of Economics and Statistics*, *10*(2), 130–153. https://doi.org/10.1086/250095

Harrison, R. T., & Leitch, C. (2010). Voodoo institution or entrepreneurial university? spin-off companies, the entrepreneurial system and regional development in the UK. *Regional Studies*, *44*(9), 1241–1262. https://doi.org/10.1080/00343400903167912

Headd, B. (2003). Redefining Business Success: Distinguishing between Closure and Failure. *Small Business Economics*, *21*(1), 51–61. https://doi.org/10.1023/A:1024433630958

Helmers, C., & Rogers, M. (2011). Does patenting help high-tech start-ups? *Research Policy*, *40*(7), 1016–1027. https://doi.org/10.1016/j.respol.2011.05.003

Hoetker, G. (2007). The use of logit and probit models in strategic managment research: critical issues. *Strategic Management Journal*, *28*, 3331–343. https://doi.org/10.1002/smj.582

Hotho, A., Nürnberger, A., & Paaß, G. (2005). A Brief Survey of Text Mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, *20*, 19–62. https://doi.org/10.1111/j.1365-2621.1978.tb09773.x

Huang, Y. F., & Chen, C. J. (2010). The impact of technological diversity and organizational slack on innovation. *Technovation*, *30*(7–8), 420–428. https://doi.org/10.1016/j.technovation.2010.01.004

Hyytinen, A., Pajarinen, M., & Rouvinen, P. (2015). Does innovativeness reduce startup survival rates? *Journal of Business Venturing*, *30*(4), 564–581. https://doi.org/10.1016/j.jbusvent.2014.10.001

Ibenholt, K. (2002). Explaining learning curves for wind power. *Energy Policy*, *30*(13), 1181–1189. https://doi.org/10.1016/S0301-4215(02)00014-9

Jackman, S. (2017). *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory*. Sydney. Retrieved from http://github.com/atahk/pscl

Kanze, D., Huang, L., Conley, M. A., & Higgins, E. T. (2017). Male and female entrepreneurs get asked different questions by VCs - and it affects how much funding they get. *Harvard Business Review*, *June*, 1–7. Retrieved from https://hbr.org/2017/06/male-and-female-entrepreneurs-get-asked-different-questions-by-vcs-and-it-affects-how-much-funding-they-get

Kaplan, S., & Vakili, K. (2013). Studying Breakthrough Innovations Using Topic Modeling: A Test Using Nanotechnology Patents. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2261731

Kebbi, A., & Valliere, D. (2007). The Double J-Curve : A Model for Incubated Start-ups. In *European Conference on Innovation and Entrepreneurship* (pp. 371–380). Academic Conferences International Limited.

Klepper, S. (2001). Employee startups in high-tech industries. *Industrial and Corporate Change*. Retrieved from https://academic.oup.com/icc/article-abstract/10/3/639/1276654

Klepper, S., & Sleeper, S. (2005). Entry by Spinoffs. *Management Science*, *51*(8), 1291–1306. https://doi.org/10.1287/mnsc.1050.0411

Klofsten, M., Bank, N., & Bienkowska, D. (2016). The Role of Incubators in Supporting Sustainable Ent- repreneurship. *The International Journal of Entrepreneurship and Innovation*, *3*(February), 1–52.

Klotz, a. C., Hmieleski, K. M., Bradley, B. H., & Busenitz, L. W. (2013). *New Venture Teams: A Review of the Literature and Roadmap for Future Research*. *Journal of Management* (Vol. 40). https://doi.org/10.1177/0149206313493325

Lee, H., Kihm, J., Choo, J., Stasko, J., & Park, H. (2012). iVisClustering: An Interactive Visual Document Clustering via Topic Modeling. *Computer Graphics Forum*, *31*(3pt3), 1155–1164. https://doi.org/10.1111/j.1467-8659.2012.03108.x

Mahoney, J. T., & Pandian, J. R. (1992). The Resource-Based View Within the Conversation of Strategic Management. *Strategic Management Journal*, *13*(5), 363–380. Retrieved from http://www.jstor.org/stable/2486455

Malerba, F. (2009). Increase learning, break knowledge lock-ins and foster dynamic complementarities: evolutionary and system perspectives on technology policy in industrial dynamics. *The New Economics of Technology Policy*. Retrieved from https://books.google.nl/books?hl=en&lr=&id=oscVXoJRKiMC&oi=fnd&pg=PA33&dq=malerba+increase+learning&ots=-YGuQhpiyU&sig=kBcihXMGwHiCj0oIECrpB5SzUbg

Malerba, F., & Orsenigo, L. (1996). Schumpeterian patterns of innovation are technology-specific. *Research Policy*, *25*(3), 451–478. https://doi.org/10.1016/0048-7333(95)00840-3

Malmström, M., Johansson, J., & Wincent, J. (2017). Gender Stereotypes and Venture Support Decisions: How Governmental Venture Capitalists Socially Construct Entrepreneurs’ Potential. *Entrepreneurship: Theory and Practice*, *41*(5), 833–860. https://doi.org/10.1111/etap.12275

Marra, M., Pannell, D. J., & Abadi, A. (2003). The economics of risk, uncertainty and learning in the adoption of new agricultural technologies: Where are we on the learning curve? *Agricultural Systems*, *75*, 215–234. https://doi.org/10.1016/S0308-521X(02)00066-5

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142. https://doi.org/10.1108/eb028592

Meyer, D., Hornik, K., & Feinerer, I. (2008). Institutional Repository. *Journal of Statistical Software*, *25*(5), 1–54. https://doi.org/10.13140/RG.2.1.4785.3929

Meyskens, M., & Carsrud, A. L. (2013). Nascent green-technology ventures: A study assessing the role of partnership diversity in firm success. *Small Business Economics*, *40*(3), 739–759. https://doi.org/10.1007/s11187-011-9400-1

Murmann, J. P., & Frenken, K. (2006). Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Research Policy*, *35*(7), 925–952. https://doi.org/10.1016/j.respol.2006.04.011

Murphy, G., Trailer, J., & Hill, R. (1996). Measuring Research Performance in Entrepreneurship. *Journal of Business Research*, *36*(1), 15–23. https://doi.org/10.1016/0148-2963(95)00159-X

Nemet, G. F. (2006). Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy*, *34*(17), 3218–3232. https://doi.org/10.1016/j.enpol.2005.06.020

Nemet, G. F. (2009). Demand-pull, technology-push, and government-led incentives for non-incremental technical change. *Research Policy*, *38*(5), 700–709. https://doi.org/10.1016/j.respol.2009.01.004

Office of the Press Secretary. (2015). Remarks by the President on Sustainable Development Goals. Retrieved March 27, 2017, from https://obamawhitehouse.archives.gov/the-press-office/2015/09/27/remarks-president-sustainable-development-goals

Ortín-Ángel, P., & Vendrell-Herrero, F. (2014). University spin-offs vs. other NTBFs: Total factor productivity differences at outset and evolution. *Technovation*, *34*(2), 101–112. https://doi.org/10.1016/j.technovation.2013.09.006

Páez-Avilés, C., Van Rijnsoever, F. J., Juanola-Feliu, E., & Samitier, J. (2016). Multidisciplinarity breeds diversity: The influence of innovation project characteristics on diversity creation in nanotechnology. https://doi.org/10.1007/s10961-016-9553-9

Parrish, B. D. (2010). Sustainability-driven entrepreneurship: Principles of organization design. *Journal of Business Venturing*, *25*(5), 510–523. https://doi.org/10.1016/j.jbusvent.2009.05.005

Peña, I. (2004). Business Incubation Centers and New Firm Growth in the Basque Country. *Small Business Economics*, *22*(3), 223–236.

https://doi.org/10.1023/B:SBEJ.0000022221.03667.82

R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rasmussen, E., Bulanova, O., Jensen, A., & Clausen, T. (2012). The Impact of Science-Based Entrepreneurial Firms -a Literature Review and Policy Synthesis. *Unreleased*, 117.

Rauch, A., & Rijsdijk, S. A. (2013). The Effects of General and Specific Human Capital on Long-Term Growth and Failure of Newly Founded Businesses. *Entrepreneurship: Theory and Practice*, *37*(4), 923–941. https://doi.org/10.1111/j.1540-6520.2011.00487.x

Richard, P. J., Devinney, T., Yip, G., & Johnson, G. (2009). Measuring organizational performance: towards methodological best practice. *Journal of Management*, *35*(3), 718–804. https://doi.org/10.1177/0149206308330560

Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of Documentation*, *60*(5), 503–520. https://doi.org/10.1108/00220410410560582

Rogner, H. (1998). Hydrogen technologies and the technology learning curve. *International Journal of Hydrogen Energy*, *23*(9), 833–840. https://doi.org/10.1016/S0360-3199(97)00124-9

Rothaermel, F. T., & Thursby, M. (2005). University-incubator firm knowledge flows: Assessing their impact on incubator firm performance. *Research Policy*, *34*(3), 305–320. https://doi.org/10.1016/j.respol.2004.11.006

Rubin, R. (1987). *Multiple imputation for nonresponse in surveys*. (John Wiley & Sons, Ed.). New York.

Schwartz, M., & Hornych, C. (2010). Cooperation patterns of incubator firms and the impact of incubator specialization: Empirical evidence from Germany. *Technovation*, *30*(9–10), 485–495. https://doi.org/10.1016/j.technovation.2010.05.001

Sepulveda, F. (2012). The Difference Between a Business Accelerator and a Business Incubator. *Inc.*, 25–27. Retrieved from http://www.inc.com/fernando-sepulveda/the-difference-between-a-business-accelerator-and-a-business-incubator.html

Shane, S., & Khurana, R. (2003). Bringing individuals back in: the effects of career experience on new firm founding. *Industrial and Corporate Change*, *12*(3), 519–543. https://doi.org/10.1093/icc/12.3.519

Shannon, C. (1948). The mathematical theory of communication. *Bell Technical Journal*, *27*(3), 379–423.

Shepherd, D. A., & Wiklund, J. (2006). Successes and Failures at Research on Business Failure and Learning from It. *Foundations and Trends in Entrepreneurship*, *2*(5). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1629315

Slaper, T. F., & Hall, T. J. (2011). The Triple Bottom Line : What Is It and How Does It Work ? *Indiana Business Review*, *2*, 1–9.

Soetanto, D. P., & Jack, S. (2016). The impact of university-based incubation support on the innovation strategy of academic spin-offs. *Technovation*, *50–51*, 25–40. https://doi.org/10.1016/j.technovation.2015.11.001

Soetanto, D. P., & Jack, S. L. (2013). Business incubators and the networks of technology-based firms. *Journal of Technology Transfer*, *38*(4), 432–453. https://doi.org/10.1007/s10961-011-9237-4

Song, M., Podoynitsyna, K., van der Bij, H., & Halman, J. I. M. (2008). Success Factors in New Ventures. *A mEta-Analzsis*, *25*, 7–27. https://doi.org/10.1111/j.1540-5885.2007.00280.x

Spencer, A. S., & Kirchhoff, B. a. (2006). Schumpeter and new technology based firms: Towards a framework for how NTBFs cause creative destruction. *International Entrepreneurship and Management Journal*, *2*(2), 145–156. https://doi.org/10.1007/s11365-006-8681-3

Srivastava, A., & Shami, M. (2009). *Text Mining Classification, Clustering, and Applications*. *Taylor & Francis Group* (Vol. 53). https://doi.org/10.1017/CBO9781107415324.004

Steyvers, M., & Griffiths, T. (2007). Latent dirichlet allocation. In *Handbook of latent semantic analysis* (pp. 424–4402). https://doi.org/10.1016/s0364-0213(01)00040-4

Stirling, A. (2007). A General Framework for Analysing Diversity. *Royal Society Interfaces* , *44*(156), 1–50. Retrieved from https://core.ac.uk/download/files/153/7372085.pdf

Stirling, A. (2010). Multicriteria diversity analysis. A novel heuristic framework for appraising energy portfolios. *Energy Policy*, *38*(4), 1622–1634. https://doi.org/10.1016/j.enpol.2009.02.023

Su, J., & Liao, W.-P. (2013). Latent Dirichlet Allocation For Text And Image Topic Modeling. Retrieved from http://people.cs.umass.edu/~jcsu/reports/Project3.pdf

Sullivan, D. M., & Marvel, M. R. (2011). Knowledge Acquisition, Network Reliance, and Early-Stage Technology Venture Outcomes. *Journal of Management Studies*, *48*(6), 1169–1193. https://doi.org/10.1111/j.1467-6486.2010.00998.x

Tamásy, C. (2007). Rethinking technology-oriented business incubators: Developing a robust policy instrument for entrepreneurship, innovation, and regional development? *Growth and Change*, *38*(3), 460–473. https://doi.org/10.1111/j.1468-2257.2007.00379.x

Teece, D. J., Pisano, G., & Shuen, a M. Y. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, *18*(7), 509–533. Retrieved from http://www.jstor.org/stable/3088148

Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2005). Sharing clusters among related groups: Hierarchical Dirichlet processes. *Advances in Neural Information Processing Systems*, (1), 1385–1392. https://doi.org/10.1198/016214506000000302

Toft-Kehler, R., Wennberg, K., & Kim, P. H. (2014). Practice makes perfect: Entrepreneurial-experience curves and venture performance. *Journal of Business Venturing*, *29*(4), 453–470. https://doi.org/10.1016/j.jbusvent.2013.07.001

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing*, *26*(3), 341–358. https://doi.org/10.1016/j.jbusvent.2009.09.004

USPTO. (2018). Cooperative Patent Classification. Retrieved March 16, 2018, from https://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y.html#Y02

Van den Bergh, J. C. J. M. (2008). Optimal diversity: Increasing returns versus recombinant innovation. *Journal of Economic Behavior and Organization*, *68*(3–4), 565–580. https://doi.org/10.1016/j.jebo.2008.09.003

Van Rijnsoever, F. J., Van Den Berg, J., Koch, J., & Hekkert, M. P. (2015). Smart innovation policy: How network position and project composition affect the diversity of an emerging technology. *Research Policy*, *44*(5), 1094–1107. https://doi.org/10.1016/j.respol.2014.12.004

van Rijnsoever, F. J., van Weele, M. A., & Eveleens, C. P. (2016). Network brokers or hit makers? Analyzing the influence of incubation on start-up investments. *International Entrepreneurship and Management Journal*, 1–25. https://doi.org/10.1007/s11365-016-0416-5

van Weele, M., van Rijnsoever, F. J., & Nauta, F. (2017). You can't always get what you want: How entrepreneur's perceived resource needs affect the incubator's assertiveness. *Technovation*, *59*, 18–33. https://doi.org/10.1016/j.technovation.2016.08.004

Verheul, I., & Thurik, R. (2001). Start-Up Capital: "Does Gender Matter?" *Small Business Economics*, *16*, 329–345.

Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy*, *45*(3), 707–723. https://doi.org/10.1016/j.respol.2015.11.010

Verspagen, B. (2007). Mapping Technological Trajectories As Patent Citation Networks: a Study on the History of Fuel Cell Research. *Advances in Complex Systems*, *10*(1), 93–115. https://doi.org/10.1142/S0219525907000945

Witajewski-Baltvilks, J., Verdolini, E., & Tavoni, M. (2015). Bending the learning curve. *Energy Economics*, *52*, S86–S99. https://doi.org/10.1016/j.eneco.2015.09.007

Witt, P. (2004). Entrepreneurs' networks and the success of start-ups. *Entrepreneurship & Regional Development*, *16*(5), 391–412. https://doi.org/10.1080/0898562042000188423

Wright, M., & Stigliani, I. (2012). Entrepreneurship and growth. *International Small Business Journal*, *31 (1)*(9783319505466), 3–22. https://doi.org/10.1177/0266242612467359

Yu, C. F., Van Sark, W. G. J. H. M., & Alsema, E. A. (2011). Unraveling the photovoltaic technology learning curve by incorporation of input price changes and scale effects. *Renewable and Sustainable Energy Reviews*, *15*(1), 324–337. https://doi.org/10.1016/j.rser.2010.09.001

Zahra, S. a. (1996). Technology strategy and new venture performance: A study of corporate-sponsored and independent biotechnology ventures. *Journal of Business Venturing*, *11*(4), 289–321. https://doi.org/10.1016/0883-9026(95)00128-X

Zeileis, A., & Hothorn, T. (2002). Package “lmtest”; Diagnostic Checking in Regression Relationships. *R News*, *2*(3), 7–10. Retrieved from https://cran.r-project.org/web/packages/lmtest/lmtest.pdf

Zhang, N., Zhou, P., & Choi, Y. (2013). Energy efficiency, CO2 emission performance and technology gaps in fossil fuel electricity generation in Korea: A meta-frontier non-radial directional distance functionanalysis. *Energy Policy*, *56*, 653–662. https://doi.org/10.1016/j.enpol.2013.01.033

Zhao, W., Chen, J. J., Perkins, R., Liu, Z., Ge, W., Ding, Y., & Zou, W. (2015). A heuristic approach to determine an appropriate number of topics in topic modeling. *BMC Bioinformatics*, *16*(Suppl 13), S8. https://doi.org/10.1186/1471-2105-16-S13-S8

Zhong, S., & Verspagen, B. (2016). *The role of technological trajectories in catching-up-based development: An application to energy efficiency technologies*. *UNU-Merit* (Vol. 13). Retrieved from http://blackwell-synergy.com/doi/abs/10.1111/j.1467-629X.1980.tb00220.x

7. Appendix A: Operationalisation table

Table 5: Operationalisation of theoretical concepts into variables

Concept	Indicators	Calculation of scores	Measure-ment
Start-up business performance	Cumulative Investment	The absolute, cumulative amount of external investments made into the start-up between it's foundation and the most recent performance survey. 1= €0-250,000 2= €250,000-500,000 3= €500,000-1 million 4= >€1 million When including 2014 and 2015 data the measure is a binary variable: 1= €0-250,000 2= >€ 250,000	Ordinal, 1-4 AND Binary 1,2
	Firm size (employee)	The number of employees working for the start-up at the time of the performance survey.	Count, 0- ∞
	Firm size (revenues)	The absolute, cumulative amount of revenues of the company in the year of the performance survey. 0= no revenues 1= €0-10,000 2= €10,000-100,000 3= >€100,000	Ordinal, 0-3
Start-up climate impact	Potential climate impact (assessed)	A five-point scale assessment of the start-up's potential for CO_2 reductions if the start-up's product/service is successfully introduced on the market.	Ordinal
Start-up technological diversity change	The technological diversity change caused by a start-up	The difference between the entropy of the population of climate-focused start-ups (H_1) and a hypothetical population in which that particular start-up does not exist (H_0). $\Delta H = H_1 - H_0$	Ratio, -∞ - +∞
Company age	The age of the start-up	The age of the company at the moment the performance measures are evaluated, calculated as the year of the performance survey minus the start-up foundation year.	Count, 0- ∞
Initial team size	The number of initial founders	The absolute amount of founders at the time of the start-up's foundation.	Count, 0- ∞

Working experience	The amount of working experience of the founders	The cumulative amount of years of working experience of the founding members of the entrepreneurial team at the time of the start-up's foundation. A robustness check is performed with the average amount of years of working experience per team member.	Count, 0- ∞
Industry experience	The previous industry experience of the founders	A binary measure that represents whether any founder has previous working experience in an industry related to the start-up at the time of the foundation.	Binary, 0-1
Manageme nt experience	The previous management experience of the founders	A binary measure that represents whether any founder has previous experience in a management position at the time of the start-up's foundation.	Binary, 0-1
Start-up experience	The previous start-up experience of the founders	A binary measure that represents whether any founder has previous experience in founding a company. This is measured at the time of the start-up's foundation.	Binary, 0-1
Product type	The start-ups type of product	A categorical variable indicating whether a start-up has a software or a hardware product (combinations are coded as hardware). This variable is measured at the time of entering the accelerator, which is used as a proxy for the start-ups time of foundation.	Dummy variable, 0-1
Customer type	The main type of customer of the start-up	A categorical variable indicating whether a start-up's main customer is a business (B2B) or a consumer (B2C). This variable is measured at the time of entering the accelerator, which is used as a proxy for the start-ups time of foundation.	Dummy variable, 0-1
Gender	The percentage of males	The percentage of members of the initial founding team which are males.	Ratio 0-1
Specific Climate-KIC accelerator program	The specific accelerator	A categorical variable indicating the specific accelerator in which the start-up is located: • The Netherlands • The Nordics (Denmark, Norway, Finland, Sweden) • DACH (Berlin, Munich, Frankfurt, Switzerland, Austria)	Dummy variable, 0-1

8. Appendix B: Climate performance variable

During the data collection process I was able to find panel reviews for 127 out of the 197 start-ups. The panel scores are given by a group of around 65 different panel members, each assessing a different group of start-ups. From the panel members a total of 633 pairs of unique panel member/start-up climate assessments were obtained. Details about the distribution of the available data across the regions are shown in table 6.

Table 6: Overview of the available panel member review data per region.

	Netherlands	**DACH**	**Nordics**
Panel reviews available	78/112	18/47	31/38
Number of panel members	37	8	22
Number of unique panel member/start-up scores	390	104	139

The one-way, single-measure ICC is used to generalize the reliability of the scores for this original subset of start-ups (127) to the full set of 197 start-up ratings. As the panel member scores are relatively sparse data, on average there are less than 10 ratings per panel member, it is not feasible to calculate the ICC in comparison to every panel member. Therefore, a single score for the panel member group is constructed by calculating the mean. In addition, I also test the robustness of this measure by rounding these means to the nearest complete number, creating the rounded mean. These central tendencies are then used to calculate the ICC by comparing them with the author scores (Table 7).

Table 7: The ICC scores based on the initial data.

	Mean	**Rounded Mean**
ICC	0.543	0.565
N	127	127

These ICC scores can be assessed as acceptable (0.4-0.6), but they are slightly below the 0.6 threshold for good inter-reliability scores (Cicchetti, 1994; Hallgren, 2012). However, a big part of the difference between the author and expert panel scores can be found in the lower scores, which are given less frequent by the experts (Table 8 and Figure 3). In addition, the mean score also results in a smaller deviation of the scores and more scores in the middle of the scale. Thus indicating that the reliability of the author scores is potentially higher.

Table 8:Table with the author and panel member scores

	1	**2**	**3**	**4**	**5**
Author CO_2e score	13	24	52	33	5
Rounded mean CO_2e score of panel members	4	18	64	38	3

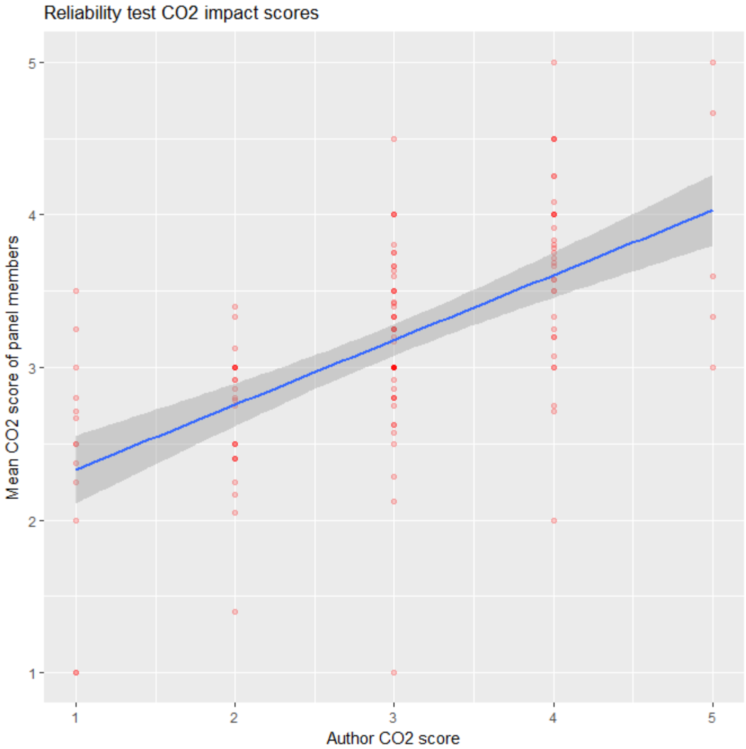


Figure 3: Scatterplot of the author and panel member mean climate performance scores

As described in the main text there are two particular concerns when taking a closer look at the data. First, there are a number of panel members who only provided climate impact evaluations for a single start-up. And second, there is a group of panel members whose assessments show very poor correlations (below 0.2) with the mean score of the panel members. When these two groups of panel members are removed from consideration, data on about 50 panel members who have scored 122 start-ups for a total of 509 pairs of unique panel member/start-up climate assessments remains. The resulting ICC values, which are reported in the main text, for are now above 0.6 showing a good inter-reliability-rating and as such proving that the climate impact assessment of the author is a reliable measure (Table 9).

Table 9: The ICC scores based on the cleaned data.

	Rounded Mean	**Rounded Mean**
ICC	0.627	0.681
N	122	122

Finally, to test the robustness of the variable I also calculated the ICC for each of the respective accelerator regions separately. The ICC scores prove to be robust between the different accelerator regions as the ICC scores are similar across the three regions (Table 10).

Table 10: The ICC scores based on the cleaned data.

	Netherlands Mean	**Netherlands Rounded Mean**	**DACH Mean**	**DACH Rounded Mean**	**Nordics Mean**	**Nordics Rounded Mean**
ICC	0.572	0.663	0.700	0.751	0.662	0.664
N	78	78	15	15	29	29

9. Appendix C: Topic modelling results

The LDA uses a corpus (a collection of documents) as it's primary input. The corpus is used to build a term-document matrix. This matrix contains all the words in the corpus as the rows and the separate documents as the columns. It is subsequently filled out with the frequency that each term occurs in the corresponding document. The pre-processing steps necessary to fit the LDA on the corpus are shown in table 2 (Blei & Lafferty, 2009; Grün & Hornik, 2011).

Table 11: Data pre-processing steps

Pre-processing step	Argumentation
Removing all punctuation characters, numbers and non-Latin characters	'R' cannot process these characters
Removing stop words and words of less than three characters	To prevent these frequent but meaningless words from influencing the topics
Converting all characters to lower-case and stemming the words	To homogenise different forms of the same words
Pruning the vocabulary	Words which occur in very little documents or in nearly every document (e.g. non-discriminating terms) can seriously skew the topics

The pruning of the vocabulary is done using the term-frequency inverse document scores (TF-IDF). The TF-IDF is a measure of how often a word occurs in a document and weighs this in comparison to the number of documents in which the word occurs (Robertson, 2004). The median TF-IDF score is slightly above 0.05 and I thus use 0.05 as the cut-off value to include terms into the analyses. Before the pre-processing steps the technology descriptions consisted of an average of 203 words per description, after the pre-processing process there are 47 words per document remaining. The term-document matrix which forms the input to the LDA now consists of 2991 rows (words) and 303 documents (start-up technologies.

As described in the operationalization of this thesis, the α has to be determined before performing the LDA. In the VEM algorithm α can be estimated or set to a fixed value of 50/k (Grün & Hornik, 2011). The choice between an estimated and fixed value for α is made by comparing the perplexities through 10-fold cross validation (Grün & Hornik, 2011; Su & Liao, 2013; Teh et al., 2005). This is done for multiple topics with an increment of five to keep computing power down. Table 12 shows that the perplexities of the estimated α minus the fixed α models. The results show that the perplexities of the fixed α model are consistently lower than for the estimated models and therefore the LDA is estimated with a fixed value for α.

*Table 12: The result of the estimated α **minus** the fixed α perplexities based on 303 accepted start-ups from the Netherlands, Nordics, and DACH regions.*

Topics	*Fold 1*	*Fold 2*	*Fold 3*	*Fold 4*	*Fold 5*	*Fold 6*	*Fold 7*	*Fold 8*	*Fold 9*	*Fold 10*
5	5,0E+00	5,3E+09	3,5E+11	2,5E+11	2,6E+11	2,3E+18	5,8E+17	1,6E+18	5,4E+10	1,5E+18
10	1,0E+01	7,2E+09	4,1E+11	3,1E+11	4,4E+11	3,0E+18	1,0E+18	2,4E+18	5,5E+10	2,8E+18
15	1,5E+01	9,1E+09	5,4E+11	5,5E+11	4,8E+11	2,9E+18	6,4E+17	2,5E+18	7,8E+10	2,9E+18
20	2,0E+01	7,6E+09	4,2E+11	6,2E+11	2,3E+11	2,5E+18	7,3E+17	3,0E+18	6,7E+10	2,1E+18
25	2,5E+01	4,1E+09	5,0E+11	5,8E+11	4,2E+11	2,4E+18	7,8E+17	2,5E+18	7,9E+10	3,1E+18
30	3,0E+01	5,4E+09	3,4E+11	4,9E+11	3,2E+11	3,2E+18	1,0E+18	1,7E+18	6,7E+10	2,9E+18

35	3,5E+01	5,7E+09	3,2E+11	3,5E+11	4,5E+11	2,7E+18	1,1E+18	2,0E+18	6,2E+10	2,5E+18
40	4,0E+01	6,1E+09	3,3E+11	5,5E+11	3,3E+11	2,9E+18	8,6E+17	1,6E+18	6,3E+10	3,4E+18
45	4,5E+01	4,4E+09	3,7E+11	4,7E+11	2,8E+11	2,5E+18	7,2E+17	1,7E+18	7,1E+10	1,7E+18
50	5,0E+01	5,2E+09	3,4E+11	3,3E+11	2,0E+11	1,8E+18	6,5E+17	1,4E+18	7,1E+10	3,2E+18

The RPC graph for the fixed α VEM models is shown in Figure 4 and the first number of topics whose RPC is smaller than the following number of topics is 15 topics, which is the most appropriate amount of topics (Zhao et al., 2015).

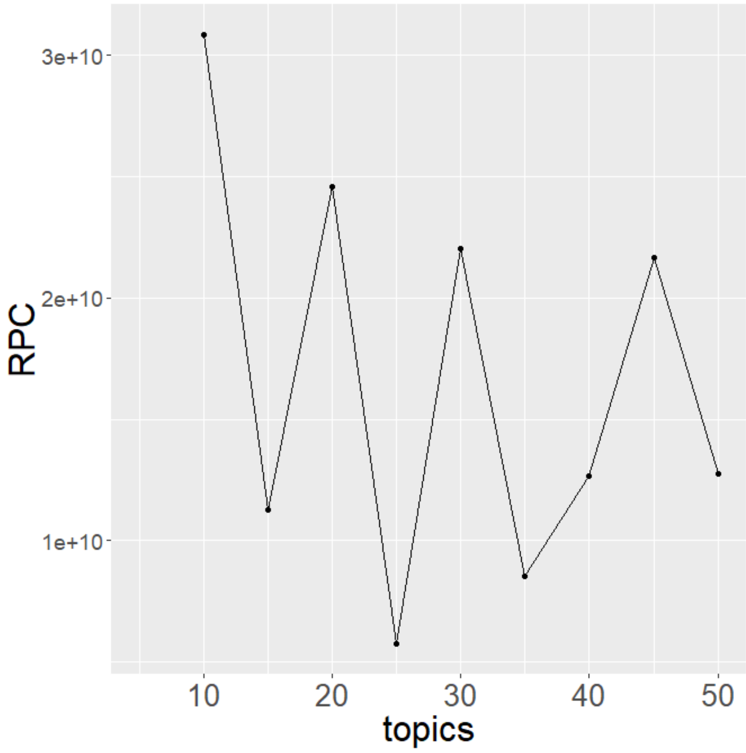


Figure 4: The RPC graph for the fixed α VEM models (15 topics).

The ten most frequent word for each of the 15 latent topics, which are the result of the LDA, are shown in table 13.

Table 13: The ten most frequent words for the fixed α LDA with 15 latent topics based on the technology descriptions of the 303 accepted start-ups.

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
solar	food	pallet	plant	charg
cell	farmer	fibr	light	vehicl
renew	crop	techniqu	pump	grid
water	chemic	databas	modul	station
flexibl	water	atmospher	laser	box
revers	fertil	print	deliveri	car
simul	agricultur	press	hydropon	street
modul	soil	transact	instagreen	meter
salt	farm	forecast	optic	raft
osmosi	yield	printer	aqysta	driver

Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
clean	cool	packag	wind	mobil
water	panel	good	air	vehicl
ship	tank	sensor	turbin	weight
fuel	thermal	organ	car	smartphon
contain	water	hydrogen	pressur	pile
robot	solar	footprint	compress	engag
captur	cycl	app	zigzagsolar	motiontag
surfac	air	pilot	yield	ducki
cleaner	investor	rotor	drive	social
load	micro	plant	facad	composit

Topic 11	Topic 12	Topic 13	Topic 14	Topic 15
greenhous	heat	water	water	wast
solar	wast	trailer	measur	recycl
output	wood	shower	sensor	batteri
treatment	dri	bike	imag	marketplac
ufb	fuel	circul	oil	layer
imageri	hous	offgrid	plastic	concret
date	glass	rider	absorb	onlin
waterbox	modheat	flexibl	cooper	applianc
cover	forest	financi	tent	home
repair	combust	advertis	filter	cybe

Figure 5 displays the highest probability with which each document belongs to a specific topic. The fact that each document has a distribution over multiple topics is well suited for the complex nature of technologies and thus an advantage of the LDA approach (Blei, 2011). The probabilities for the documents to belong to a specific topic ranges between 0.0155 and 0.767.

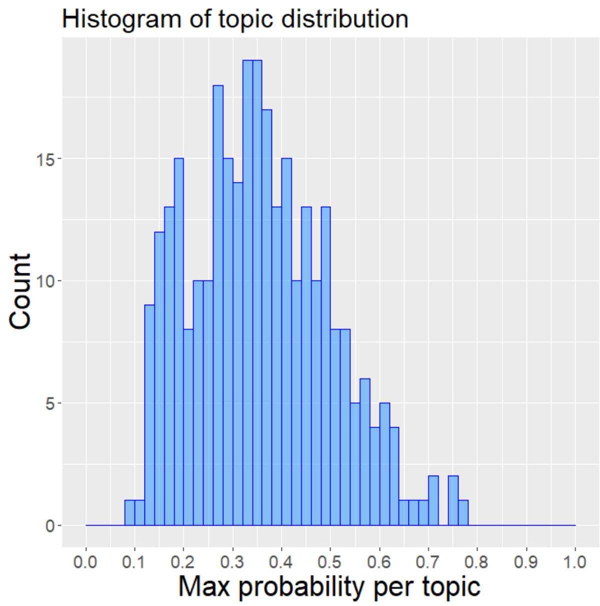


Figure 5: The highest probability for each document to belong to a topic.

The technologies that exhibit similar distribution over the topics are then clustered together (Aggarwal & Zhai, 2013). To do so clustering models with 5 to 100 clusters are fitted and the variance explained by the cluster models are used to choose the appropriate number of clusters. The most appropriate number of clusters is at the 'elbow' of the plot of the total within-clusters sum of squares and the number of clusters (Annand, 2017). This is the case for 20 clusters (Figure 6).

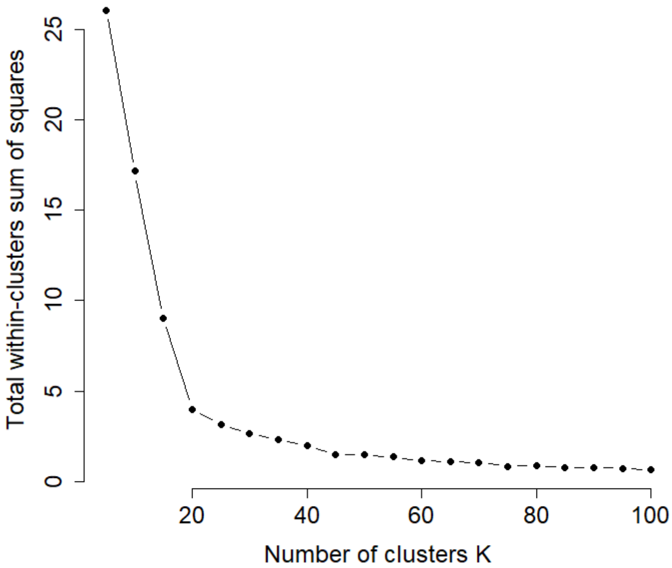


Figure 6: Plot to determine the optimal number of clusters

The clusters are then used to calculate the diversity creation for each start-up through the aforementioned Shannon-Weaver entropy index (Eq.1). The resulting histogram of the technological diversity creation is shown in figure 1 of the main thesis. For the analyses, thee diversity values are multiplied by 1000 to produce more normal coefficients (Figure 7).

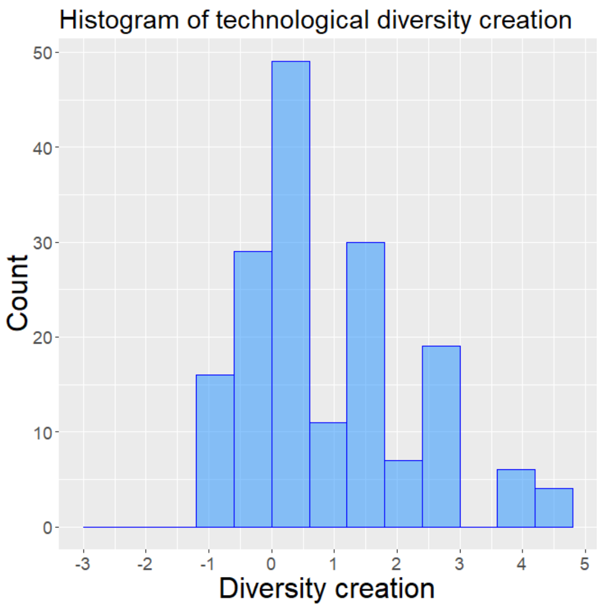


Figure 7: Histogram of the technological diversity change values of the start-ups.

The graphs for the two robustness checks are also displayed below. First, figure 8a-c shows the RPC, cluster graph and the histogram of the resulting variable when the Gibbs sampling algorithm is used instead of the VEM algorithm. Second, figure 9a-c show the same three graphs in the case where the technological diversity change variable is constructed based on the 197 start-ups included in this research. In both cases the results of the regression models proved to be robust (these are shown in Appendix E).

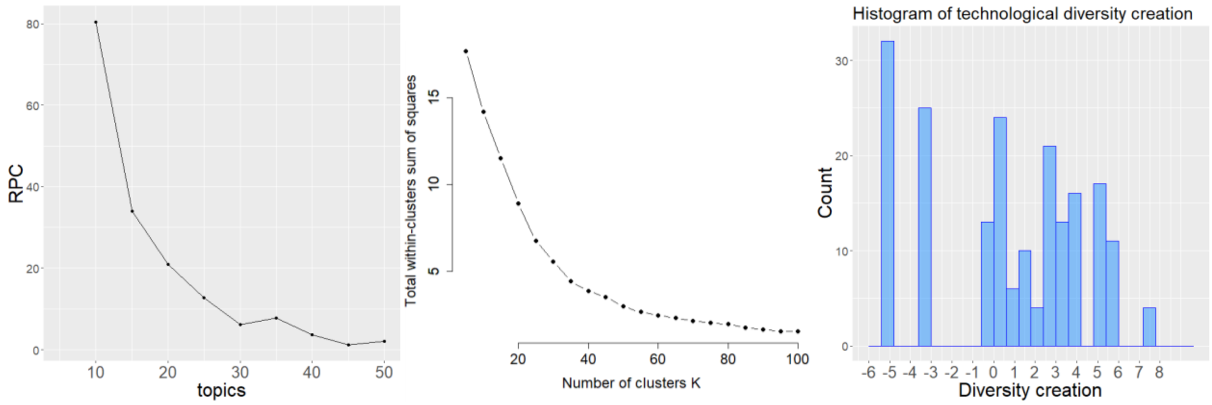


Figure 8a-c: The RPC graph (30 topics), cluster graph (25 clusters) and the histogram for the Gibbs sampling models (30 topics).

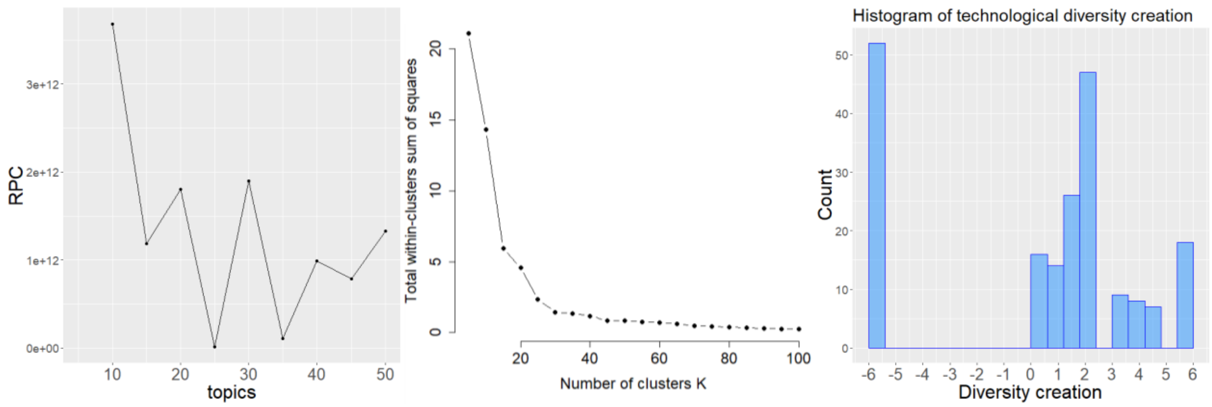
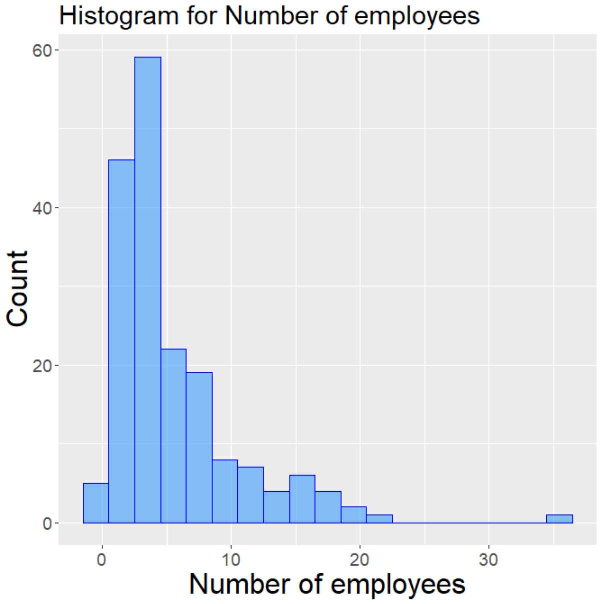


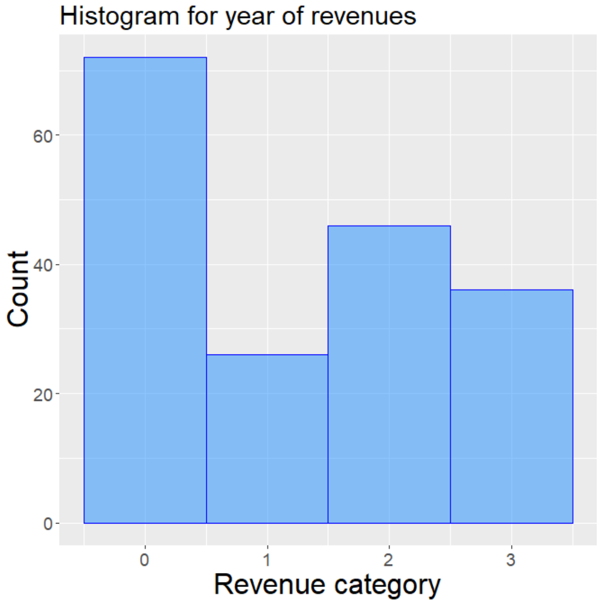
Figure 9a-c: The RPC graph (15 topics), cluster graph (15 clusters) and the histogram for the fixed α VEM model based on 197 start-ups

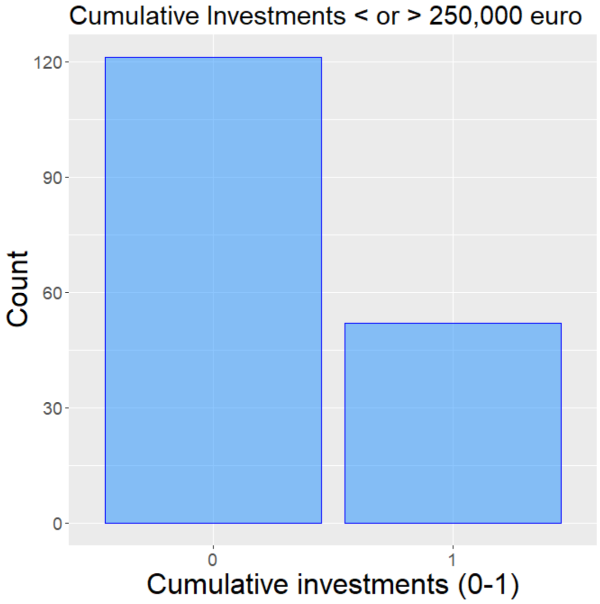
10. Appendix D: Descriptive statistics

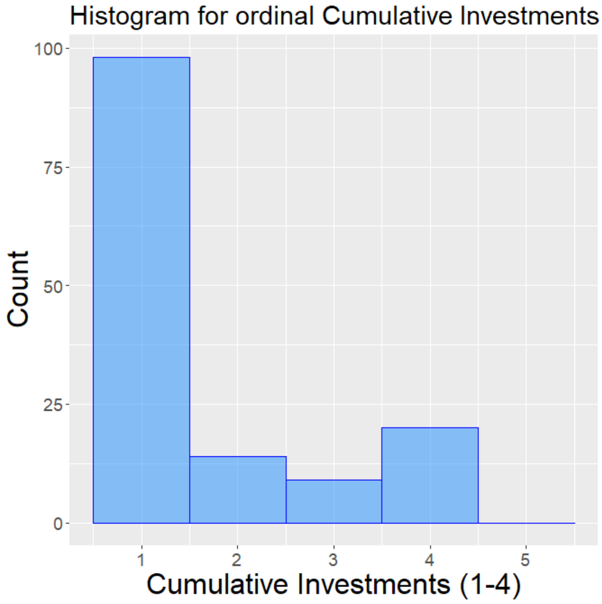
This appendix contains the histograms and bar diagrams for all the variables included in this research. In addition, for all dependent variables the scatterplots between the respective dependent variable and the independent variable, as well as the significant control variables are displayed and discussed.

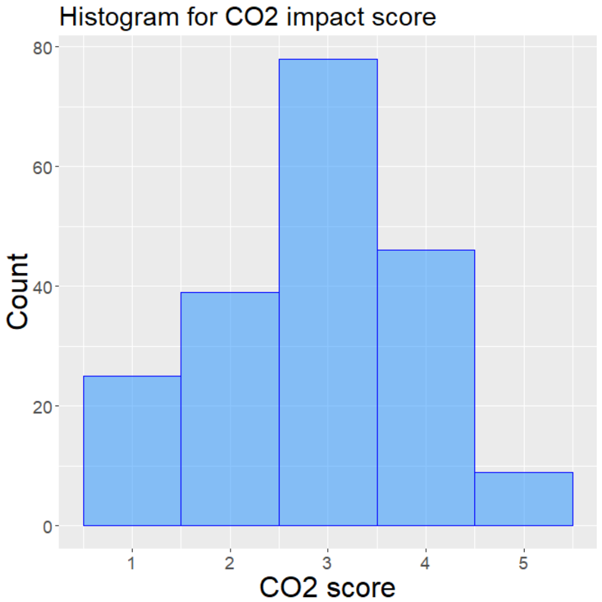
Histograms and bar diagrams

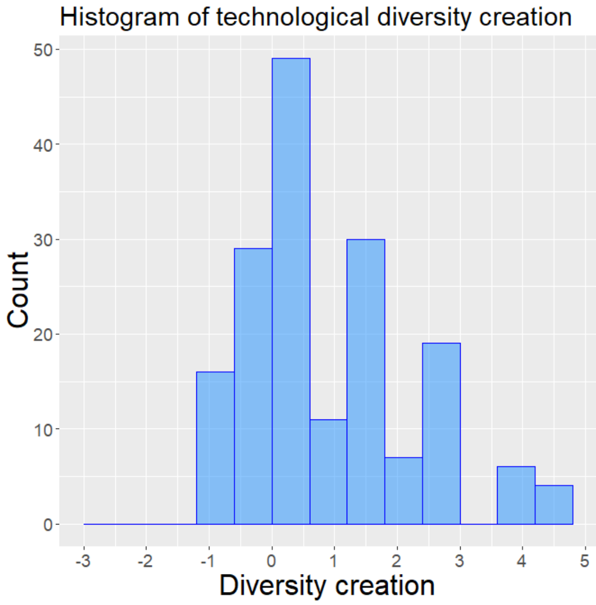


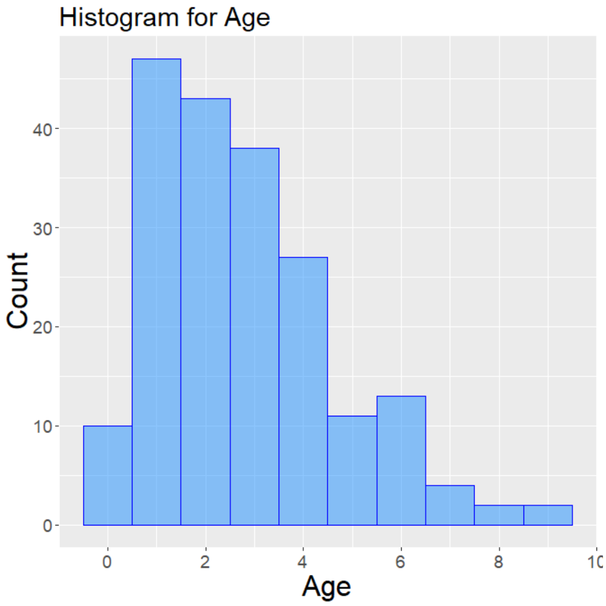


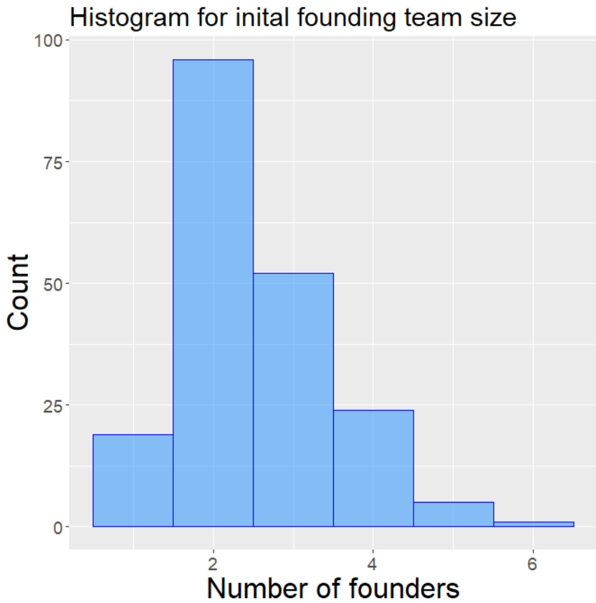


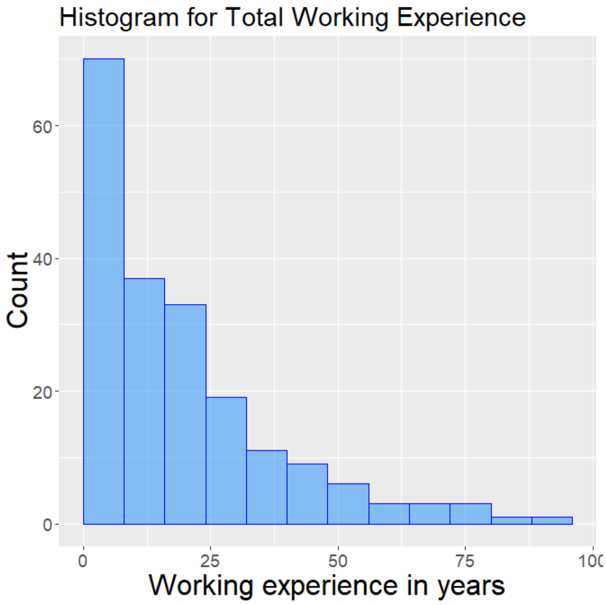


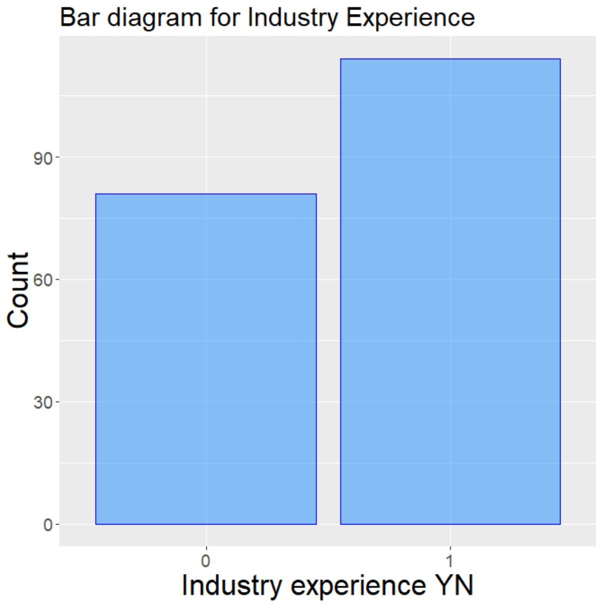

Histogram for CO2 impact score
Count
CO2 score
80
60
40
20
0
1
2
3
4
5

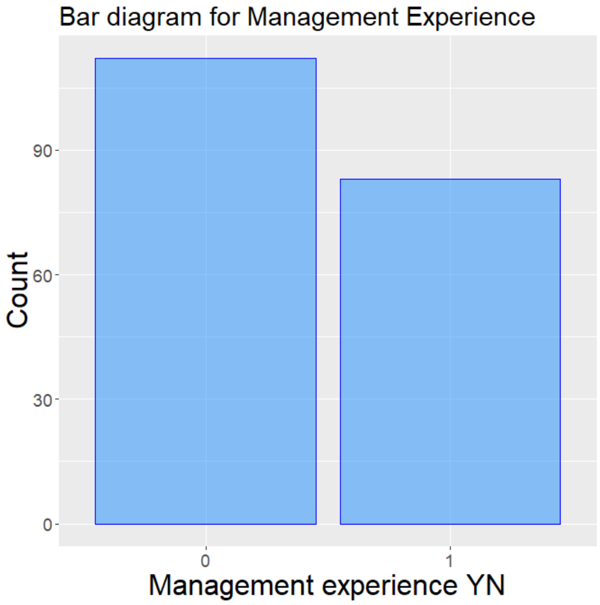

Histogram of technological diversity creation
Count
Diversity creation
50
40
30
20
10
0
-3
-2
-1
0
1
2
3
4
5

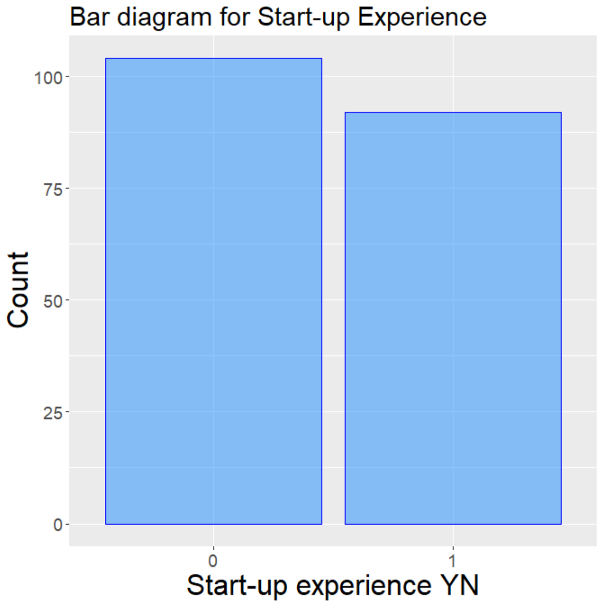

Histogram for Age
Count
Age
40
30
20
10
0
0
2
4
6
8

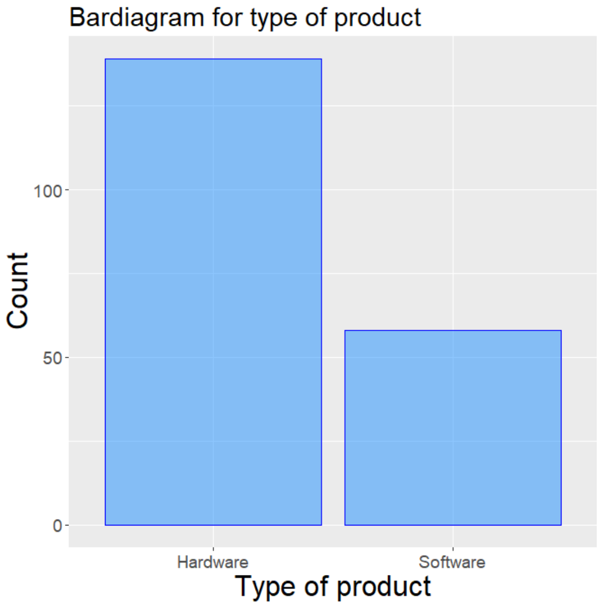

Histogram for inital founding team size
Count
Number of founders
100
75
50
25
0
2
4
6

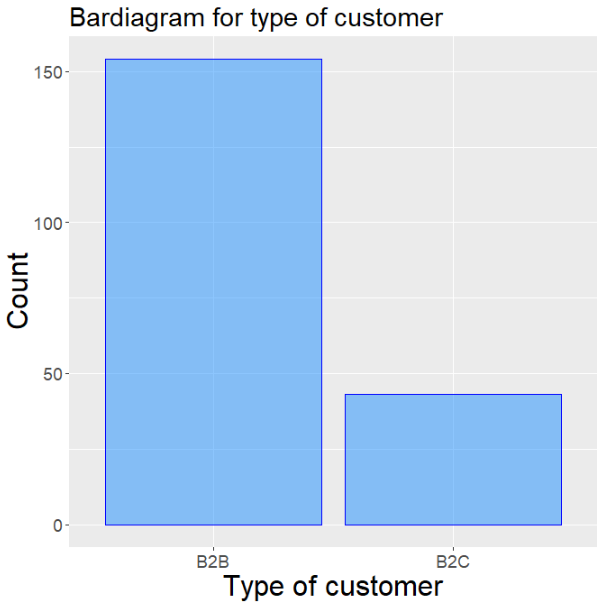

Histogram for Total Working Experience
Count
Working experience in years
60
40
20
0
0
25
50
75

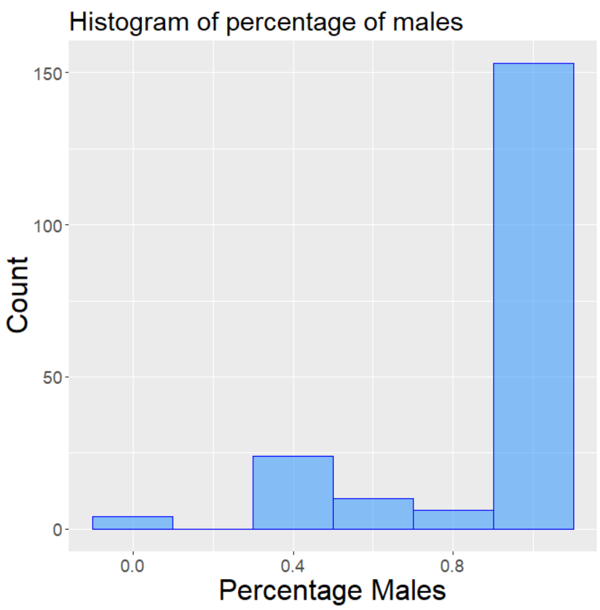

Bar diagram for Industry Experience
Count
Industry experience YN
90
60
30
0
0
1

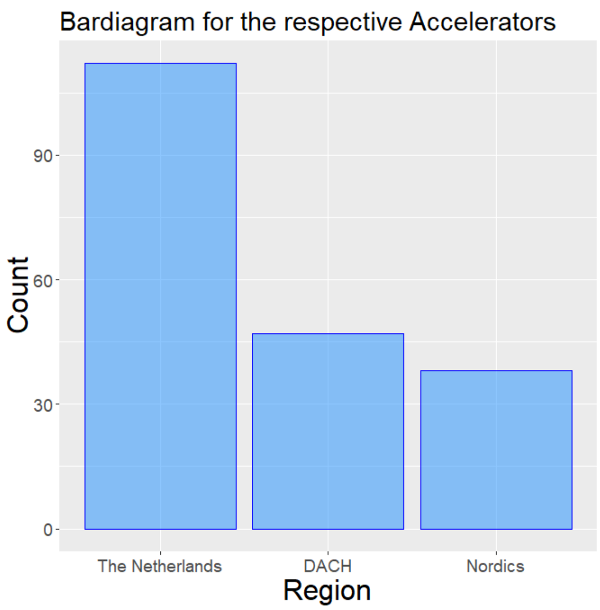

Bar diagram for Management Experience
Count
0
30
60
90
0
1
Management experience YN


Bar diagram for Start-up Experience
Count
0
25
50
75
100
0
1
Start-up experience YN

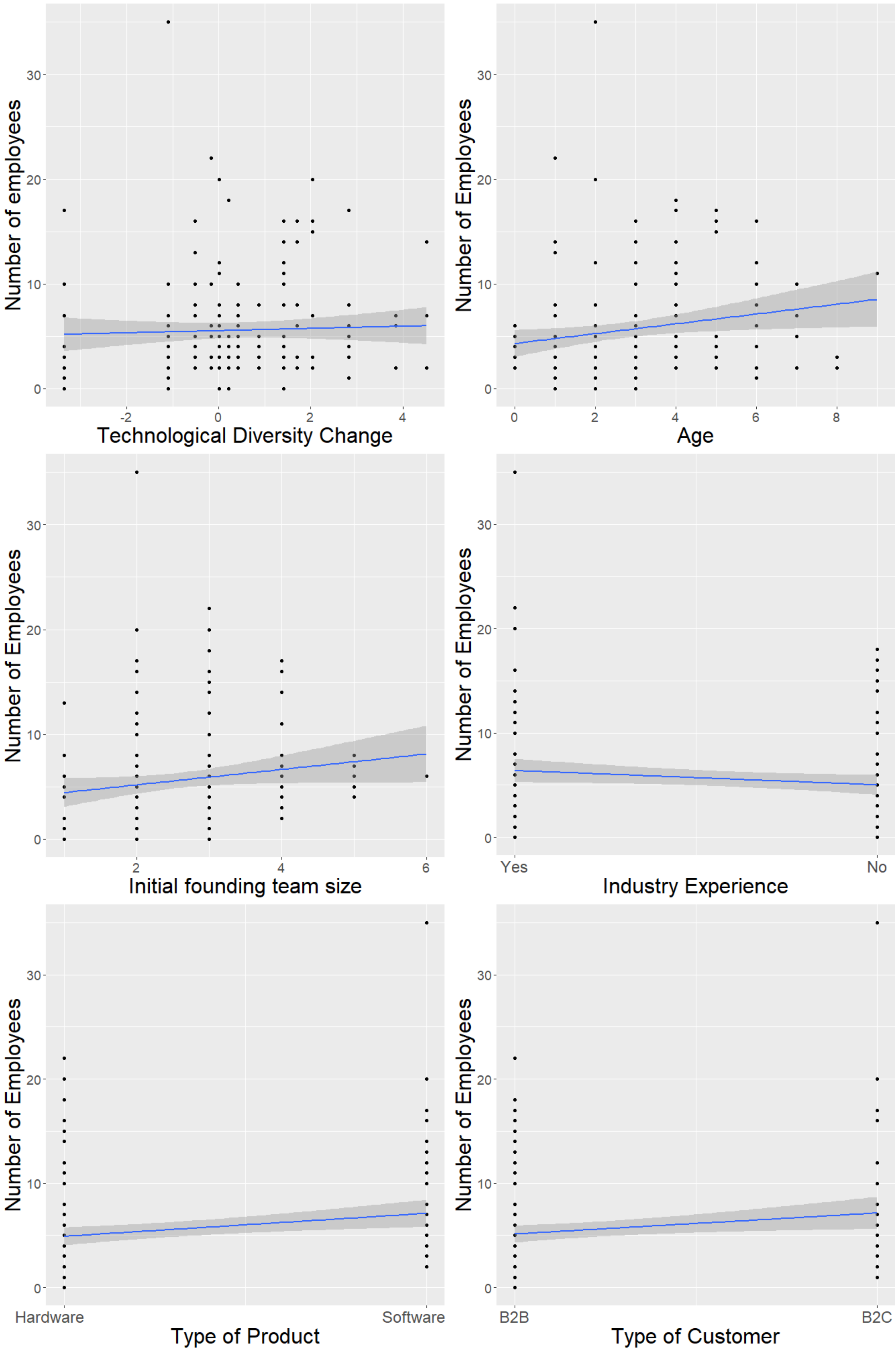

Bardiagram for type of product
Count
0
50
100
Hardware
Software
Type of product

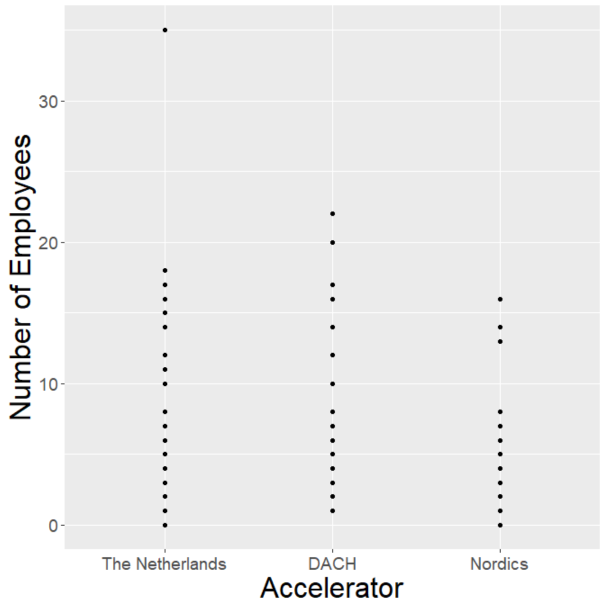

Bardiagram for type of customer
Count
0
50
100
150
B2B
B2C
Type of customer


Histogram of percentage of males
Count
0
50
100
150
0.0
0.4
0.8
Percentage Males


Bardiagram for the respective Accelerators
Count
0
30
60
90
The Netherlands
DACH
Nordics
Region

Scatterplots with the number of employees as the dependent variable

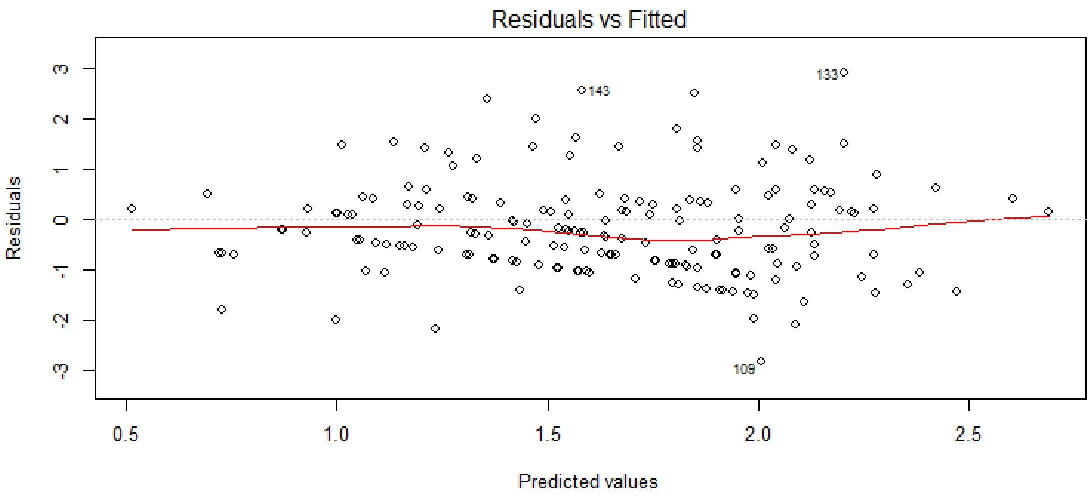


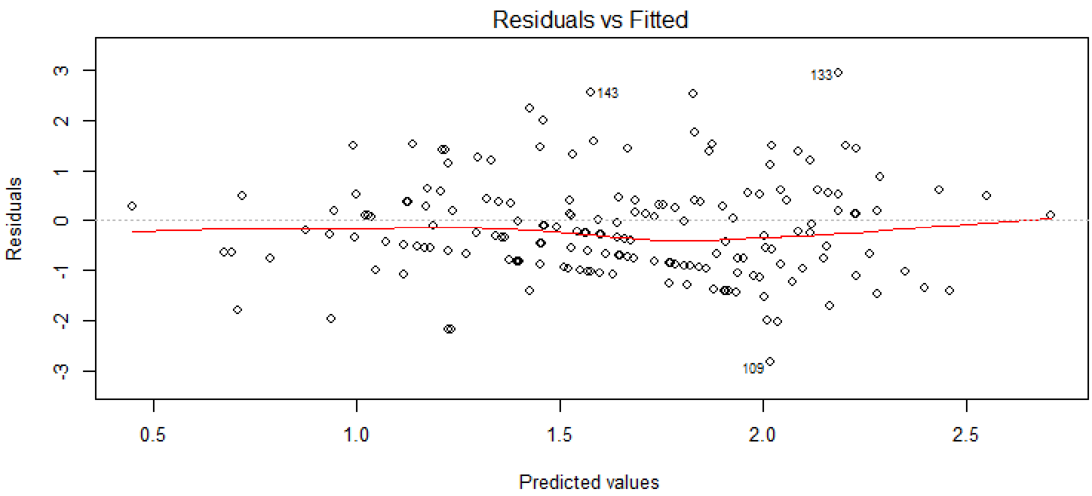


The scatterplot with the independent variable, technological diversity change, and the number of employees is homoscedastic. The scatterplots with the number of employees and the significant control variables are also homoscedastic. The wider which can be observed range for the high values of age and founding team size is due to the smaller number of observations. Also notable is that there is one firm for which the number of employees (35) is clearly higher than the next largest firm. I therefore also ran the regression analyses when recoding this value at the next highest number of employees. The results proved to be fully robust.

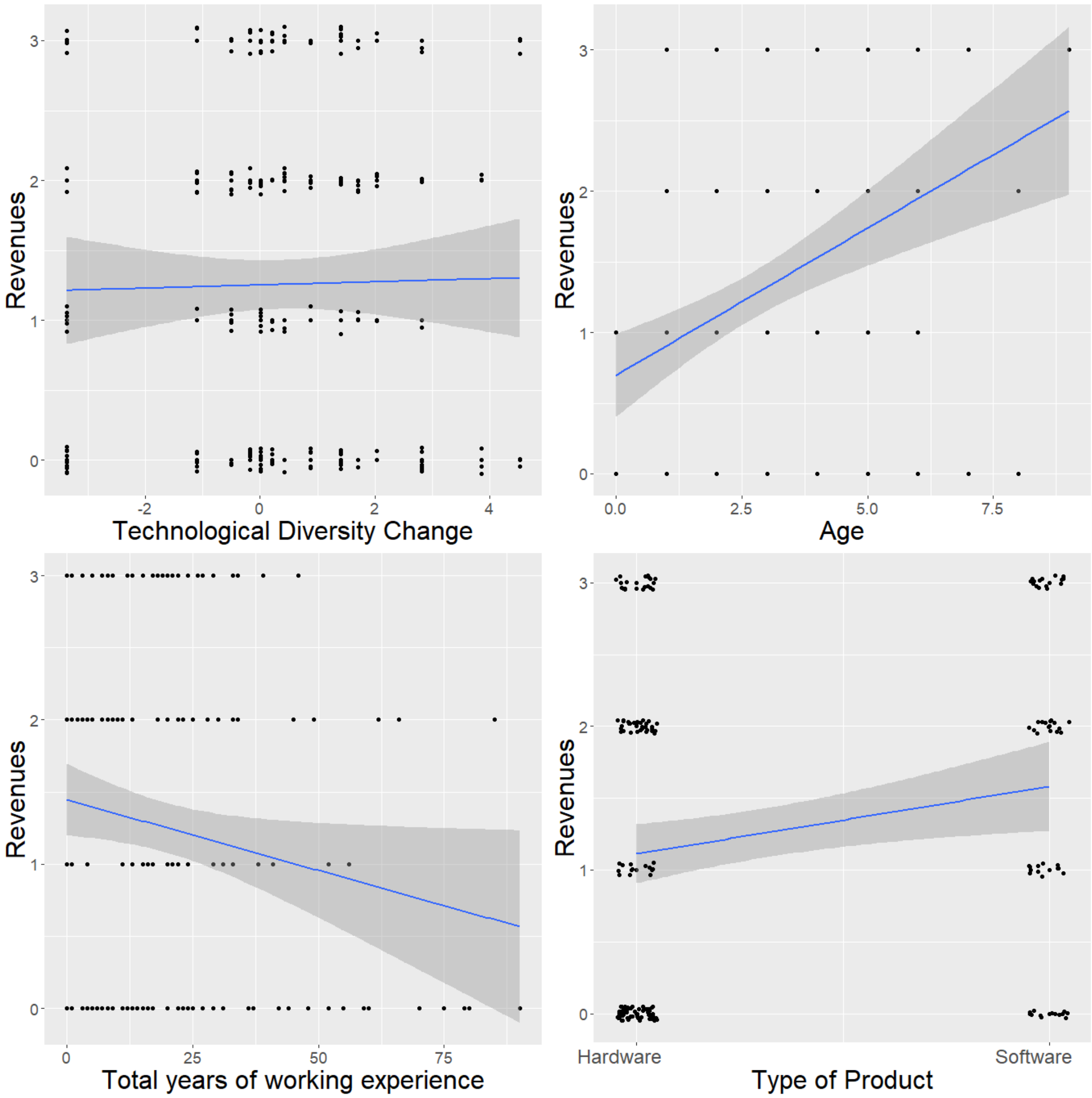
Residual plots for the model with the number of employees

The residual plots for the negative binomial control and full model are shown below and both plots show that the residuals are distributed rather evenly and as such the residuals are homoscedastic.



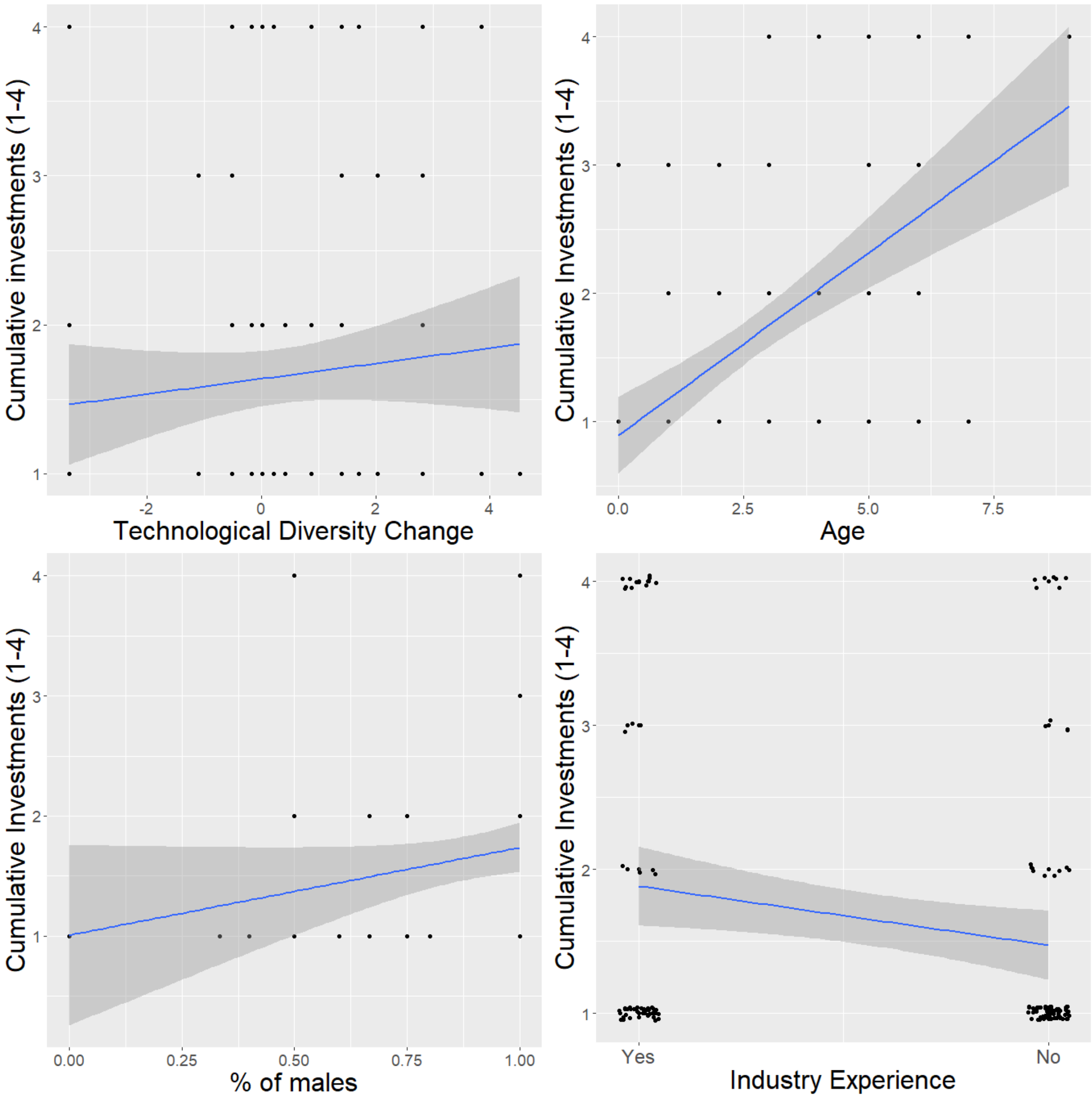


Scatterplots with revenues as the dependent variable



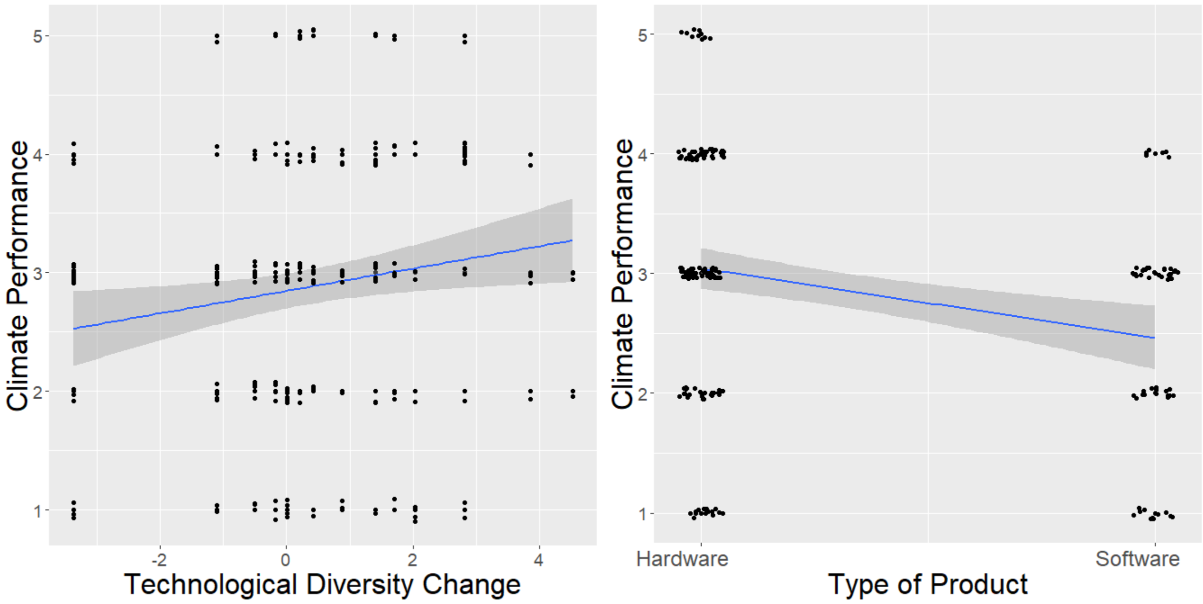
The scatterplots between the independent, technological diversity change, variable and the revenues show that, although the distibution is relatively homoscedastic, there is a broader range at the lower and higher end of the diversity values. This is partly due to the smaller number of observations and as such not evaluated as a particular concern. Regarding the significant control variables the smaller number of observations at the high values leads to a broader range for age and particularly for the total years of working experience. In the case of the total years of working experience a robustness check is therefore performed using the average years of working experience per team member. This test shows that the results remain robust. In the scatterplots with a binary control variable and an ordinal dependent variable I added some jitter to the plot to more clearly show the relationship for these variables.

Scatterplots with cumulative investments as the dependent variable



Similar to the revenues the scatterplot between technological diversity change and the cumulative investments shows a wider range at the low and high end, which is largely caused by the smaller number of observations. The broader range at the high end of age and at the low end of the percentage of males are similarly caused by a smaller number of obserations. Especially for the % of males there is some heteroskedasicity, which is the result of the high number of all male and the small number of all female start-ups. However, as this is not the main variable of this research it is not considered problematic for the results of the analyses.

Scatterplots with Climate Performance as the dependent variable



The scatterplots for the independent and the significant control variable with climate performance as the dependent variable are homoscedastic and as such do not show particular concerns.

11. Appendix E: Verifying assumptions and Robustness models

Verifying assumptions

Table 14: The overdispersion of the Poisson models with the number of employees.

	Overdispersion	P Chi-Square
Control model	3.16	1
Full model	3.18	1

Table 15: The VIF scores and maximum Cook's distance for the negative binomial model.

	VIF
Technological Diversity	1.02
Age	1.07
Initial founding team size	1.13
Total years of working experience	1.54
Industry Experience	1.15
Management Experience	1.37
Start-up Experience	1.18
Hardware-Software	1.04
B2B-B2C	1.04
Percentage of males	1.03
Accelerator	1.10
Maximum Cook's distance	0.15

Table 16a-c: The results of the test of parallel lines for the three OLM models

Test of Parallel Lines for the revenues model

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	429.364			
General	410.098	19.266	20	.505

Test of Parallel Lines for the investments model

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	222.882			
General	206.029	16.854	20	.662

Test of Parallel Lines for the climate model

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	495.278			
General	451.919	43.359	30	.054

Robustness test with the binary investment measure

The binary investment variable contains more observations (before imputation n=174) than the ordinal variable (before imputation n=142). I therefore perform use the binary investment model to test the robustness of the investments variable, to do so I perform a BLM. The McFadden value for both BLM's is 0.24. The difference between the control and full model is not significant. *The negative effect of technological diversity creation variable is not significant in the BLM for investments (5% level).* Of the control variables Start-up *age* has a positive effect on the binary cumulative investments at the 0.1% level. Furthermore, in both binary models, start-ups with a higher percentage of male founders have gathered significantly more investments (5% level).

Table 17: Results of the Binary Logit Models with survival as the dependent variable.

	Control Model	Full Model
Intercept	-5,629 *(0.000)****	-5,671 *(0.000)****
Technological Diversity	-	0,102 *(0.373)*
Age	0,628 *(0.000)****	0,629 *(0.000)****
Initial founding team size	0,284 *(0.254)*	0,263 *(0.298)*
Total years of working experience	0,000 *(0.976)*	0,002 *(0.885)*
Industry Experience	-0,730 *(0.145)*	-0,718 *(0.154)*
Management Experience	-0,404 *(0.486)*	-0,439 *(0.455)*
Start-up Experience	0,911 *(0.055)*	0,884 *(0.064)*
Hardware-Software	-0,515 *(0.256)*	-0,492 *(0.279)*
B2B-B2C	-0,031 *(0.948)*	0,011 *(0.981)*
Percentage of males	2,670 *(0.021)**	2,689 *(0.021)**
acceleratorDACH	0,704 *(0.214)*	0,707 *(0.212)*
acceleratorNordics	-0,822 *(0.167)*	-0,794 *(0.189)*
N	184	184
McFadden R^2	0.240	0.240

*Significance codes: '***' $p < 0.001$, '**' $p < 0.01$, '*' $p < 0.05$*

Robustness test using Gibbs sampling to calculate the independent variable

Table 3: Results of the regression models when Gibbs sampling is used to calculate the independent variable.

	Control employee	Full employee	Control revenues	Full revenues	Control invest	Full invest	Control climate	Full climate
Intercept	0.250 (0.381)	0.234 *(0.413)*	-	-	-	-	-	-
Technological Diversity	-	0.008 *(0.574)*	-	-0.010 *(0.805)*	-	0.003 *(0.960)*	-	0.097 *(0.007)***
Age	0.123 *(0.000) ****	0.122 *(0.000) ****	0.470 *(0.000) ****	0.470 *(0.000) ****	0.514 *(0.000) ****	0.513 *(0.000) ****	-0.003 *(0.965)*	-0.022 *(0.771)*
Initial founding team size	0.200 *(0.001) ****	0.205 *(0.001) ****	-0.065 *(0.688)*	-0.068 *(0.674)*	0.252 *(0.317)*	0.253 *(0.258)*	-0.224 *(0.158)*	-0.196 *(0.218)*
Total years of working experience	-0.006 *(0.138)*	-0.007 *(0.118)*	-0.026 *(0.035)**	-0.025 *(0.036)**	-0.001 *(0.952)*	-0.001 *(0.949)*	-0.016 (0.148)	-0.018 *(0.123)*
Industry Experience	-0.239 *(0.048)**	-0.239 *(0.048)*	0.379 *(0.258)*	0.380 *(0.258)*	-1.024 *(0.029)**	-1.025 *(0.029)**	0.452 *(0.152)*	0.454 *(0.148)*
Management Experience	0.226 *(0.114)*	0.226 *(0.115)*	0.172 *(0.656)*	0.174 *(0.652)*	-0.105 *(0.848)*	-0.105 *(0.847)*	0.542 *(0.135)*	0.584 *(0.112)*
Start-up Experience	0.117 *(0.337)*	0.123 *(0.311)*	0.197 *(0.547)*	0.193 *(0.557)*	0.502 *(0.239)*	0.505 *(0.238)*	0.329 *(0.288)*	0.370 *(0.236)*
Hardware-Software	0.307 *(0.008) ***	0.305 *(0.008) ***	0.789 *(0.011)**	0.789 *(0.011)**	-0.390 *(0.349)*	-0.390 *(0.349)*	-1.169 *(0.000) ****	-1.248 *(0.000)****
B2B-B2C	0.249 *(0.048)**	0.248 *(0.049)**	0.600 *(0.086)*	0.599 *(0.085)*	-0.133 *(0.765)*	-0.133 *(0.765)*	-0.170 *(0.592)*	-0.183 *(0.562)*
Percentage of males	0.520 *(0.033)**	0.523 *(0.033)**	1.012 *(0.138)*	1.019 *(0.135)*	2.412 *(0.034)**	2.414 *(0.034)**	0.192 *(0.729)*	0.086 *(0.879)*
Accelerator DACH	0.341 *(0.013)**	0.352 *(0.010)**	0.500 *(0.184)*	0.494 *(0.191)*	0.565 *(0.270)*	0.569 *(0.273)*	0.699 *(0.045)**	0.732 (0.036)*
Accelerator Nordics	-0.189 (0.206)	-0.185 *(0.216)*	-0.177 *(0.661)*	-0.190 *(0.641)*	-0.816 *(0.136)*	-0.813 *(0.140)*	0.697 *(0.071)*	0.729 *(0.060)*
N	184	184	184	184	184	184	197	197
McFadden R^2	0.058	0.059	0.099	0.099	0.141	0.142	0.048	0.060

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

Robustness test using 197 start-ups as the technological field to calculate the independent variable

Table 3: Results of the regression models when the technological field is determined using only the 197 start-ups for which performance data is available.

	Control employee	Full employee	Control revenues	Full revenues	Control invest	Full invest	Control climate	Full climate
Intercept	0.248 (0.386)	0.246 *(0.389)*	-	-	-	-	-	-
Technological Diversity	-	-0.005 *(0.746)*	-	-0.012 *(0.767)*	-	-0.018 *(0.726)*	-	0.074 *(0.039)**
Age	0.124 *(0.000)****	0.124 *(0.000)****	0.477 *(0.000)****	0.481 *(0.000)****	0.481 *(0.000)****	0.483 *(0.000)****	-0.004 *(0.960)*	-0.016 *(0.835)*
Initial founding team size	0.202 *(0.001)****	0.202 *(0.001)****	-0.073 *(0.653)*	-0.073 *(0.652)*	0.204 *(0.386)*	0.205 *(0.386)*	-0.228 *(0.151)*	-0.235 *(0.136)*
Total years of working experience	-0.006 *(0.127)*	-0.007 *(0.118)*	-0.026 *(0.029)**	-0.027 *(0.028)**	0.003 *(0.834)*	0.003 *(0.862)*	-0.016 (0.165)	-0.015 *(0.198)*
Industry Experience	-0.233 *(0.052)*	-0.236 *(0.052)*	0.428 *(0.199)*	0.420 *(0.209)*	-0.956 *(0.032)**	-0.975 *(0.032)**	0.462 *(0.142)*	0.539 *(0.090)*
Management Experience	0.224 *(0.117)*	0.226 *(0.114)*	0.145 *(0.708)*	0.151 *(0.696)*	-0.275 *(0.604)*	-0.2251 *(0.641)*	0.526 *(0.149)*	0.490 *(0.179)*
Start-up Experience	0.116 *(0.339)*	0.121 *(0.323)*	0.194 *(0.554)*	0.206 *(0.533)*	0.396 *(0.345)*	0.408 *(0.334)*	0.324 *(0.292)*	0.273 *(0.376)*
Hardware-Software	0.307 *(0.008)***	0.297 *(0.011) **	0.772 *(0.013)**	0.752 *(0.018)**	-0.438 *(0.287)*	-0.471 *(0.271)*	-1.169 *(0.000)****	-1.070 *(0.001)****
B2B-B2C	0.251 *(0.047)**	0.249 *(0.049)**	0.590 *(0.091)*	0.593 *(0.090)*	-0.191 *(0.675)*	-0.193 *(0.671)*	-0.165 *(0.602)*	-0.201 *(0.529)*
Percentage of males	0.519 *(0.034)**	0.525 *(0.032)**	1.009 *(0.140)*	1.034 *(0.133)*	1.906 *(0.096)*	1.916 *(0.093)*	0.198 *(0.724)*	0.027 *(0.962)*
Accelerator DACH	0.342 *(0.012)**	0.343 *(0.012)**	0.490 *(0.193)*	0.501 *(0.185)*	0.538 *(0.298)*	0.547 *(0.288)*	0.697 *(0.046)**	0.656 (0.061)
Accelerator Nordics	-0.186 *(0.211)*	-0.186 *(0.213)*	-0.172 *(0.671)*	-0.173 *(0.670)*	-0.739 *(0.176)*	-0.751 *(0.174)*	0.694 *(0.071)*	0.674 *(0.081)*
N	184	184	184	184	184	184	197	197
McFadden R^2	0.058	0.058	0.102	0.102	0.130	0.131	0.047	0.055

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

12. Appendix G: Cross table of revenues and climate potential

Table 8: Cross table of the start-up's revenues and their potential to reduce CO_2 emissions.

	€0	**€0-10,000**	**€10,000-100,000**	**€100,000-1 million**
CO2 Impact 1	10	7	4	4
CO2 Impact 2	17	4	11	7
CO2 Impact 3	31	12	19	15
CO2 Impact 4	19	3	11	10
CO2 Impact 5	7	0	1	1