Who will survive in Stockholm Archipelago?

A longitudinal analysis of firm-survival in a peripheral region

By: Youssouf Merouani

Supervisor: Marcus Box
Södertörn University | School of Social Sciences
Master's thesis 30 credits
Business studies | Spring semester 2020

SÖDERTÖRNS HÖGSKOLA | STOCKHOLM
sh.se

Acknowledgments

Many people have contributed in different ways to make this master's thesis possible. Foremost, my supervisor, Marcus Box, for academic discussions, feed-back, commitment, and sharing my passion for quantitative research. Thank you! I am also grateful to my professors Paulina Rytkönen and Tommy Larsson Segerlind, for setting me on the path of Archipelago research. I also want to acknowledge the excellent suggestions of the examiner, professor Mikael Lönnborg.

Further, a special thanks to the whole team at Sjötrafiksektionen of Trafikförvaltningen for extracting 20 years' worth of data on Stockholm Archipelago visitors. I am also thankful for the statistics and information provided by the municipalities of Haninge, Norrtälje, Nynäshamn, Vaxholm, Värmdö, and Österåker.

Finally, for the unwavering support of my family. Mom, Dad, and Claudia. Thank you.

Youssouf Merouani
Stockholm, June 2020

Abstract

This longitudinal study investigates firm-survival in a peripheral region. The analysis relies on a unique longitudinal dataset, encompassing all firms in 17 Stockholm Archipelago islands during 2000-2019, collected through data triangulation methods. The study employs multivariate analysis accomplished by three-level logistic regression models of repeated observations (level-1) of the same distinct firms (level-2) situated in specific islands (level-3). The main novel results suggest that firms on islands with higher population density have an increased probability of exit – specifically, that there exists an overcrowding of local markets and a harsher selection process in these islands.

While time-distance to the mainland have been discussed as problematic, the effect on firm-survival is insignificant. However, local access to essential services and infrastructures significantly lowers the likelihood of firm-exit. Islands with increasing institutional thickness have increased community dynamics with more knowledge spillovers, community-level learning, and entrepreneurial ecosystems, leading to a higher likelihood of firm-survival. A higher ratio of summer visitors increases the probability of exit, suggesting that firms on islands with seasonal dependability may have difficulties surviving the off-season.

Keywords: firm-survival, peripherality, multilevel modeling, three-level logistic regression, Stockholm Archipelago.

Populärvetenskaplig sammanfattning

Trots dess närhet till Sveriges största och snabbast växande marknad delar Stockholms skärgård många drag med andra regioner som befinner sig i periferin, det vill säga utkanten. Bland annat har skärgården svårigheter att dra nytta av den ekonomiska dynamik som är karakteristisk för centrala regioner. Trots den problematik som kan vara typisk för perifera områden (såsom brist på resurser och avfolkning), betonar forskare att beslutsfattare måste utgå ifrån den lokalspecifika och regionala kontexten. Små och medelstora företag i periferin förväntas inte enbart bidra med ökade intäkter till regionen, utan förväntas även tillmötesgå sysselsättningsefterfrågan. Frågan om vilka företag som överlever, och inte minst förklaringarna till företags överlevad, är därför särskilt viktig i dessa regioner. Trots att orsakerna till företags överlevnad länge diskuterats inom forskningen, råder det brist på studier som tar upp den kontext som avser periferi och faktorer som är betydelsefulla för en specifik kontext. Denna studie utgör ett bidrag till en sådan analys.

Studiens resultat visar att överlevnadsmöjligheterna är sämre för företag som befinner sig på öar med högre andel fastboendebefolkning. Detta indikerar att på större öar, med en relativt sett högre befolkningsnivå, är konkurrensen mer påtaglig. Vidare, trots att transportmöjligheter till fastlandet ofta anges som problem i diskussioner, slår studien fast att tidsavståndet till huvudorter på fastlandet inte är viktigt för företags överlevnadsmöjligheter. Däremot har företag sämre utsikter för överlevnad på öar med högre andel sommargäster i relation till vinterbesökare. Tillgången till fiber/bredband, livsmedelsbutiker som är öppna året runt, och förskolor bidrar betydligt till överlevnadsmöjligheterna. Vidare konstaterar studien att ett mer aktivt föreningsliv på en ö ökar möjligheterna för överlevnad.

För att komma fram till dessa samt andra resultat, har studien använt sig av en longitudinell ansats, som innebär att studieobjekten (företagen) följs över tid, liksom ett unikt empiriskt underlag. Detta material har samlats från flera olika typer källor, underlag och databaser som beskriver både öar och företag i Stockholms skärgård, där undersökningsperioden är mellan år 2000 och år 2019. Vidare tillämpar studien en i forskning om företags överlevnad tidigare mindre beprövad statistisk metod: i korthet undersöker studien företag och öarna där de finns, som om dessa vore enskilda ”individer”. Dessa ”individer” följs från sin födsel –från att de bildas – till dess att de dör, det vill säga läggs ned. Under tiden ”individerna” följs, så undersöks även vilken betydelse de yttre betingelserna på öarna liksom faktorer i företagens omgivning har för deras överlevnadsförmåga. Specifikt studeras 1 341 företag på 17 av Stockholms skärgårdens öar (kärnöar).

Table of Contents

List of tables

List of figures

I. Introduction

1.1. Background

Scholars have identified peripherality as one of the constant causers of disparity between regions; an issue previously sought to be solved almost exclusively through investment in transport. With the start of the new century, some researchers proposed that new information and communications technologies (ICT) could solve issues for local firms as the geographical location would become less important (Copus 2001, p. 539). European and national policymakers, including those in Sweden, were quick to adopt ICT-centered development plans despite no empirical foundation (European Commission 2020). Thus, attempts to reduce these spatial differences within Europe have had only limited impact at best (Kinossian 2019, p. 61).

Peripheral regions, specifically those in the Baltic Sea Region, are burdened with depopulation, deficient infrastructure, increased risks for local economies, seasonal economic activities, and various economic challenges of the locals (Rytkönen & Kinossian 2019, p. 38). However, for decades, entrepreneurship studies have shown that an increase in new ventures leads to competitive advantage, increased productivity, increased innovation, economic growth, job creation, increased social mobility, and other benefits. Such benefits would alleviate many of the burdens linked to peripherality; however, these benefits are contingent upon the *survival* of firms (Yang & Aldrich 2012, p. 478).

Furthermore, policies are an essential tool to come to terms with the effects of peripherality. As Rodríguez-Pose bluntly declared, "*It's the Institutions, Stupid!*", emphasizing the importance of institutions for regional development (2013, p. 1036). However, researchers emphasize that development policies need to go beyond universal economic approaches and take into consideration the specific features of a region in order to be effective (Rodríguez-Pose 2013, pp. 1036–1039; Kinossian 2019, p. 64). Regional development authorities in Sweden, following OECD[1] definitions, have recently acknowledged that there exists a difference between rural areas that are closer to the cores and those significantly farther away. Specifically, areas where car transportation to the core region takes longer than 45 minutes. For the Stockholm region, only islands without bridges satisfy the definition (Tillväxt- och regionplaneförvaltningen 2018, p. 28). However, policymakers still do not fully take into consideration the unique characteristics of

[1] Organization for Economic Co-operation and Development (OECD), is an international organization working to shape policies, Sweden is a member country (OECD 2019).

these peripheral islands. For instance, regional investment is focused on tourism despite varied demography of local firms and the seasonal dependability of the sector (Rytkönen, Larsson-segerlind, Onn, Degerstedt & Kaipainen 2019).

Due to uneven and longstanding regional development within and across countries, there has long been an interest in core-periphery divides (Huggins, Prokop & Thompson 2017, p. 359). Scholars have proposed various definitions to define what constitutes a peripheral region. Some researchers use an economic approach, looking at regions that are lagging in terms of growth. Others emphasize geography and argue that regions far away from the centers, remote or in the periphery, constitute peripheral regions (Danson & De Souza 2011, pp. 6–11).

Kinossian (2019, p. 61), however, argue that peripherality should not only be defined by geographical proximity to the core regions, but should be determined by various dependencies on the core regions, either in terms of power, culture or stigmatization. According to Syrett (2012, p. 67), boundaries and borders, together with peripherality, are social constructions that produce processes of marginalization through the central aspect of who is an insider and who is an outsider. Eriksson (2008, p. 369) considers that regional development is often hindered due to narratives and strategies of "internal-otherness," whereby a fictionalized narrative of the region is created and normalized. These narratives are argued to impede the growth of regions as their actual characteristics are not taken into consideration; instead, proposals are made based on beliefs.

1.2. Stockholm Archipelago

> The key feature of islands is of course the shoreline – the fixed but ever shifting perimeter … Those who live inside this shoreline have a powerful sense of community – of communal home – dictated by geography - Harry Baglole, in Hay (2006, p. 21)

As will shortly be presented, the Archipelago of Stockholm possesses many characteristics of a peripheral region. Stretching about 150 km from Arholma in the north to Landsort in the south, the Archipelago extends along the Swedish east coast and stretches almost uninterruptedly eastwards, across to the Åland and Åbo archipelagos. Further, it is characterized by a landscape of joint-valleys, that give rise to cliffs and rifts, encompassing roughly 30 000 islands of varying sizes. Together, these islands create the largest connected archipelago in Swedish waters (Hill & Wallström 2008, p. 309). Roughly 200 islands are inhabited, and traveling between

them and the mainland is especially draining on resources of firms and inhabitants (County Board of Stockholm 2019, pp. 1–6).

This landscape, as well as the culture and environment of the Archipelago, have inspired many literary works. Perhaps even more significantly, these works have helped shape a persistent romantic narrative of the Archipelago. Widholm (2019, p. 59) describes the Stockholm archipelago as symbolic to the Swedes to such a degree that local entrepreneurs have to consistently negotiate with a fiction of the archipelago, making it exceptionally difficult to bring institutional change. Thereby, it does not come as a shock that Hay (2006, p. 10), similarly to other scholars, warns against the tendency to construct different narratives about islands. Particularly, Hay argues, these narratives often shape islands as conservative environments, sometimes as a romantic paradise, and the island inhabitants as "others."

Before the first half of the 19th century, workers and inhabitants in the archipelago were mainly fishers and farmers. This was primarily due to the prohibitions of the guild system. During the second half of the 19th century, with the advent of modern society and the removal of the guild system, the archipelago saw a short-lived boom in the transport sector but gave way to an inevitable decline. At the same time, the wealthy families and inhabitants of the city made the archipelago a place of recreation and relaxation during the summer months (Widholm 2019, p. 57).

Thereby, by the 20th century, local entrepreneurs were busy selling and renting land as well as running transportation, hotel, and spa services during the summer season. During the rest of the year, locals turned to traditional fishing and farming as well as few innovative projects. By the 21st century, the middle-class of the mainland followed the trend of the wealthy by heavily investing in archipelago real estate, thereby increasing tourism, and providing further opportunities for local entrepreneurs. Since then, subsidies and investments have been heavily focused on tourism (Widholm 2019, p. 58). However, researchers find that in Swedish Archipelagos, tourism benefits rarely reach the local communities. Furthermore, the adverse effects of tourism are found to affect traditional culture and resource use (Tunón, Kvarnström, Boström & Utbult Almkvist 2019, p. 45).

The number of permanent inhabitants in the islands has been continuously declining for years. Together with other sweeping changes, especially during the last century, the lives and means of support of archipelago inhabitants in the Baltic Sea have therefore profoundly changed; however, locals still have a strong relation to the archipelago's history, nature, and culture. And perhaps more importantly, in many islands, a strong sense of community. For instance, many

fishers still use the surplus of the catch to give to friends and neighbors (Tunón et al. 2019, p. 41).

> We small business owners who live here year-round are used to piecing together various business endeavors to make ends meet -A local entrepreneur from Utö (Lidström 2019, p. 71).

Similarly to fishers, fortunate local entrepreneurs often seek to invest and employ from their island and express desires to see the island flourish with more inhabitants (Lidström 2019, pp. 69–70). However, for many other small business owners in the Stockholm Archipelago, entrepreneurship is a question of necessity. Business owners need to cooperate and juggle multiple activities in order to make a living in the islands. Furthermore, the stark contrast of the winter months is an exceptionally challenging period, as many firms are seasonally dependent (Lidström 2019, p. 71). Rytkönen et al. (2019, p. 74) find that despite their physical closeness to the largest market in the country, entrepreneurs in the Stockholm Archipelago have entirely different preconditions than those in the core. The authors describe it as "operating in the shadow of the urban capital." Furthermore, they elucidate that while in the same archipelago, different islands have entirely different contextual and institutional settings.

The County Board Administration argues that providing and developing adequate infrastructure and service for all the inhabited islands in the archipelago is impossible. Thus, some "island communities" [skärgårdssamhällen] were pointed out, where investment in businesses and infrastructure will be concentrated. More specifically, the islands of Yxlan, Blidö, Ljusterö, Svartsö, Ingmarsö, Möja, Sandhamn, Ornö and Utö. With time, and due to the development plan of the Stockholm region, RUFS 2050[2], Arholma, Tjockö, Ramsö, Gällnö, Runmarö, Nämdö, Landsort, and Gräskö were further included. Together these 17[3] islands are designated "core-islands" (Kärnöar) (County Board of Stockholm 2019, p. 1). On top of the agenda for the development is "better accessibility," where investment is said to be focused on public transport on land as well as on water, more parking for boats, bicycles and cars in strategic locations, better transport of goods, as well as access to broadband and cell phone reception (Tillväxt- och regionplaneförvaltningen 2018, pp. 10–11). I argue that studying these specific islands becomes of interest, not least due to the plans of long-term investment.

[2] RUFS2050 [Regionala utvecklingsplan för Stockholmsregionen] is the name of the Stockholm regional development plan with a timeframe well into 2050, see (Tillväxt- och regionplaneförvaltningen 2018).
[3] See Appendix III for a map over the islands of the study.

In summary, I have mentioned several peripheral characteristics of the Stockholm Archipelago. First, natural barriers and other geographical aspects, beyond those of mere distance, puts inhabitants and local entrepreneurs at a disadvantage. Second, the history of Stockholm Archipelago is that of a region dependent on the greater metropolitan Stockholm region. Third, the entrepreneurial challenges and stories are typical of peripherality. Fourth, as discussed previously, fictionalized narratives of "otherness," in the archipelago taking the form of the "romantic idyll," is also a testimony to its peripherality. Thus, in the context of the empirical analysis presented in this thesis, Stockholm Archipelago, can be viewed as a peripheral region, despite its proximity to the Swedish capital.

1.3. Problematization

> The dearth of research on the role of spatial context is particularly precarious in relation to rural entrepreneurship on the conceptual as well as the practical level - Korsgaard, Ferguson & Gaddefors (2015, p. 578)

Peripheral regions and their specific features have been neglected in research, as considerable research has focused on central areas of Europe, especially cities and city-regions and their characteristics (Danson & De Souza 2011, p. 20; Burnett & Danson 2017, p. 28). In archipelago and island studies, much importance has been given to the relation between islands and the mainland/cities while ignoring the interplay between and among the islands, thereby "vanishing" the archipelagos and islands; the very context in which entrepreneurship in these places takes place (Stratford, Baldacchino, McMahon, Farbotko & Harwood 2011, pp. 114–116). Similarly, while institutions are critical for the development of regions, most studies on their role have focused on core-regions in developed regions (Zukauskaite, Trippl & Plechero 2017, p. 327). Further, effective development policies are dependent on research sensitive to local conditions (Kinossian 2019, p. 64).

Scholars are asking entrepreneurship and small firm researchers to pay more attention to space and time (Coviello & Jones 2004, p. 487; Aldrich 2009, p. 22), particularly the context of rural and peripheral entrepreneurship is less investigated (Korsgaard, Ferguson & Gaddefors 2015, p. 578; Clausen 2020, p. 114). In the present study, context is defined as "those variables that shape the characteristics of a *setting* and the behavior of different actors within that setting" (Zahra 2008, p. 248). Further, there is a need for longitudinal studies with extended time perspectives, significant populations, as well as the use of multiple levels of analysis, and analysis of change and transitions (Box, Gratzer & Lin 2017, p. 431). Scholars argue that there has been a failure in understanding of the multilevel nature of organizational events (Morgeson, Mitchell

& Liu 2015, p. 515). Finally, despite a growing body of research on firm-survival, much remains to be discovered about the factors affecting survival rates at a micro-spatial level (Huggins, Prokop & Thompson 2017, p. 357).

1.4. Aim and research question

As a result of an apparent knowledge and research gap, this thesis aims to analyze some of the factors impacting upon firm-survival across *time*, on *multiple levels* of analysis *within* a peripheral region. The analysis is based on the case of the Stockholm Archipelago, and the *micro-spatial* environment of its islands. The quantitative data of the study, *Archipelago Organizations Database,* collected and compiled by me for the present study, is a unique and relatively complete[4] database spanning over a century in all 17 core-islands of Stockholm Archipelago. While the quantitative data spans a century, the focus of the study is on the last 20 years (2000-2019). Thereby, the study's research question is:

- What are the challenges and conditions for firm-survival in a peripheral region?

1.5. Clarification of terms

In the present study, the term *firm* is used when referring to entities with one of the following legal structures; joint-stock companies, trading and limited partnerships, sole proprietorships, and economic associations. Furthermore, when the term *organization* is used, it indicates the *inclusion* of other types of organizational entities, such as foundations and non-profit organizations. The differentiation is primarily due to practicality, as the exit date for non-firm organizations is, at times, unsure. Thus, these organizations will be used to develop contextual variables. However, they will not be part of the primary statistical (survival) analysis.

[4] This description of the data will be further elaborated upon in the methodology chapter.

II. Theoretical Framework

In this chapter, the previous research upon which this study builds and the proposed theoretical framework through which this study contributes to the research field are presented and discussed. The chapter begins with a discussion on entrepreneurship in peripheral regions and concludes with the majority of hypotheses on the island level of analysis. This discussion is followed by a presentation of the study's perspective, organizational ecology. Further, I account for some of the variables that may influence firm-survival on the firm-level of analysis as well as the regional economy level. Throughout the present chapter, I will formulate hypotheses that are possibly falsifiable through the quantitative data. The chapter concludes with a conceptual model and a summary of the hypotheses.

2.1. Entrepreneurship in the periphery

2.1.1. Dynamic economy and entrepreneurship

The connection between region and entrepreneurship has been the focus of substantial attention in the academic discussion (Fritsch & Storey 2014, p. 940). In this discussion, entrepreneurship has is associated with more successful core regional economies, especially cities. In the core regions, university and research knowledge spillovers have been found to increase the opportunities of entrepreneurial activities due to flexibility and innovation typical of agglomeration economies (Audretsch, Hülsbeck & Lehmann 2012, pp. 498–499). Thus, peripheral regions with less economic dynamism and success frequently look to increase economic development by increasing entrepreneurship (Stephens, Partridge & Faggian 2013, p. 808).

Whereas excessive market crowding and competition, reduce survival prospects of firms (Geroski, Mata & Portugal 2010, p. 526), the argument has been that there exists a relationship between economic growth and the growth of new firms as well as the dynamic competition between these firms (market mobility). In short, existing firms are confronted by new firms. Those firms that can adapt are assumed to be better positioned relative to others. This "reshuffling" of firms is proposed to increase the performance of firms as well as employment growth. Thereby, economic growth and development are achieved (Koster & van Stel 2014, p. 215). However, for peripheral regions, these conditions and effects are not found. Fritsch & Storey (2014, p. 950) argue that, for peripheral regions, a point of departure is to recognize that real economic development - other than temporarily, growth of a significant employer, or through

the discovery of natural resources- is likely to take several decades. The authors base this argument on empirical evidence suggesting a lack of a "simple and automatic" link between the formation of new firms and job creation.

2.1.2. Opportunity and necessity

Why is it then that despite policies to increase entrepreneurship in peripheral regions, no similar effects have been reached? For this relevant question, Schumpeterian theories may provide an insightful answer. Schumpeter held that entrepreneurial skills are the "fundamental phenomenon of economic development." Thereby, increasing the number of entrepreneurs leads to the growth of the economy (Schumpeter 1983, pp. 74–75). However, Schumpeterian theories differ between self-employment originating from necessity and those originating from creativity and perceived opportunities where *innovation* is vital. Thus, though peripheral regions frequently employ policies for increased job creation, it may not necessarily increase the appropriate type of entrepreneurship (Stephens, Partridge & Faggian 2013, p. 780).

Investigating the relationship between entrepreneurship and economic growth, the GEM[5] studies, have shown that in high-income countries, levels of entrepreneurship are positively correlated with economic growth, while in low-income countries, it was negatively correlated. The reason? Entrepreneurship in developing countries is, to a high degree, due to necessity. It is then proposed that necessity entrepreneurship, contrary to opportunity entrepreneurship, does not translate to economic growth (Acs 2006, pp. 98–101). Thereby, we can assume that peripheral areas in developed countries, where lack of jobs necessitates self-employment, would show effects of necessity type entrepreneurship. This assumption would be correct, as empirical studies have shown that despite more entrepreneurs in rural areas, the entrepreneurial activity is due to necessity. Therefore, more entrepreneurship in peripherality is not necessarily more growth (Low, Henderson & Weiler 2005, pp. 68–69). While the present study does not aim to investigate opportunity and necessity, the results of previous studies indicating a prevalence of necessity entrepreneurship in peripheral areas narrows the search for relevant hypotheses.

2.1.3. Peripherality and islandness

Clausen (2020, pp. 115–116) coined the term "liability of rurality" to describe the particular disadvantage in terms of survivability and other organizational outcomes faced by firms in the context of the periphery. In essence, the author synthesizes empirical findings in the fields of

[5] Global Entrepreneurship Monitor (GEM), is a yearly global study measuring entrepreneurship in several countries.

geographic economy and entrepreneurship studies into three broader dimensions that may explain the liability of rurality. First, these rural (or peripheral) spaces are found to lack a *supply of resources*. Specifically, the author mentions difficulties obtaining research from universities, investments from financial institutions, access to recent technology, and more. Second, in these regions, there is a lack of knowledge spillovers from the community, lower community-level learning, lack of interaction between community members, and lack of entrepreneurial ecosystems and clusters. Together, the author describes these as a *lack of community dynamics*. Third, *lack of economic dynamism* is explained as the presence of a small market, lack or remoteness of users and suppliers, and lacking competition.

Further, a prominent characteristic of the peripheral context, including islands in the Stockholm archipelago, is that of declining population numbers. A study on the effects of population decline on entrepreneurship by Delfmann, Koster, McCann & Van Dijk (2014, p. 1048) finds that while the formation of new firms in urban regions tends to be negatively influenced by population change, in peripheral regions, the impact remains positive. The authors suggest that at first, population decline captures the attention, but as the decline continues, the regions adjust to the shock by increasing entrepreneurship. This increase is proposed to be a response to the community's need for minimum levels of services as the decline continues. I interpret these findings as suggesting that not only is necessity entrepreneurship in these regions prevalent, but it is also a reaction to the context of peripherality itself.

Moving on to the specific case of the present study brings us to the concept of "islandness." Baldacchino & Fairbairn (2006, p. 331) explains it as the various implications of being a business owner in a context of limited space, resources, client base, market, and increased transportation costs. Even in best-case scenarios, islands have difficulty reaching skilled workers, better research, and financing. Further, entrepreneurs seeking to be innovative, often do so within the geographical limits of their island. The authors find that these characteristics are prevalent in all island societies. Rytkönen et al. (2019, pp. 71–82) studied entrepreneurship on three islands of Stockholm Archipelago, while they find that archipelagos share many of the characteristics of islandness, the authors argue that the interplay between islands and the mainland brings a new contextual aspect. Notably, the authors find that despite physical proximity, there are substantial differences in the social capital and informal institutions of the different islands.

2.1.4. Institutional thickness

A significant amount of studies has found that institutions shape regional development (Rodríguez-Pose 2013, p. 1036). North (1990, p. 3) defined institutions as "*rules of the game*"

and described them as being of formal (e.g., laws and policies) and informal (e.g., traditions and culture) nature. Furthermore, in his seminal work, Baumol (1990, p. 918) successfully argued for the critical role institutions play in the allocation of entrepreneurship into productive, unproductive, and destructive forms. Thus, understanding how institutions change regions has been of interest in economic geography and entrepreneurship literature (Zukauskaite, Trippl & Plechero 2017, p. 326). In this literature, different strands have focused on different types of institutions. Mainly, concerning informal institutions, the notion of "*institutional thickness*" has been influential and linked to regional development (Rodríguez-Pose 2013, p. 1038). Amin & Thrift (1994, p. 15) define it as the "simultaneous collectivization and corporation of economic life, fostered by particular institutional and cultural traditions."

Amin & Thrift's (1994, p. 15) definition of institutional thickness originates from their contribution to the embeddedness hypothesis. In short, the hypothesis proposes that small firms in a local economy are in a stronger position of *survival* when embedded in a network formed by local loyalties, trust, and a collective response to threats. In their development of the institutional thickness concept, Amin & Thrift linked it to four factors; strong local institutional presence, high levels of interaction between local organizations, mutual awareness of being involved in a common enterprise, and structures of domination and patterns of coalition (Coulson & Ferrario 2007, pp. 592–593). The first factor regards the presence and variety of organizations, as well as their density, commitment, ownership, locality, and accountability. This factor, thus, concerns the ties of the organizations to a specific spatial locality (Coulson & Ferrario 2007, p. 599).

The second factor, high levels of interaction between local organizations, simply regards the intensity of the ties mentioned previously as well as the interaction between the nodes in the network (Coulson & Ferrario 2007, p. 601). The third factor regards the awareness of these organizations that they have a common goal. Specifically, it is linked to a strong sense of local identity and cohesiveness (Coulson & Ferrario 2007, p. 604). The last factor concerns the relative power of the local organizations. This factor is expressed as formal competencies of these organizations, as well as the perceptions of the local actors (Coulson & Ferrario 2007, p. 607).

Many studies have linked the outcomes of regional economic development to the thickness of the local informal institutions. Thin institutional context was found to hamper the capacity of learning while the opposite was found to foster the clustering of entrepreneurial activities and stimulate entrepreneurship (Rodríguez-Pose 2013, p. 1039). Therefore, similarly to the arguments of Clausen (2020, p. 116), peripheral regions are often found in empirical studies to lack

supportive organizations and dynamic clusters (Tödtling & Trippl 2005, pp. 1208–1208). However, there are some issues with the conceptualization of institutional thickness. Some of these are questions of cause and effect, conflating organizations with institutions, issues of geography and boundaries, and simplistic policy recommendations. The first three issues can be solved through proper research operationalization of the concept, which will be discussed and solved in the methodology chapter. As for simplistic policy recommendations, the issue at heart is that institutions cannot simply be duplicated or created. Further, success and failure are possible outcomes of any institutional combination and structure. Therefore, care must be taken when discussing and forming policies with an institutional aim (Coulson & Ferrario 2007, p. 594).

The aspect of institutional thickness goes hand in hand with what Clausen (2020, p. 116) discussed in regards to peripheral entrepreneurship and the effects of community dynamics, and clusters. Further, in the following discussion on organizational ecology, the diffusion of competence through networks is vital for firm-survival and linked to institutional thickness. Thus, linking these empirical findings together with the previously discussed aspects of peripherality and islandness, I argue that a thicker institutional context may play a significant role in the survival of firms in the periphery.

As discussed above, entrepreneurship and institutions are critical for the development and growth of the economy, in both core and peripheral regions. However, due to the particular context of peripheral areas, such as Stockholm Archipelago, necessity entrepreneurship, while less conducive to the overall growth of a region, is seen as more prevalent. However, within these peripheral regions, exceptionally depopulated areas are proposed to be more resilient due to having adapted to the initial shock.

To my knowledge, no prior studies have established links between firm-survival and specific contextual variables of infrastructure and service (e.g., access to fiber broadband and schools). Thus, I argue that on a micro spatial scale, and due to the context of the periphery, these factors are essential for local entrepreneurs and small business owners. Similarly, due to the effects of agglomeration economies, I further argue that the physical proximity to urban centers may function as a proxy of market distance. Finally, I argue that islands with a higher proportion of summer visitors may be more seasonally dependent and therefore have higher levels of firm-exit. Thereby it is possible to formulate the following hypotheses:

Hypothesis 1a: Lower island population density decreases the likelihood of firm-exit.

Hypothesis 1b: Longer distance to the nearest urban center (market) increases the likelihood of firm-exit.

Hypothesis 1c: The island level of service significantly affects the likelihood of firm-exit.

Hypothesis 1d: Increasing activity, in the form of visitors, increases the likelihood of firm-exit.

Hypothesis 1e: Increasing island land area decreases the likelihood of firm-exit.

Hypothesis 1f: Institutional thickness decreases the likelihood of firm-exit.

2.2. Organizational dynamics

The term "ecology" derives from the Greek words "Oikos," meaning dwelling place or household, and "logos," language of reason. Presently, this term means the "science of the relations of the organism to its surrounding outside world" (Schwarz & Jax 2011, p. 145). This same perspective of entities, environment, and relations, when applied to the study of organizations, offers a useful view from which to investigate firm-survival (Xi, Block, Lasch, Robert & Thurik 2020, p. 3).

The organizational ecology model describes organizational results as the variation and distribution of collections of organizations; population demography. The focus is on variations in the populations through the birth and failure of organizations. Further, the relation between organizations is considered. The organizations within the population are subject to environmental forces and dependent on specific resources niches, leading to competition and cooperation both within the same population and other populations (Aldrich & Ruef 2006, p. 35). The resource niche is explained as a limited resource that can supply a limited size of the organizational population. For instance, an abundant resource leads to a large population (Carroll & Hannan 2000, p. 246). The survival and growth of organizations are explained to be dependent on this competition and cooperation between organizations, resulting in a particular population adapted to the specific environment in which it is present (Aldrich & Ruef 2006, p. 35).

Inherent qualities of individual organizations, even if they play a role in how firms adapt to the environment, are largely neglected. Thereby, the perspective shifts the focus away from the individual organization and to the population (Xi et al. 2020, p. 3). Organizational ecologists concentrate on patterns of *vital events*, such as birth and failure of organizations, over time (Carroll & Hannan 2000, p. 39). Traditionally ecologists have downplayed events of transformation, as generalizing to the entire population has been of interest, and significant transformations have been statistically rare, almost always taking place within more prominent organizations. However, recently transformations within the population have been given increased attention in the organizational ecology literature (Aldrich & Wiedenmayer 2019, p. 63)

Furthermore, the processes of variation, selection, and retention are useful in the ecological approach: changes occurring in the population reflect their very operation. As entrepreneurs are founding organizations, either purposely attempting to adapt their plans to current circumstances or by chance stumbling onto opportunities: *variations* between organizations occur (Aldrich & Ruef 2006, p. 35). The higher frequency of these variations leads to more significant opportunities for change. Some of these variations will prove to be better adapted to the environment as they better acquire and use resources. Institutional, competitive, market and other forces set the *selection* conditions, and with time populations *retain* the characteristics of the surviving organizations. What is effectively retained in the population is the competence that all organizations use as a collective (Aldrich & Wiedenmayer 2019, pp. 63–64). The figure below illustrates this process quite well.

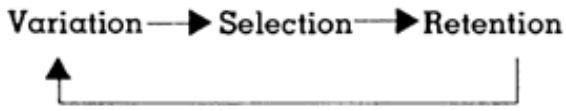

Figure 1: Process of variation, selection, and retention (Van De Ven & Poole 1995, p. 520).

Aldrich and Wiedenmayer (2019, p. 64), argue that these processes are not occurring sequentially throughout the population. Thus, contrary to the implication of the figure above, these processes are co-occurring. As new variations emerge, other organizations are in parallel being selected against, while yet others are retained, leading to simultaneous competition for resources. Further, the authors argue that the processes of *diffusion* and *competitive struggle* take place. The competence held by the individuals inside organizations is passed on to others through networks and links they are a part of, leading to the diffusion of the valuable competencies. Thereby, isolated organizations lack the linkage necessary to diffuse their competencies to the population. A competitive struggle over resources occurs once an organizational form thrives, again leading to variations and the selection of those forms that manage best to fit the environment at a given point in time.

2.2.1. Critique of Organizational Ecology

Finally, using the perspective of organizational ecology to study organizations is not without issue, at least when relying on registration data. To start, founders of firms often do not instantaneously establish new firms but rather do different actions to varying degrees, in different orders and different times. For instance, a firm founder may start developing products and seeking customers many months, even years, before seeking funds and registering the firm. Thereby, the age of firms is almost always underestimated since the birth of an organization happens

before its registration. Many of the organizations may be left entirely outside of the data as they are still in the nascent phase, while others may have exited the population without reaching the maturity level to register (Yang & Aldrich 2012, p. 483). Altogether, these issues regard the left-truncation of data and will be discussed in detail in the methodology chapter.

The organizational ecology perspective has had a limited impact on entrepreneurship studies, with a decline starting in the mid-2000s. The decline is explained partly by the challenging requirements set on the methodology, that the perspective was mainly popular in a limited circle of scholars, and that new institutional theory flourished and incorporated ecological rationale (Al-Turk & Aldrich 2019, pp. 108–110). However, organizational ecology offers a dynamic and multilevel model following populations over time and seeking to encompass all organizations in a population. Further, in this school of thought, vital events are essential. (Al-Turk & Aldrich 2019, p. 100). Finally, I am aware that, unlike the present study, homogenous populations of firms are often the subject of study in organizational ecology. However, other strands in the literature still make use of the dynamic perspective for more diverse populations (Box, Gratzer & Lin 2017, p. 435). Thus, I argue that this perspective offers an excellent tool for analysis in the present study.

2.3. Firm-survival and the macro-environment

In the literature, several general factors are found to influence the performance of firms. These factors range from the micro-level individual to the macroeconomic environment, and in-between; the structural factors of the firm (Box 2008, p. 379). While the present study does not delve into the individual-level factors, it does consider structural factors and macroeconomic factors. On the macroenvironment level, empirical studies have shown that firm-survival and exit vary as changes and transitions occur in the external environment of firms, either social, environmental, or economic (Box, Gratzer & Lin 2017, p. 434). In short, periods of macroeconomic contraction (recession) are found to increase the probability of exit due to decreasing investment and consumption. Conversely, periods of economic expansion (boom) lowers the probability of exit due to increased economic activity and aggregate demand (Box 2008, p. 388).

Geroski, Mata & Portugal (2010, p. 526) find similar effects of macroeconomic fluctuations on firm survivability. Furthermore, firms born during economic booms have permanently higher survival rates than those born during recessions. In other words, both current and initial macroeconomic conditions affect survival. Thus, the conditions improving or impairing survival go

beyond those of the firm's current market. National and international macroeconomic conditions will influence the survival likelihood of firms operating in any region due to the cyclical nature of the economy (Huynh, Petrunia & Voia 2010, p. 686; Congregado, Golpe & Parker 2012, pp. 1257–1258). Similarly, regional resources and conditions of the economy influence the outcomes of local firms (McAdam, Reid & Shevlin 2014, p. 84). Finally, multiple studies have found that during an exogenous shock, such as the global financial crisis probability of firm-exit increase substantially (Cefis & Marsili 2019, p. 573; Türkcan & Erkuş-Öztürk 2019, p. 15).

Hypothesis 2: The likelihood of firm-exit increases during worsened economic conditions

2.4. Structural factors of survival

Imprinting theory suggests that new firms obtain their characteristics from the environment in which they are founded. The imprinting then manifests as an impact on organizational behavior and outcomes, such as performance, adaptation, and survival (Simsek, Fox & Heavey 2015, p. 301). Thus, I shall now discuss some of these factors on the firm-level.

2.4.1. Age

In his seminal work, Stinchcombe (1965, p. 148) coined the term "liability of newness." The author explained that "as a general rule, a higher proportion of new organizations fail than old." Since then, empirical studies have established several links between various variables and higher probabilities of either survival or exit. The liability of newness signified that newer firms are more liable to exit than older firms, presenting *age* as a factor of survival (Aldrich & Yang 2012, p. 1). Several explanations for why age has this effect introduced in the literature revolve around internal factors (e.g., routines and experienced workers) and external factors (e.g., legitimacy and investor wariness) (Coad 2018, pp. 24–25).

A study on six million US firms showed that the probability of survival increases with age, with the most improvement ensuing shortly after 5-7 years with the rate staying flat onwards (Axtell 2016, p. 810). Similarly, following 240 thousand private sector firms born in the UK during 1998, a study showed that two out of three firms fail within the first five years. The hazard (risk of failure) rate was shown to be continuous but declining, and 15 years after birth, nine out of ten firms had failed (Anyadike-Danes & Hart 2018, p. 15). Several other empirical studies show support for the liability of newness, with increasing hazard until roughly after the 5-year mark, and then declining with age (e.g., Holmes, Hunt & Stone 2010, p. 186; Box, Gratzer & Lin 2017, p. 450).

However, some studies have shown that the liability of newness is not necessarily a universal rule that applies in all cases. Wagner's (1999, p. 77) study of firm exits in the German manufacturing industry showed that only a quarter of exits were by firms having an age of five years or less, though the probability of exit in general declines with age. Similarly, a study by Box (2008, p. 390) concerning exits of Swedish joint-stock companies between 1899 and 1950 while confirming the liability of newness for earlier birth cohorts, found no such effect for later cohorts.

Several other liabilities of age have been proposed in the literature to account for the varying relation between age and survival. For instance, the "liability of adolescence," proposes that firms will not have the highest probability to exit when entirely new, but instead will have a period where they can use their resources even when facing issues in the immediate time after entering the population. In a way, the liability of adolescence is a sort of honeymoon period for firms. Further, the "liability of senescence" explains that as firms age, they become rigid and inert due to accumulated policies, rules, and structures. This rigidity then leads to stagnation and eventual decay (Coad 2018, pp. 17–18). Finally, the "liability of obsolescence" originates in the changing environment of the firm. As these external changes progress, the firm is unable to adapt quickly enough and finds itself obsolete in a market that has already moved on, leading to a reduced likelihood of survival (Coad 2018, p. 25).

In summary, the relation between age and firm-survival has been established in the literature. Studies on firm-age and survival using cross-sectional methodology cannot make assumptions on increasing or decreasing age, only the present age of firms. In these studies, an apparent pattern emerges: older firms have lower rates of exit. However, longitudinal studies, following the same population of firms through time, have found that the hazard of exit changes. Notably, in the initial five years, the probability of exit increases substantially (liability of newness and adolescence), following this increase, the probability of exit decreases with age. However, in later years probability of exit increases yet again (liabilities of obsolescence and senescence). Thus, due to the non-linear relation of age and firm-exit, clarity is warranted when formulating hypotheses to avoid ambiguity:

Hypothesis 3a: Younger firms have a higher likelihood of exit than older firms due to the liability of newness.

Hypothesis 3b: The likelihood of new-firm exit increases until approximately year 5.

Hypothesis 3c: The likelihood of firm-exit decreases after reaching a peak after approximately year 5.

Hypothesis 3d: In the long term, the likelihood of firm-exit increases due to the liabilities of obsolescence and senescence.

2.4.2. Size

Economists often provide theories founded upon economies of scale, explaining that larger firms gain advantages, such as those of cost. Small firms are, in turn, showed to have a liability of smallness, providing *size* as a factor. Scholars mean that large organizations are hardier than smaller organizations, and therefore have a lower hazard rate (Carroll & Hannan 2000, p. 313). However, Carroll and Hannan (2000, p. 319) further describe the liability of smallness as being founded on "a thin empirical base." The authors mean that few studies follow organizations throughout their lifetime, thereby showing firm size as constant, instead of varying with time.

Higher numbers of employees were negatively associated with increased hazard rate in a study of three cohorts in the German manufacturing industry; however, the association was decreasing. Thereby, hinting to a limit when size no longer matters for survival (Wagner 1999, p. 77). Smaller firm-size has been found to negatively impact Swedish joint-stock companies' ability to survive (Box 2008, pp. 388–389). Recent studies on firm-survival have found that SMEs are more sensitive to macroeconomic downturns compared to larger firms (Box 2008, pp. 388–389; Türkcan & Erkuş-Öztürk 2019, p. 19) A study on founding conditions impact on Portuguese firms have shown that size indeed matters not only during the entry phase but its effects are still influential for a long time (Geroski, Mata & Portugal 2010, p. 523). However, the liability of smallness may not apply to all cases of firms. A study on a small population of UK firms showed that bigger initial plant size led to increasing hazard rates for micro-firms, but decreasing rates for SMEs (Holmes, Hunt & Stone 2010, p. 186). As will be presented in the results chapter, the vast majority of firms in the sample are micro to small firms. Thereby, a size hypothesis would be unfalsifiable.

2.4.3. Legal form

Despite many empirical studies on determinants of firm-survival, surprisingly, few studies address the relationship between legal forms and survival of firms (Baumöhl, Iwasaki & Kočenda 2020; Iwasaki & Kim 2020). For instance, a search on Scopus for *legal form* and *firm-survival* results in only seven references[6], when limiting for relevant fields. In comparison, a Scopus search for only *firm-survival*, excluding non-relevant fields of study, yields 5,835 results[7].

Harhoff, Stahl & Woywode (1998, pp. 470–471) made one of the earliest efforts to study the relationship between the legal form of a firm and its survivability. The researchers studied firm

[6] Search was conducted on the 2020-04-20 using the following search string, while excluding non-relevant study fields: survival AND (firm* OR venture* OR business* OR organi*ation*) AND "legal form*"
[7] Search was conducted on the 2020-04-20 using the following search string, while excluding non-relevant study fields: survival AND (firm* OR venture* OR business* OR organi*ation*)

survivability of German firms while distinguishing between voluntary exit, such as liquidation due to retirement, and involuntary exits, such as bankruptcy. Their results show that limited liability firms have higher insolvency rates compared to unlimited liability firms, such as sole proprietorship. However, the researchers find that voluntary exit is much more common for sole proprietorship than other legal forms. The authors argue that these results are more reasonable than prior[8] results in their study when not considering the reasons for exit.

Studies conducted in Spain (Esteve-Pérez & Mañez-Castillejo 2008, p. 244) and Portugal (Mata & Portugal 1994, p. 242), propose that limited liability firms, survive longer than those of unlimited liability. Further, a study on the survivability of Russian firms during 2007-2015, find that, especially in the service industry, all things being equal, the limited liability of firms is negatively associated with higher hazard rates. Moreover, the researchers found a relationship, albeit non-linear, between the organizational openness of legal forms and survival rates of firms (Iwasaki & Kim 2020, p. 23). Similarly, in Turkey, a study found that the corporate legal form of firms in the tourism sector is negatively related to increased hazard rates. In other words, corporations are more likely to survive than non-corporations (Türkcan & Erkuş-Öztürk 2019, p. 16).

Similarly, an extensive study on the survivability of 41,496 firms operating in Czech, Slovakia, Hungary, and Poland during 2006-2015 was conducted by Baumöhl, Iwasaki & Kočenda (2020, p. 14). The study found that the legal form of a company is a significant economic factor for firm-survival. Notably, while there are some differences between countries, the study finds that, in general, limited liability firms have higher survival rates than those of unlimited liability. Furthermore, when considering reasons for firm-exit, limited liability firms have higher insolvency rates than those of unlimited liability.

In Sweden, organizations may have several legal forms. In order to better analyze the effects of these forms on survival, it is essential to understand not only the extent of their liability but also their requirements and purpose. The liability economic associations' [ekonomisk förening] members are limited to the contribution paid to the association; the association is thus regarded as a legal entity (Swedish Companies Registration Office 2020a). Further, at least three persons should be members, and if the number of members drops below this level, the association will be liquidized (Hemström & Giertz 2014, p. 31). Economic associations have many similarities

[8] Prior results by Harhoff, Stahl & Woywode (1998) showed that limited liability brought higher hazard rates, something the authors called into doubt.

to joint-stock companies, with the defining difference being that influence is linked to membership, irrespective of stake scope. Furthermore, since the ’80s, economic associations have been founded for primarily social causes (Lithander 2005, pp. 27–36).

Joint-stock companies [aktiebolag], have the highest requirements for creation, in terms of initial investment compared to other legal forms. Further, in joint-stock companies, the invested capital of the shareholders forms the prerequisite for the company’s operations. However, for trading partnerships [handelsbolag] and limited partnerships [kommanditbolag], the central aspect is the cooperation between the founders and their shared goals (Hemström & Giertz 2014, p. 58). Further, while the liability of joint-stock companies is limited, it is not the case for trading and limited partnerships as either one or all of the owners are responsible for the company’s debt (Swedish Companies Registration Office 2018a, 2019, 2020b). Finally, sole proprietorship [enskild näringsverksamhet], can be created by any person over the age of 16. Moreover, while the fees of registry for this firm are small, the owner is entirely and solely liable for the firm’s debts and contracts (Swedish Companies Registration Office 2018b).

In summary, studies on the legal structure and firm-survival find that with unlimited liability comes increasing hazard rates. Thereby the following hypothesis is proposed:

> *Hypothesis 4: Firms with limited liability have a lower likelihood of exit compared to firms with other legal forms.*

2.4.4. Firm-industry

Empirical studies have shown that firm-survival, similarly to other structural factors, depends on the firm’s activities. In other words, the industry in which the firm operates. Notably, the industry life cycle and the technological demands of the industry shape the survivability of firms. Depending on the firm's ability to either find niches in the market or its size to shoulder burdens of initial research and development investment, the firm may manage to grow and survive. Further, it is argued that in expanding industries, firms are found to have better survival probabilities (Agarwal & Audretsch 2001, pp. 39–40).

Studies on firm-survival accounting for different industries find differing results between countries. However, these results may be due to the different classification of industries. For instance, a study on Swedish joint-stock companies showed that firms with manufacturing activities had a significantly lower hazard rate than that of companies in the trade and service industries (Box, Gratzer & Lin 2017, p. 450). Similarly, A study on Russian firms’ survivability shows similar effects whereby firms with construction and service activities show the highest exit rate followed by agricultural firms. Manufacturing and mining companies had the lowest

exit-rate (Iwasaki & Kim 2020, p. 31)[9]. However, for Italian firms, the highest survival probabilities are found in the clothing and footwear industry. Conversely, retail trade, furniture, and mechanical manufacturing have the lowest probabilities of survival (Ferragina, Pittiglio & Reganati 2012, pp. 369–371). Thus, I formulate the following hypothesis:

Hypothesis 5: Firms in manufacturing have a lower likelihood of exit than firms in service and trade.

2.5. Summary

Events occur at every level of an organization, from the most micro individual level to the most macro-environmental level. The effects of these events, further, move up, down, and within these levels, and over time, they shape thoughts, feelings, and actions. (Morgeson, Mitchell & Liu 2015, p. 515). I argue that firm-exit is an organizational event that does not randomly happen. As scholars have established in the literature, several factors, from the most micro to the most macro, are linked to increased or decreased chances of survival. While I have not accounted for the individual factors of the entrepreneurs, such as their human capital and growth orientation, these factors play a definite role. Furthermore, for the study, I have established factors on the firm and more macro levels presented by the conceptual model below (Figure 2). Further, while the model implies a direct influence on firm-exit, in reality, other variables cross-interact and are possibly moderated or mediated by yet other variables outside the scope of the study. Due to the complexities[10] of multilevel statistical analysis, limitations of the *Archipelago Organizations Dataset*, and, more importantly, the aim of the study, only *some* of the factors in the conceptual model are hypothesized. Here follows a summary of the study's 14 hypotheses:

Hypothesis 1a: Lower island population density decreases the likelihood of firm-exit.

Hypothesis 1b: Longer distance to the nearest urban center (market) increases the likelihood of firm-exit.

Hypothesis 1c: The island level of service significantly affects the likelihood of firm-exit.

Hypothesis 1d: Increasing activity, in the form of visitors, increases the likelihood of firm-exit.

Hypothesis 1e: Increasing island land area decreases the likelihood of firm-exit.

Hypothesis 1f: Institutional thickness decreases the likelihood of firm-exit.

Hypothesis 2: The likelihood of firm-exit increases during worsened economic conditions

Hypothesis 3a: Younger firms have a higher likelihood of exit than older firms due to the liability of newness.

Hypothesis 3b: The likelihood of new-firm exit increases until approximately year 5.

[9] The author mainly discusses legal structures, however, table 2 in the appendix presents the results discussed.
[10] These complexities will be discussed in multilevel modeling section of the methodology chapter.

Hypothesis 3c: The likelihood of firm-exit decreases after reaching a peak after approximately year 5.

Hypothesis 3d: In the long term, the likelihood of firm-exit increases due to the liabilities of obsolescence and senescence.

Hypothesis 4: Firms with limited liability have a lower likelihood of exit compared to firms with other legal forms.

Hypothesis 5: Firms in manufacturing have a lower likelihood of exit than firms in service and trade.

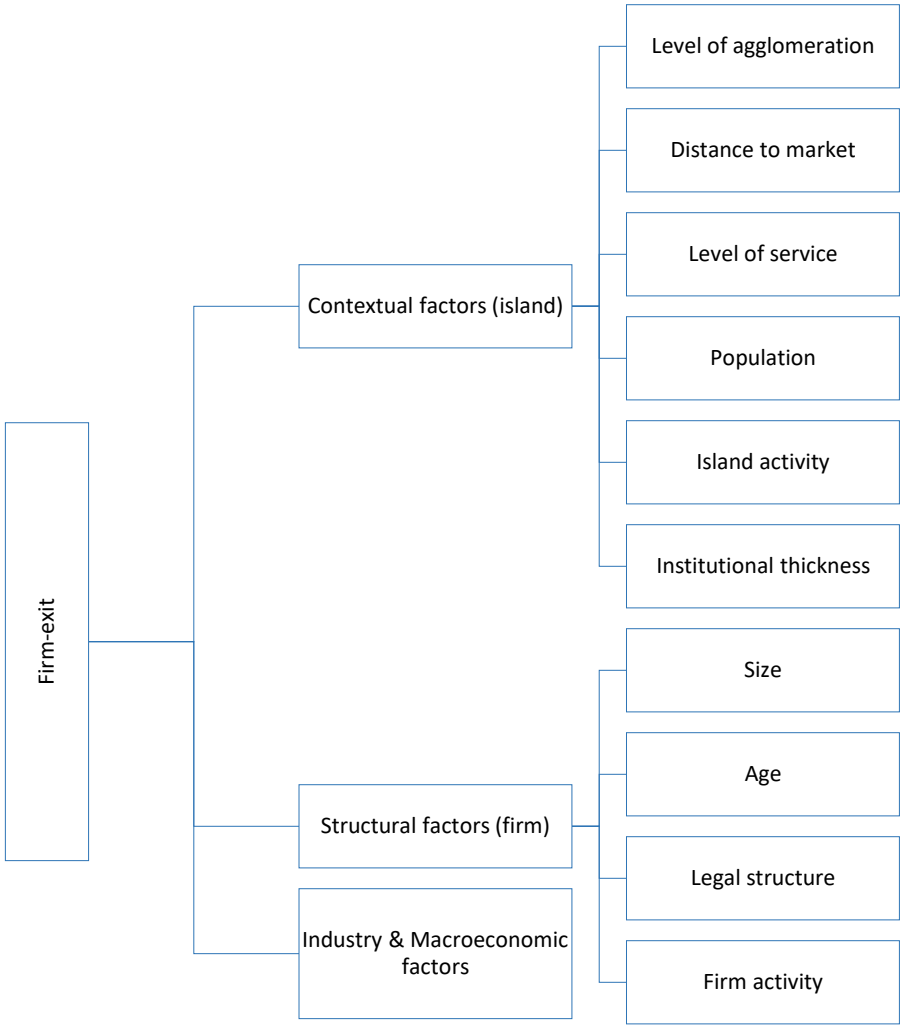

Figure 2: Conceptual model of firm-exit and linked factors (own figure).

III. Research Methodology

In the introduction, a research question was presented inquiring on the challenges and conditions for firm-survival in the Stockholm Archipelago. It is an exploratory question that defines the direction of the study's investigation into the context of entrepreneurship in this peripheral region. The study aims at encompassing the firm and island levels of analysis through a period of 20 years, from 2000 to 2019. The nature of this research, to *quantify* a phenomenon through statistical analysis, warrants a quantitative approach. Thereby, the present chapter presents the methodological choices made for this study, what has motivated them, and their consequences.

The chapter begins with a discussion on the longitudinal nature of the study, exploring the importance of the extended time perspective and its importance for the strength of inference. The discussion then moves on to the principal statistical analysis of the study, multilevel survival modeling. The discussion mainly focuses on why such an approach is warranted and then how to conduct it. After which, I shall present the sampling strategy, and describe the collection of data. The chapter concludes with short but important discussions on the analysis strategy, validity, reliability, and ethical considerations of the study.

3.1. Longitudinal research

> Time is a river, a violent current of events, glimpsed once and already carried past us, and another follows and is gone. – Marcus Aurelius (Marcus Aurelius & Hays 2003, p. 84).

The stated objective of the study is to analyze firm-survival in the peripherality context of Stockholm Archipelago through *time*. Thereby, the research objective should be addressed by using longitudinal data together with statistical methods that help us detect patterns across these multiple subject histories (Smelser & Baltes 2004a, p. 9058). A natural result of thinking across space and time is that it leads to a focus on change, dynamics, and longitudinal phenomena (Morgeson, Mitchell & Liu 2015, p. 516). In longitudinal data, each subject (firm) is observed at several occasions over time, making it possible to study individual change, both due to the passage of time and explanatory variables (Rabe-Hesketh & Skrondal 2012, p. 5)

Furthermore, the study is looking forward, from a specific point in time, thereby making it a prospective study. Longitudinal data have, thus, advantages over cross-sectional data, for making a more robust inference, identifying causality, and causal mechanisms (Box-Steffensmeier & Jones 2004, p. 4; Hedström & Ylikoski 2010, p. 64). Furthermore, longitudinal investigations

are suited for learning with an open mind, which ensures identifying and investigating crucial variables over time (Edmondson & Mcmanus 2007, p. 1162). As firms are born, they enter the population of the study, and as they die, they exit the population (Smelser & Baltes 2004b). In summary, the study reaches a higher level of quality by incorporating the element of time and thereby making it possible to draw a more robust inference on the challenges and conditions of survival for the firms in the sample.

3.2. Multilevel survival modeling

In the following section of the methodology chapter, I start with an explanation of survival analysis, multilevel modeling, and the type of statistical analysis used. Moreover, I will discuss the issues, solutions, and considerations relevant to the data and results of the study. This discussion will be critical for all subsequent parts of the thesis, as it sets what inference can be made and the limits of the statistical analysis.

3.2.1. Survival analysis

Survival, failure-time, time-to-event, reliability, or duration analysis are all names for *event history modeling*. The rich vocabulary arises from the usefulness of different kinds of applications that these models make possible. The main interest is the time from one particular event to another event—for instance, time from the entry of a firm into the population to its exit. The unit of analysis, the *firm*, is then said to become at risk of the exit event after the initial event, entry, has occurred (Box-Steffensmeier & Jones 2004, p. 2; Rabe-Hesketh & Skrondal 2012, p. 743). History event analysis is a comparative analysis that makes possible claims on survival and risk, but also comparative inference across cases. Application of history event models is warranted when researchers have a significant problem to study, requiring comparative analysis of longitudinal data. As these types of problems are frequent in social sciences, it makes it a "natural model" to use for analysis (Box-Steffensmeier & Jones 2004, pp. 4–5).

In survival analysis, the term "hazard" is often used to imply risk. Hazard, in the case of the present study, is the probability that a firm who is under observation at a time *t* has an exit event at that time. The hazard function is then the cumulative rate of hazard that the event will take place (Chakraborty 2018, p. 43). The most crucial consideration for time-to-event analysis is perhaps the origin of the time scale. In the present study, the origin (point in time) is when the individual firm first becomes at risk for exit (Rabe-Hesketh & Skrondal 2012, p. 743). In the firm-survival literature and this thesis, the origin is also the firms' time of entry. A further discussion on timing and survival analysis follows bellow.

3.2.2. Considerations: left-Truncation, right-Censoring, and competing risks

The study of organizational ecology, and especially vital events, is not without issue, as gathering data on emerging organizations is challenging. Registration data, such as the one used for the present study, contain information on many organizations in the population. However, the "date of registration" is not always close in time to when the business owner started creating the firm. The date of registration is simply when the firm owners officially submitted relevant paperwork to the Swedish Companies Registration Office [Bolagsverket].

In other words, the individual firms, and other organizations, have been at risk of exit before being registered by the authorities. Furthermore, many other organizations have not achieved the level of performance necessary to be officially created, while others have already exited the population before this event. Thereby, the real vital event of birth is missed (Yang & Aldrich 2012, p. 480). I will explain these issues and their consideration through the following illustration:

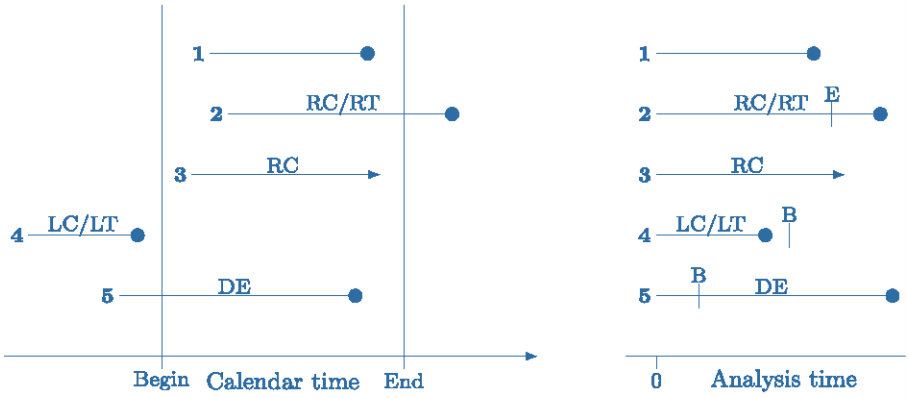

Figure 3: Illustration of different types of censoring and truncation in calendar time (left) and analysis time (right). Dots are events, and arrowheads are censoring (source: Rabe-Hesketh & Skrondal 2012, p. 745).

Firm 1 in the above illustration is an example of an ideal firm that enters and exits the population between the set time-frame chosen in the research. Firm 1, can be analyzed from the moment that the founders have started the venture to its exit. In the present study, I do not know (exactly) when the firms were created, only the date of registration. Therefore, I cannot be sure of any firm having ideal circumstances in the data. Firm 2 has a form of right censoring as the exit event does not happen during the time of analysis. In other words, during the time-frame of the study, Firm 2 continues to survive (Rabe-Hesketh & Skrondal 2012, p. 745).

Firm 3, represents a right censoring form for firms that stop being at risk for the exit event during the time of the investigation. This type of censoring is not possible in the present study as the exit is only defined as *ceasing operations*. However, if the exit was defined explicitly as

bankruptcy, and the firm went through a merger or other type of exit before bankruptcy could happen, it would be represented by Firm 3. In other words, the firm would have succumbed to a *competing risk* before the specific risk analyzed in the study. Competing risks are of particular interest for firm-survival studies that analyze the different options available to firms to leave the market, using "multiple-spell" models. For instance, in the literature chapter, I covered the study of Harhoff, Stahl & Woywode (1998) that demonstrated different effects of explanatory variables on survival estimates depending on the type of exit. However, for studies where the principal analysis lies in what determinants leads to the firm continuing or exiting the market, such as the present study, "single-spell" models are useful (Manjón-Antolín & Arauzo-Carod 2008, p. 15)

Firm 4 represents the first issue discussed at the beginning of this section: left-truncation. Firm 4 has entered and exited the population without ever being registered and, therefore, is not known in the present study. Firm 5, representing the second issue discussed at the beginning of the section, likewise have left-truncation but have survived long enough to reach a phase where registering with the authorities is necessary. Firm 5 has, in other words, a *delayed entry*, and have been at risk for failure for a longer time than the date of registration, by the authorities, indicates (Rabe-Hesketh & Skrondal 2012, p. 746). This type of left-truncation is problematic for almost all entrepreneurship and business studies relying on registration data. In the present study, it is therefore highly probable that almost all firms in the data have a slight form of delayed entry.

In the literature, studies on firm-survival have controlled for left-truncation by several methods. Some studies have only included cases with the complete observation of entry and exit, or including cases started within a short time before the beginning of the study. While others have used dummy variables, continuous variables, or statistical models to account for truncation-effects (Yang & Aldrich 2012, p. 483). Similarly, Rabe-Hesketh & Skrondal (2012, p. 773) describe a statistical method to correct left-truncation. However, the method is dependent on knowing the real-time of entry, when the subjects first became at risk for the event.

For the present study, these solutions are not easily feasible. Simply removing subjects that are not entirely observed from the data would lead to an empty database: rendering the study impossible. Another solution would be to calculate the time between the real entry date and the date of registration by the authorities for each organization. However, this is only possible by contacting the founders of each organization within the *Archipelago Organization Database* and asking when they started with activities such as planning or organizing. Not only is such a

process time consuming, but even in the best of scenarios, it would bring issues that often plague retrospective studies. For instance, people often forget or inaccurately recollect when in time events happened (Taris 2000, p. 8). Furthermore, these solutions can only account for the delayed entry type of left-truncation. Thereby, still leaving us with issues of missing subjects such as those illustrated in the above figure by Firm 4.

Fortunately, a study on the new-venture survival of firms established during 1998 in Sweden, found legitimating organizational activities that involve government entities, dependent on official registration with the authorities, to be the first step for most new ventures (Delmar & Shane 2004, p. 405). Furthermore, a recent study on new firms in Stockholm finds that 70 % of firm founders create the firm within a year of having the initial idea (Lundström, Nordström & von Friedrichs 2014, p. 13). Therefore, I argue that it is more probable that the date of registration in the *Archipelago Organization Database* is relatively close in time to the real-time entry of the organization. Thereby, going forward, the registration data is seen as the entry date for all organizations. However, following the recommendations of Yang & Aldrich (2012, p. 490) and the results of Delmar & Shane (2004, pp. 400–405), I will keep in mind that the sample being analyzed is, at least to a degree, biased toward resilient firms that have survived the nascent phase of their lives. In practice, the inference will only apply to firms that have survived the nascent phase, not the whole population.

3.2.3. Multilevel modeling: Why?

> People don't buy what you do, they buy *why* you do it. – Simon Sinec (2009, p. 46).

In their research article, Molina-Azorín, Pereira-Moliner, López-Gamero, Pertusa-Ortega & Tarí (2019, p. 2) seek to answer the above question through an organizational research perspective. To start, the authors find that outside the domains of organizational psychology and behavior, other domains within the organizational studies field have primarily ignored this methodological approach. The authors argue that a multilevel approach offer opportunities to improve and expand research through the development of new theory, as well as the improvement and development of data collection techniques, analytic techniques, as well as managerial solutions. More specifically, the implementation of multilevel approaches helps bridge the gap between the micro and macro when studying a phenomenon, and perhaps more importantly, avoiding statistical fallacy, which I will describe shortly.

The *Archipelago Organizations Dataset* used for the study is coupled with data variables on the islands to reach a three-level structure represented in the illustration below. The observa-

tions of the same distinct organizations over time become longitudinal (level-1); these organizations (level-2) are, in turn, nested in specific islands (level-3) (Liu 2020, p. 3030). It is imperative to note that not all multilevel models are hierarchical. If I had included industry variables in the study, such as industry-wide capacity, it would have been necessary to implement a cross-classified model. The reasoning is that it would not have been possible to argue that islands fit within industries, or that industries fit within islands (Hox, Moerbeek & Schoot 2018, p. 161). However, it is possible to fit firms within islands and consequently add time-varying island level variables that, together with the age of firms, vary from year to year. Therefore, in the present study, the observations at level-1 function as a yearly "snapshot" (Hox, Moerbeek & Schoot 2018, p. 72).

Subjects nested in different locations experience different contextual factors. Conversely, observations nested within subjects tend to have a high correlation within. This high correlation is due to the similarities of the within-group members. For instance, the observation within a single firm will be more similar than observations between several other firms. Likewise, firms within one island will be more similar than firms from different islands. This similarity, or correlation, is called intraclass correlation (ICC). Ignoring a moderate to high ICC value violates the assumption of the interdependence of the observations, an essential assumption in statistical analysis. For instance, analyzing a three-level structure as if it was a one-level structure, leads to smaller standard errors (large bias), and thus results in falsely significant results (Seltzer, pp. 259–260; Hox, Moerbeek & Schoot 2018, p. 4).

Altogether, I argue that as the purpose of the study is to study firm-survival in the context of the islands in which they are situated, employing a multilevel model from the study design level, all the way to the statistical analysis level is vital. Not only, may the study draw a more robust inference, but more importantly, statistical results become more reliable. Finally, it is essential to remember that on the island-level, every variable will be time-varying (change year-to-year), while on the firm-level, only the age and alive-status of the firm varies with time.

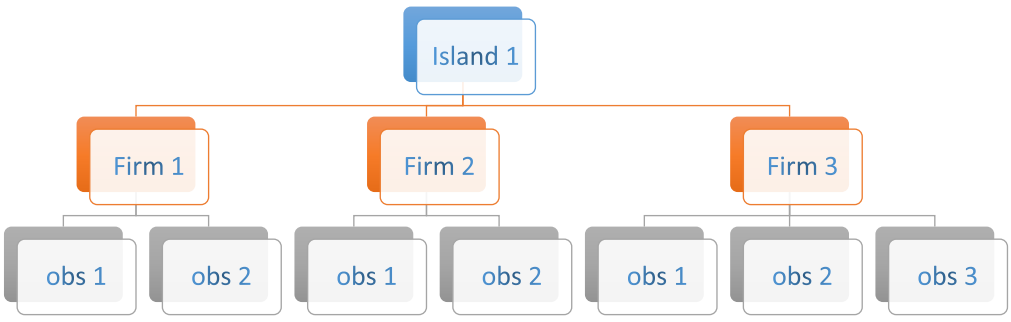

Figure 4: Illustration of the study's hierarchical model structure.

3.2.4. Multilevel modeling: from simple to multilevel logistic regression

Previously, I established *why* a multilevel model is warranted, but the question of *how* to conduct it is yet to be explained. In survival analysis, the outcome variable (y) describes the occurrence or censoring (absence) of the exit event. The outcome variable, generally, takes values of 1 when the event has occurred, and values of 0 when the event is censored. In the present study, when the firm exits the market, the outcome variable takes values of 1, and when the firm continuous to survive, it takes values of 0. Thus, logistic regression is especially useful as it is employed in cases where the outcome variable is of a dichotomous (binary) nature (Hox, Moerbeek & Schoot 2018, p. 107). To explain a multilevel logistic regression, I shall, to a degree, rely on the article by Sommet & Morselli (2017, pp. 204–209), which I find to provide an excellent explanation of two-level models. However, the example used will be of my invention. Further, I will adapt the explanation to three-level models.

(1) $$Y_{\mathrm{i}} = B_0 + B_1 * X_i + e_{\mathrm{i}}$$

Let us start with a simple linear regression represented by equation 1, where Y_i is the outcome (dependent) variable. For our purposes let it be thesis grade (1-100) for student i, explained by the predictor (independent) variable $\mathrm{X_i}$. $\mathrm{X_i}$ in our example is the number of weekly hours spent on a cardio workout, perhaps swimming, while writing the thesis. B_0, The intercept is the predicted value of the outcome variable if the student spends 0 hours per week swimming. $B_1$is the estimate describing the relationship between time spent swimming ($\mathrm{X_i}$) and the grade (Y_i). If $B_1 = 5$ it would indicate that an increase of one-hour swimming per week, would increase the grade by 5 points. Linear regression then results in the predicted mean value of the dependent variable (grade of the thesis) at a specific value of a predictor variable (number of hours per week spent swimming). The residual, e_{i}, is simply the "distance" between the value that the model predicts and the observed value in real life. If we plotted this function on a graph, we would get a straight line that starts from the intercept value and has a slope equivalent to the $B_1 * X_i$ value; hence *linear* regression (Sommet & Morselli 2017, p. 204).

B_1 is then particularly important to understand the link between time spent swimming and the grades of the thesis. Two things can happen during such a regression. If B_1 is not significantly different from zero, then there is no significant relationship between the outcome and predictor variable. In this case, we have to accept the null hypothesis. If, however, B_1 is significantly different from zero, then there is a significant relationship between the outcome and predictor variables, and we accept the alternative hypothesis (Sommet & Morselli 2017, p. 204).

Now, assume instead, we no longer have the possibility to achieve a grade of 1-100. Instead, we employ a Swedish grading system of either U (fail) or G (passing). In such a case, our predictor variable takes the value of 0 for fail and 1 for passing. This case is similar to that of the present study. Firms can either exit (1) or survive (0). The outcome variable is binary, and therefore linear regression is no longer possible. Instead, as explained at the beginning of this section, logistic regression is warranted. Logistic regression (equation 2) results in the *conditional probability* that an outcome variable equals 1, $P(Y_i = 1)$, at a particular value of a predictor variable. The exponent of $B_0 + B_1 * X_i$Functions as in the linear model described in equation 1, but now with *probability*. In other words, the equation provides the likelihood of the student obtaining a passing grade (1), when spending a certain number of hours per week swimming (Sommet & Morselli 2017, p. 205).

(2) $$P(Y_i = 1) = \frac{exp(B_0 + B_1 * X_i)}{1 + exp(B_0 + B_1 * X_i)}$$

If I plotted this function on a graph, we would not see a straight line. Instead, we would see a slightly straightened out S-shaped line that is flatter at the top and the bottom. Using the equation as is, would make it difficult for us to interpret the results. Furthermore, the relationship would, by definition, no longer be linear. However, by *transforming*, specifically logit transforming, both sides of the equation, we would instead *linearize* the relation. Thus, reforming the S-shape into a line (Eq. 3). First, compared to Equation 1, now we do not have a term for the residual e_i, as it is no longer necessary. Second, the equation results in the log-odds that the grade is G (1) over the probability that the grade is U (0). In short, the possibility that something will occur rather than not occurring is called the *odds*. Now, if we tried to interpret the meaning of having $B_1 = 1$, it would still be hard to understand (Sommet & Morselli 2017, p. 205).

(3) $$Logit(odds) = B_0 + B_1 * X_i$$

It would be more meaningful to raise it to the exponent, which will formulate the numeric value in terms of *odds ratio* (OR). For instance, $odds\ ratio = OR = exp(1) \approx 2.7$. As the OR value is larger than 1, we multiply it by 1. The resulting value, 2.7, means that students are 2.7 times *more* likely to have a passing grade when swimming one more hour per week. If $B_1 = -1$, it would mean that the $odds\ ratio = OR = exp(-1) \approx 0.4$. This value is smaller than 1; thus, we divide 1 by the value of the odds ratio, $\frac{1}{OR}$, whereby it becomes 2.7. This resulting value means that the student is now 2.7 times *less* likely to have a passing grade when swimming one

more hour per week. In short, odds ratio values that are smaller than 1 mean that the predictor decreases the likelihood, while values larger than 1 increases the likelihood. As in the case with linear regression, two things may happen. Either, OR is not significantly different from 1 (or B_1 not significantly different from 0), and we accept the null hypothesis. Otherwise, OR is significantly different from 1 (or B_1 significantly different from 0), and we accept the alternative hypothesis (Sommet & Morselli 2017, p. 206).

Now assume that instead of one student, there are many more students (i) in several classrooms (j). The study now has become a two-level model, and if we leave it empty of the predictor variables, it is represented by equation 4. This function results in the log-odds that the grade is G (1) instead of U (0). B_{00} is the fixed intercept just as in equation 3, while u_{0j} represents the distance from the class-specific intercept (level-2) to the fixed intercept. Hence, u_{0j} is the residual on the level-2. If we would continue to add levels by adding the university-level (k). We would have equation 5 (Sommet & Morselli 2017, p. 206).

(4) $$Logit(odds) = B_{00} + u_{0j}$$

(5) $$Logit(odds) = B_{00} + u_{0jk} + v_{0k}$$

Equation 5, effectively have one more term, v_{0k}, which is the distance from the university-specific intercept to the fixed intercept, thus, the residual on the level-3. Equation 5 is, therefore, our empty (intercept-only) three-level model. The difference between simple logistic regression (Eq.3) and multilevel logistic regressions, (Eq.4) and (Eq.5), is that we have allowed the log-odds to *vary* from one class to another, and from one university to another. In other words, in the last equation, we have a fixed intercept and two random intercept variances. The fixed intercept explains the overall log-odds of having a passing grade in the thesis for a typical student in a typical classroom and typical university. If we would give it a value and calculate its OR, we would get the likelihood to pass the thesis across all classes and universities. However, as the log-odds is allowed to vary, the level-2 residual gives us an indication of how large the variations may be across classes within universities, while the level-3 residual gives us an indication of the variation across universities. Higher random intercept variance of level-1 and level-2 residuals means that students have more chances of passing the thesis in some classrooms and some universities, respectively (Sommet & Morselli 2017, p. 206).

Now if we would include level-1 predictor variables, we would denote them X_{ijk}, for level-2 X_{jk}, and for level-3 X_k. X_{ijk} would then be variables on the student level. For instance, previous

grades, hours spent swimming, and time spent studying. For the class level, X_{jk}, we could have variables such as the supervisor experience. Finally, for the university-level predictor variables, X_k, we could have variables such as the allocated supervision hours and the quality of statistics courses available. If adding only the number of weekly hours spent swimming, on the level-1, our model will look as in equation 6 (Sommet & Morselli 2017, p. 206).

$$Logit(odds) = B_{00} + (B_{10} + u_{jk} + v_{1k} +) * X_{ijk} + u_{0jk} + v_{0k} \quad (6)$$

Moving on to the case of the present study, I am analyzing the log-odds of *exit* (1) instead of surviving (0) during observations (i), of firms (j), in islands (k). This hierarchical structure, as stated previously, is a three-level model and thus also represented by equation 6. In summary, I have now explained how to interpret and analyze the three-level logistic regression of the present study. The random intercept variance, v_{0k}, will determine if there exists a difference between the islands in terms of firm-survival. Further, predictor variables will determine whether we accept the null hypotheses or the alternative hypotheses according to the literature chapter. The established description of the models will be built upon in a later discussion, the strategy of analysis.

3.3. Sampling strategy and sample

I collected and refined the database that constitutes the empirical data of this study, *Archipelago Organization Database*, solely for the present study. The collection resulted in 1511 organizations of all types, in the 17 core islands of Stockholm Archipelago between the years 2000-2019. Out of these, 1 341 firms are suitable for the principal survival analysis. In terms of firm-years for survival analysis, 10 082 total observations (years) are created for analysis. The criteria underlying data collection was to capture as much of the population of firms registered in the Archipelago as possible. As the sampling starts with a purpose in mind, and the sample is selected to include only firms registered in the archipelago, and exclude those that are not, it becomes a purposive sampling method. Furthermore, when the entire population meeting specific criteria is included in the research conducted, it is called Total Population Sampling (TPS) (Etikan 2016, p. 3).

Purposive sampling methods emphasize continuing to sample until no new substantive information is acquired. Therefore, the sample size is determined by data saturation (Etikan 2016, p. 4). Further, it should be noted that beyond the issues of left-truncation discussed earlier, there

may be some firms active in the core-islands but missing from the sample. For instance, it is not unheard of for a firm to be registered geographically far from the location of its operations.

3.4. Data collection

As mentioned previously, the purpose of the study is to analyze firm-survival on multiple levels of analysis, including the firm and island levels. Thereby, it becomes necessary to collect complete and appropriate data on all levels while seeking data saturation due to the sampling strategy. To this end, the empirical material in the study is based on multiple digital sources. For the firm-level, all data are from printed sources that will be mentioned below. However, for the island level, some parts of the material were accessed from unprinted sources by contacting municipalities and other relevant authorities.

3.4.1. Firm-level data collection

Finding appropriate sources of data

To start, I searched through two databases, Orbis and Retriever Business, that provide information on organizations[11] located in Sweden. However, both databases showed an unexpected decline in the number of Archipelago organizations regarding some of the years. Furthermore, Orbis seemed to lack the date of incorporation (official registration) for some organizations. These unexplained results wavered my confidence in how complete the data sources may be. Consequently, I searched for alternative sources of data and was provided with limited access to a database on organizations in Sweden, Bisnode Infotorg. Bisnode Infotorg, in turn, collects its information from the Swedish Tax Agency [Skatteverket], Swedish Companies Registration Office, Statistics Sweden [Statistiska Centralbyrån, SCB], and Swedish Courts [Domstolsväsendet] (Bisnode Infororg 2016, p. 5).

The database of Bisnode Infotorg makes it possible to search for organizations by several different search criteria, among these administrative areas such as the name of the area [*postort*], or the postal code [p*ostnummer*], in which the organizations are registered. For most islands in the study, the area name corresponds to the island name or a variation of it[12]. Some of the islands, however, share area names with neighboring islands or parts of the mainland and need

[11] In this section I will use the term organization, as opposed to "firm," since the dataset used for the study, Archipelago Organizations Database, includes information on all types of organizations in the core-islands of Stockholm Archipelago.
[12] For instance, Svartsö is actually called Skälvik, Gällnö is called Gällnöby, and Landsort is sometimes called Öja.

additional filtering, which will be discussed later. Thus, I searched and filtered by the appropriate area name for different islands. The results, however, were not readily available for download as a database file for limited users. Instead, data came in a table of links that lead to an information page for one organization at a time; this information can either be copied manually or downloaded as a pdf file.

Collecting and cleaning initial data

Downloading the information as pdf-files does not make the process of creating a database simpler than copying the information manually. Therefore, this option was ignored. However, manually copying the information for each organization could lead to random errors for which I have no way to account. Instead, I chose to repurpose a small software tool that made it possible to copy raw information on one organization at a time and save it in a text-file. The tool did not diminish the time nor the effort needed for writing down the data one letter at a time[13]. Instead, the time saved in this step is lost when systematically cleaning and reorganizing the data. However, it made it possible to remove the risk of making manual errors due to the human factor.

As soon as I had collected, then merged, organized, and cleaned the raw data from Bisnode Infotorg, I set out to find systematic errors due to the less than ideal nature of data collection. Errors were easier to find and remedy, as finding any error led to finding all similar errors. Furthermore, I conducted random tests on the database to make sure the information was accurate and corresponded to information on Bisnode Infotorg. Thereby, through Bisnode Infotorg, I had arrived at 2931 organizations, in 16 different organizational forms, registered between the years 1916 to 2020. To further encompass as much of the population as possible, I went back to the databases Orbis, and Retriever Business to see whether I could find any organizations missed by Bisnode Infotorg. If no more organizations can be added to the dataset, it would mean that data saturation had been reached.

Seeking data saturation

Retriever Business collects its information from the Swedish Tax Agency, Swedish Companies Registration Office, Statistics Sweden, and UC and makes its data readily extractable as excel files (Retriever Business, p. 2). Through Retriever Business, I managed to find a total of 1714 organizations using the same search parameter as that of Bisnode Infotorg. The organizations

[13] Roughly for each day spent collecting raw data, between two to three days, were spent on data cleaning.

were registered between the years 1936 and 2020, encompassing 12 different organizational forms. Once the databases had been organized similarly in order to facilitate a merger, only 123 organizations of 1714 were found to be unique (the rest had already been found using Bisnode Infotorg). Orbis, on the other hand, was not used as it failed to find new organizations. Thereby, I deemed that sample saturation was achieved, resulting in a total of 3054 organizations encompassing the period between 1916 to 2020.

Final data cleaning

As mentioned briefly, filtering away organizations that are situated within the same administrative area but not within the geographical boundaries of the islands has been the inconvenience of the data sources. Data sources usually take into consideration the administrative division of a region, not the geographic division. For instance, the postal code of Blidö includes many smaller islands, including Gräskö. Ramsö, a small island near the center of Stockholm, is found within a postal code that it shares with other areas in the municipality of Vaxholm. Finally, as with Ramsö, Landsort had a similar situation within the municipality of Nynäshamn. Thereby, to remove organizations outside of the islands, I searched the address of each organization on different digital maps available online[14]. The organizations falling outside of the physical boundaries of islands were given a specific code so that they may be easily filtered away.

The number of organizations, thus, arrived at 2979, encompassing all the core-islands between the periods of 1916 to 2020. Further, out of these, 1511 organizations were created during the 2000-2019 time period. However, only 1341 organizations with legal[15] forms where information is complete allow for the principal statistical analysis. For other forms of legal structures, the exit date could not always be sure. Therefore, these organizations will only be used to develop contextual variables. Finally, I changed the exit date for several firms to reflect reality better. Particularly, sole proprietors that have not had any updates in the Swedish Companies Registration Office for ten years; are sent a letter asking if they still conduct their activity. If no answer is provided or the firm owner confirms the operations have ceased, the firm is given an exit date (Swedish Companies Registration Office 2015). The exit date, however, is

[14] Specifically, OpenStreetMap, Google Maps, and Hitta.se. For many firms OpenStreetMap provided a list of road names, making it possible to create dynamic filtering within Excel.
[15] For the following legal forms data was complete and accurate: sole trader, joint-stock companies, trading partnership, limited partnership and economic association. However, for other types of organizations, such as non-profit associations the exit date did not always show accuracy, making it difficult to conduct appropriate survival analysis.

purely administrative. Keeping in mind that these firms had not been active for roughly ten years, I removed ten years from the exit dates in order to get closer to the real exit date.

3.4.2. Industry variables

SNI 2007 (henceforth SNI) stands for Swedish Standard Industrial Classification 2007 [Standard för svensk näringsindelning 2007] and is the revised version of former SNI classifications. SNI builds on NACE[16], the statistical classifications of economic activities in the European Union. (Statistics Sweden 2007a, p. 10; Eurostat 2008, p. 13). The difference, in short, is that NACE has four levels, while SNI adds a fifth level to the classification. The first level is a letter code designating the section. For instance, manufacturing has the letter "C." The second level is that of the division, which is a two-number code, e.g., food manufacturing has the code "10". The third level is a three-number code referring to the group, e.g., "10.5" is for the manufacturing of dairy products. The fourth level (the last level for NACE) refers to the class. For instance, "10.52" is the code for the manufacturing of ice cream. Finally, the fifth level, found only in SNI, adds additional detail if the fourth level is not descriptive enough (Statistics Sweden 2007a, p. 13).

By the end of the firm-data collection, almost precisely 60% (802) of the 1341 firms of the sample had SNI codes available, leaving 40% of the sample without codes. Roughly 700 of these firms had a written statement in Swedish, describing the economic activity of the firm. Thus, using the SNI manual (Statistics Sweden 2007a) as well as a webpage provided by Statistics Sweden (2007b), I categorized all these 700 firms according to the division level of SNI classification. However, at the end of this process, many firms were found to have significantly diverse activities. Therefore, it made better sense to classify them according to three broader categories: trade, service, and manufacturing.

Furthermore, this categorization of firms has previously been used in similar studies (e.g., Box 2008, p. 384). Many issues were thus solved. For instance, an entrepreneur that repaired vehicles and provided transport could now simply be categorized as a service provider. For some firms, the activities were even more diverse and did not neatly fall into these categories. For these firms, I simply chose categories that best fit the overall description and name of the firm.

3.4.3. The island-level data collection

For the island level, I needed data that could comprehensively describe the context of the islands. Further, as the study employs a longitudinal design, almost all variables on the island

[16] NACE in turn builds on the international ISIC Rev 4 (Eurostat 2008, p. 13).

level need, therefore, to *vary with time* (except land area and distance to market). To this end, information from several sources was used to collect and establish all the variables. Most island level variables thus contain 340 unique measurements (20 years multiplied by 17 islands). Further, as there are 13 unique variables on the island-level, out of whom 11 are time-varying, in total, I collected 3774 unique island data-points for the years 2000-2019.

Population density

Estimates of the growth of the human population are invaluable, as they are used for research and operational applications, including providing essential data to study the labor force and entrepreneurial activity (Millán, Román & Congregado, p. 94). However, obtaining reliable population estimates at the spatial resolutions required for the present study was a significant challenge. Population data in Sweden is often provided at broader administrative divisions[17], none of which solely encompasses geographical aspects, such as the islands of the study. In a way, most of the data provided by Statistics Sweden, vanish the islands into the broader municipalities in which they are situated.

To solve the issue, I contacted the municipalities in which the islands are located. For half the islands, it was possible to gain access to appropriate data. However, for the other half, there were many years in which information was lacking or limited by broad geographical division. Thus, for some of the islands, the provided data had numbers well beyond what secondary, printed, sources were indicating. Gällnö and Ramsö, in particular, were difficult to distinguish from neighboring archipelago islands.

Furthermore, all municipalities (except Haninge), could only provide island inhabitant statistics for a few years. For these municipalities, several sources from Statistics Sweden (2014, p. 7, 2019a), and County Board of Stockholm (2012, p. 8, 2016, pp. 14–20) provided more information. However, once all information was collected, I deemed it best to establish dummy (categorical) variables for the population variable since no statistical imputing method could account for the vast number of years for which no exact inhabitant numbers were available.

[17] To see available divisions for statistics in Sweden see Statistics Sweden (2020).

Island service access, infrastructure, and distance to the urban center (market)

For the study, it was essential to have indicators reflecting the level of service and infrastructure available on each island every year. However, such an indicator is not readily available. Therefore, I used several variables as a proxy for the level of service available on an island in any given firm-year. The first two variables are indicators of availability of preschools and primary level schools (class 1-5). The third variable indicates the availability of year-round grocery stores with full service. The last variable is an indicator of broadband fiber infrastructure in the islands.

Several different sources have been used to collect the above information. For education accessibility, websites of the municipalities were used, together with newspaper articles and regional authorities and local websites. For instance, the island of Utö has a school from kindergarten level to class nine elementary (Värmdö Municipality 2018). However, only through a newspaper article describing a fire accident, it was possible to know that the school had been there since 1886 and inactive during 2003 before being rebuilt the year later (Gyllenberg & Maluszynski 2004). Likewise, only through the local library website, I was able to know that Nämdö school had been closed since 2007 due to a lack of school-aged children (Nämdö Library 2020).

Similarly, for the it-infrastructure and existence of year-round grocery stores, I used websites of local associations (Nämdö Fiber 2020; Ornö Fiber EF 2020), Island by Island [Ö för Ö] (2020) and County Board of Stockholm (2012, 2014, 2015, 2016, 2017, 2019). The measures of distance in time (minutes) to the nearest urban center were found on the website of Island by Island (2020). Island by Island is a website for an archipelago inventorying project that provided excellent information on many of the above variables.

Island activity

Archipelago municipalities [skärgårdskommuner] have a significantly higher number of visitors and inhabitants during the summer season (County Board of Stockholm 2016, p. 32; Värmdö Municipality 2020). An increase in visitors is assumed to influence the island and the organizations within. Therefore, having a variable measuring the number of visitors to the islands could be beneficial. To quantify this measure, I compiled Waxholmbolaget[18] passenger-

[18] Waxholmbolaget is responsible for the county's public transport on water (Region Stockholm 2020).

data obtained from the department of Sea Traffic [Sjötrafiksektion] at the Public Transport Administration [Trafikförvaltningen] (Region Stockholm 2020). The data compilation led to the construction of three variables—two variables on the number of passengers during the off-season and the high-season, and a third variable on the ratio of off-season passengers compared to high-season passengers. Further, as Waxholmbolaget is not the only means of transportation to the core islands, the first two variables may not be significant due to other transportation actors being more active in individual islands. However, the proportion (ratio) of the off-season and high-season passengers is probably more or less the same for all actors. This third variable should then be significant if the real number of passengers affects the likelihood of survival at all.

Institutional thickness

Measuring institutional thickness is not without issue and depends on the construction of creative measurements. Prior studies have used many different proxy variables to quantify institutional thickness, among these the number of investment incentives, local authorities, SME investment, number of trade associations, number of NGOs (Ersoy 2018, p. 111). Many of these proxy measurements are challenging to gain access to, especially considering the more extended time perspective of the present study. However, as stated previously, the *Archipelago Organizations Dataset* includes information on many non-profit associations, foundations, and economic associations in the islands of the study. Thus, I created a variable that sums the number of these associations for each year and island. I deem that the variable is particularly suitable, considering the definition of institutional thickness presented in the literature framework. Finally, I am aware of the debate concerning the difference between the two concepts of organizational thickness and institutional thickness. Even if my operationalization of institutional thickness may be argued to be of organizational thickness, scholars find that the two concepts are sufficiently interrelated that they may represent each other (Zukauskaite, Trippl & Plechero 2017, pp. 329–331). In practice, a higher number of associations is thus argued to be an indicator of institutional thickness.

3.4.4. Macroeconomic variables

I have included variables that test for the effect of changing regional macroeconomic conditions. Stockholm as a region is higher on the hierarchical grouping than the islands. Strictly speaking, the study is not comparing between different regions. Therefore, macroeconomic variables, simply, enter the model on the island-level. The first variable is a time-lagged (1 year)

continuous variable measuring the regional GDP of Stockholm. For this variable, data was only available for the years 2000-2017 (Statistics Sweden 2019b). In practice, this means that for any given year, the variable represents last year's regional GDP. Thus, the most recent year available (2017) moves forward by a year to 2018. Therefore, subsequent years, 2019 and forward, become missing values. The second variable is a dummy variable for years when the regional Gross Domestic Product (GDP) decreased significantly and stayed low for consecutive years. For the Stockholm regional GDP, a significant fall takes place shortly after the year 2007, whereby the change in regional GDP drops to 0.3% or lower for two consecutive years, 2008 (0.2%)-2009 (0.3%). This variable then, in practice, functions as an indicator of the global financial crisis (Türkcan & Erkuş-Öztürk 2019, p. 10).

3.4.5. Summary of study variables

Altogether, the collection and transformation of data have resulted in two sets of variables. The first set is those variables that may explain exit/survival on the firm-level. The age variable is the only measure on this level that varies with time and therefore is on the observation level. The second set of variables is that of the island level; these variables all vary with time and place except for the macroeconomic variables, which only vary with time. Below are the summary tables of these variables.

Table 1:Summary table of the firm-level variables.

Variable	Operationalization
Firm-exit	Dependent variable that equals "1" if the firm ceases its activities. Dummy variable equals "0" if operations continue during the period of observation. Thus, the values of "0" are censored.
Age	Time-varying variable starting from "1" during the first year of firm activity, and increase by one for each subsequent year.
Age (cluster-mean-centered)	The mean-centered, age variable for multilevel analysis. See section on analysis strategy for more detail.
Sole Proprietorship	Dummy variable that equals one if the firm has the legal structure of Sole Proprietorship. Dummy variable equals "0" if the firm has any other legal structure.
Joint-stock Company	Dummy variable that equals "1" if the firm has the legal structure of Joint-stock Company. Dummy variable equals "0" if the firm has any other legal structure.
Trading/Limited Partnership	Dummy variable that equals "1" if the firm has the legal structure of Trading or Limited Partnership. Dummy variable equals "0" if the firm has any other legal structure.
Trade	Industry dummy variable that equals "1" if the firm belongs to the trading industry. Dummy variable equals "0" if the firm belongs to any other industry.
Manufacturing	Industry dummy variable that equals "1" if the firm belongs to the manufacturing industry.
Service	Industry dummy variable that equals "1" if the firm belongs to the service industry.

Table 2:Summary table of the island level variables

Variable	Operationalization
Low population density	A time-varying dummy variable taking values of "1" if the number of permanent inhabitants on the island for the year is equal to or smaller than 100. Otherwise, the variable takes values of "0".
High population density	A time-varying dummy variable taking values of "1" if the number of permanent inhabitants on the island for the year is larger than 100. Otherwise, the dummy variable takes values of "0".
Distance to urban center (market)	A continuous variable that measures the amount of time, in minutes, to the closest urban center.
Island area size	The variable measures the size of the island, in ha.
Fiber broadband	A time-varying dummy variable that equals "1" if the island has a functioning fiber-network during a firm-year. Dummy variable equals "0" if no fiber was reported during the firm-year.
Primary School (access)	A time-varying dummy variable that equals "1" if the island has an active school ranging from kindergarten to fifth class primary during a firm-year. The variable equals "0" if no school was reported active during the firm-year.
Kindergarten (access)	A time-varying dummy variable that equals "1" if the island has an active kindergarten during a firm-year. The variable equals "0" if no kindergarten was reported active during the firm-year.
Grocery store (access)	A time-varying dummy variable that equals "1" if the island has a year-round grocery store with access to postal services and during a firm-year. The variable equals "0" if no such grocery store was reported active during the firm-year.
Island activity (Passengers in) off-season	A time-varying variable of the number of passengers transported to the island by Waxholmsbolaget from January to March and October to December
Island activity (Passengers in) high season	A time-varying variable of the number of passengers transported to the island by Waxholmsbolaget from April to September.
The ratio of high-season passengers to off-season	A time-varying rate of high-season to off-season passengers. $Ratio = \frac{No.high-season\ passengers}{No.off-season\ passengers}$
Institutional thickness	The cumulative number of associations and non-profit organizations on each island, starting from the year 2000. The source for the information is the *Archipelago Organizations Dataset*.
Regional GDP (in 100 billion)	Lagged, by a year, time-varying and continuous variable measuring the regional GDP of the Stockholm region in hundreds of billions. $\frac{Regional\ GDP\ in\ SEK}{100\ 000\ 000\ 000}$
Financial crisis	Time-varying dummy variable that equals "1" for years when the volume change of the Stockholm GDP is 0.3% or less. The variable equals "0" when the regional GDP change in volume is 0.3% or lower. Precisely, the years of 2008-2009 correspond to this operationalization, and therefore, the variable in effect becomes a dummy variable representing the years of the 2008 financial crisis.

3.5. Analysis strategy

So far, I have argued for the use of multilevel modeling, described, and discussed the design of the study, explained multilevel logistic regression, accounted for the collection of data, and

provided a summary of the variables and their operationalization. Before moving on to a discussion on reliability, validity, and ethics, I will discuss the strategy of multilevel logistic regression analysis following the data collection.

3.5.1. Data preparation

Survival analysis of longitudinal data requires the dataset to be transformed from a wide-format to a long-format. In practice, this means that instead of each subject (firm) having one row with all the variables, each subject should have one row for *each* observation. Thus, each independent firm has between 1 to 20 rows of data depending on how long it survived. For instance, a firm established in 2000 and survived the entire study has 20 rows of observations, while a firm established in 2005 and exited in 2010 has exactly five rows of observation. To reshape the data, first, I moved it from the Excel table in which data collection took place, to the STATA statistical software package. Then, in STATA, I used the *expand-and-sort* procedure as described by Rabe-Hesketh & Skrondal (2012, pp. 752–753).

Sommet & Morselli (2017, p. 2011) recommend that researchers mean-center their independent variables when performing a multilevel logistic regression. Further, for longitudinal study designs with several observations by subject, the authors recommend to cluster-mean-center level-1 variables. The argument is that mean-centering makes it possible to estimate the pooled within-subject fixed effects. Centering is explained to facilitate interpretation of the model since the fixed intercept becomes the log-odds that the independent variable equals one when all the independent variables are at their mean. In the present study, the age of firms is the only level-1 independent variable that varies with time. Thus, in the multilevel logistic regression, the age variable will be mean-centered. However, this will bring the issue of analyzing the average age of the firm, not the actual age (Sommet & Morselli 2017, p. 214). Therefore, a viable solution will be to use a hazard estimate graph to analyze this variable.

3.5.2. Intra-class correlation and beyond

Following the preparation of data, the next step is to analyze an intercept-only model. The intercept-only model is also called an “empty” model as no predictor variables are included, only the fixed and random intercepts. Previously, I described this model during multilevel logistic regression (equation 5). The purpose of analyzing the intercept-only model is to measure the intra-class correlation (ICC) (Hox, Moerbeek & Schoot 2018, p. 44). ICC is an indicator of the total variance explained by the structure of the population. In other words, it calculates how large the log-odds variation between having a “0” or “1”, firm-exit or survival, is explained by either the level-1, -2, or -3 (Hox, Moerbeek & Schoot 2018, p. 13). As explained previously in

the discussion of multilevel modeling, ignoring the correlation between the outcome for subjects in the same context results in falsely significant predictors (Hox, Moerbeek & Schoot 2018, p. 4).

The range of ICC is between 0 and 1. A value of 0 means that the residuals have absolute independence. In practice, this means that the higher-level intercepts are not needed, and therefore multilevel modeling is unnecessary. Moderate to high values of ICC with a significant intercept-only model warrant a multilevel-approach (Sommet & Morselli 2017, p. 212). If the geographical context has a strong influence on outcomes, such as exit or survival of firms, the ICC should be high. However, the contextual level ICC, for instance, the island level, is low when the sample size is large. Further, ICC will also be low if the neighborhoods from which the sample is drawn are not sufficiently different (Merlo, Wagner, Austin, Subramanian & Leckie 2018, pp. 34–36). Following the ICC calculation, researchers are recommended to add their predictor variables one level at a time and experiment with different types of interaction effects, all the while comparing the different models to arrive at a model that explains most of the variation (Sommet & Morselli 2017, p. 2013; Hox, Moerbeek & Schoot 2018, p. 44).

3.6. Validity, reliability, and source criticism

Reliability and validity are concepts used to describe and assess the quality of a study. The terms have different meanings in quantitative and qualitative research. In quantitative research, reliability refers to the repeatability of the findings. In essence, a study that is possible to replicate, with time-stable measurements, and easily comparable, reaches high reliability. The concept of validity is a question of whether the research is measuring the intended phenomenon. Validity is thus, a question of accuracy and credibility (Golafshani 2003, pp. 598–599).

Further, as I have used diverse sources of information to collect the data necessary, one could argue that the study incorporates subtle hints of qualitative methodology. Thus, assessing what is probable and generally being beware of sources of information (source criticism) is essential. Thurén (2003, p. 21) explains that whereas it is difficult to set rules for conducting source criticism, in general, users of information should assess the following: authenticity, time, dependency, and tendency (bias).

I have described in detail the process and sources of data collection, as well as the statistical techniques and strategies. This transparency should make it possible for other researchers to replicate the study, and thus increase the reliability of the study. Further, I have to a high degree made use of constructs established in numerous firm-survival studies. Moreover, I have used

data-triangulation techniques commonly used in qualitative methodology to improve the validity and reliability of the study (Golafshani 2003, p. 603). Essentially, by using multiple sources for the same information, it becomes possible to obtain more accurate and reliable measurements on which to construct the variables of the study.

Moreover, I made sure to use information as close in time and place to the relevant phenomenon as possible. For instance, to construct the variable of population density, I made use of published statistics by Statistics Sweden, as well as contacted the municipalities of the islands. Similarly, to construct the contextual variables, such as access to fiber broadband, I made use of information available by the Island by Island[19] project as well as webpages of local fiber associations and archived news articles on the islands. Thus, I deem that the study altogether is reliable and valid.

3.7. Ethical considerations

Research with a geographic aspect brings important ethical considerations as it affects study design, collection and analysis of data, and results presentation. Ethical considerations for lower levels of analysis are mainly concerned with complying with laws and standards concerning issues such as data protection, confidentiality, and security (Griffith 2008, pp. 238–240). In the case of this study, a firm-level of analysis, especially of Sole Traders, is critical. To meet these ethical standards, first, I collected all data on the firm-level on my personal computer. This method ensures that the database is only available on devices known by me. Second, the data has only been shared with my supervisor due to sensitive information such as the national identification number of individuals. Third, in the later stage of the study, a new data file was created, in which sensitive information was deleted, and subjects were given a unique identifier. Thereby, information on the final dataset for analysis is effectively anonymized and is difficult to be traced to any specific individual or firm without access to the original file.

[19] Island by Island (Ö för Ö) is an on-going inventorying project by the County Administration Board of Stockholm of the core-islands in Stockholm Archipelago (Island by Island 2020)

IV. Results & Analysis

The chapter will present the statistical analysis of the study and its results. The chapter begins with frequency, descriptive, and correlation analysis. Next is an analysis of the multilevel model criteria, followed by the principal statistical analysis of the study, multilevel logistic regression. The statistical analysis is then followed by a hazard estimation analysis of the firm-age. The chapter then closes with hypothesis testing and a summary of results.

4.1. The islands

The data collected in the *Archipelago Organizations Dataset* enables a more in-depth analysis of the entrepreneurial context of the islands. Specifically, it is possible to analyze the firms as well as the islands. Thus, in practice, this section provides an opportunity to develop a deeper understanding of the islands to reach a more in-depth analysis.

Table 3: The context of the islands and their firms

Islands	Population density	Land area (ha)	Fiber access year	Grocery store access (2019)?	Active primary school (2019)?	Active kindergarten (2019)?	**No. associations	***No. firms (2000-2019)	% firm-exit (island)
Arholma	Low	541	-	Yes	No	No	1	21	47.6
Blidö	High	2223	2018	Yes	No	Yes	34	176	36.4
Gräskö	Low	120	2014	No	No	No	2	8	25.0
Gällnö	Low	540	-	No	No	No	2	25	40.0
Ingmarsö	High	612	2013	Yes	Yes	Yes	5	45	26.7
Landsort	Low	173	2006*	No	No	No	4	13	61.5
Ljusterö	High	6160	2015	Yes	Yes	Yes	50	481	45.3
Möja	High	1285	2015	Yes	Yes	Yes	15	75	42.7
Nämdö	Low	1046	-	Yes	No	No	4	33	30.3
Ornö	High	4767	2019	Yes	Yes	Yes	9	77	40.3
Ramsö	Low	105	-	No	No	No	2	24	41.7
Runmarö	High	1353	2018	Yes	Yes	Yes	7	81	38.3
Sandhamn	High	224	-	Yes	No	Yes	5	57	29.8
Svartsö	Low	715	-	Yes	Yes	Yes	3	47	21.3
Tjockö	Low	293	-	Yes	No	No	2	13	7.7
Utö	High	2936	2017	Yes	Yes	Yes	14	77	45.5
Yxlan	High	1695	-	Yes	Yes	Yes	18	88	39.8
Total	-	-	-	-	-	-	-	**1341**	**40.0%**

*ADSL during 2006-2017. No broadband today. Source: Archipelago Organizations Dataset
**Number of associations founded between 2000-2019 and active by the end of 2019.
***Survival sample (observations)

It would be challenging to present data for longitudinal changes on the islands, as 20 years multiplied by 17 islands amounts to a quite large table. However, the above table provides the most relevant information (table 3). The population density of the islands is roughly equal, with eight islands having a population of less than 100 permanent inhabitants and the other nine

islands having more than that. However, in practice, there is significant variability between the islands in each category. The land area (ha) of the islands is significantly different, with the smallest island being 105 ha (Ramsö), and the largest island has an area of 6160 ha (Ljusterö). More than half of the islands lack fiber broadband infrastructure today, and islands with access do not always lead to full coverage. Landsort/Öja is a particular case as it had ADSL broadband access since 2006. However, the technology is no longer supported, and thus the local network has been shut down since 2017.

The majority of islands have local access to year-round grocery store access, which is quite different from islands that have not been designated "core-islands" (Island by Island 2020). However, local access to primary schools and kindergartens is even rarer. Furthermore, going back in time on the *Archipelago Organizations Dataset* would show that these services and infrastructure were even rarer just a few years back. However, while the situation on some islands has progressed, other islands have not been fortunate. The last two columns present the sample allocation and the percentage of firms that went through the event of exit during the time of analysis. The number of firms per island ranges from 8 to 481. Thus, there are substantial differences between islands.

Further, while the overall percentage of firm-exit is 40% of the sample, on an island-by-island basis, there are significant differences, ranging from 7.7% to 61% mortality. Moreover, some islands with low population density come quite close to have the same number of firms as islands with higher density. Similarly, the number of associations on each island seems to follow the number of firms operating. However, there are some differences concerning proportion. Finally, while the above table has provided much information on the islands, it does not tell us enough about firm-survival, as the number of years a firm manages to survive is essential.

4.2. Descriptive analysis

The purpose of this section is to develop some intuition of the underlying variables of the data. Thus, descriptive statistics will be used (table 4). First, variables that have minimum values of 0 and maximum values of 1 are all dummy variables. A comparison between the summary tables at the end of the methodology chapter and the above descriptive table should confirm this assumption. The number of observations for most variables is 10082, the expected number of firm-years analyzed. However, four variables reach a lower count of observations, regional GDP, and industry variables. Regional GDP for the year 2019 was declared missing in the

methodology chapter, thereby explaining the missing observations. Similarly, for the three industry variables, I explained in the methodology chapter that a small percentage of firms lacked SNI codes as well as a description of the firm's activity. In conclusion, all missing observations come from expected sources.

Table 4: Descriptive table of study variables

Variable	Obs	Mean	Std.Dev.	Min	Max
Firm-exit (outcome variable)	10082	.053	.224	0	1
Age	10082	6.092	4.432	1	20
Age (cluster-mean-centered)	10082	0	3.559	-9.5	9.5
Joint-stock companies (reference)	10082	.299	.458	0	1
Sole proprietorship	10082	.581	.493	0	1
Trading/limited partnership	10082	.1	.3	0	1
Economic association	10082	.02	.14	0	1
Service industry (reference)	9687	.779	.415	0	1
Trade industry	9687	.098	.297	0	1
Manufacturing industry	9687	.124	.329	0	1
Small island population	10082	.175	.38	0	1
Large island population	10082	.825	.38	0	1
Distance to urban center	10082	76.062	24.412	25	127.5
Land area of island	10082	3176.356	2373.827	105	6160
Fiber broadband infrastructure	10082	.253	.435	0	1
Access to Primary School	10082	.712	.453	0	1
Access to kindergarten	10082	.84	.367	0	1
Access to a year-round grocery store	10082	.94	.238	0	1
Off-season Passengers	10082	6451.317	6209.36	0	31421
High-season Passengers	10082	19238.46	16202.7	0	79390
Ratio of summer passengers	10082	7.445	13.862	.131	91.429
Institutional thickness (number of associations)	10082	17	16.203	0	50
Regional GDP	9207	11.265	2.206	6.500	14.396
Financial crisis	10082	.085	.279	0	1

Source: Archipelago Organizations Dataset

The low mean value (0.053) of the outcome variable, firm-exit, is explained by how event-history data is constructed. Subjects have a value of 0 when the event is censored and one value of 1 when the event takes place. Thus, most observations of the outcome variable have values of 0. Similarly, sole proprietorship has the highest mean value of the legal structure variables as it is the most common legal form on the observation[20] level. Sole proprietorships are then followed by joint-stock companies, trading and limited partnerships, and economic associations. Moreover, the service industry is the most prevalent sector, followed by manufacturing and trade.

Two particularly exciting variables to compare are the age and mean-centered age variables. The age variable has minimum values of 1 and maximum values of 20, meaning that the youngest firm analyzed attained the age of 1 year while the oldest firms of the sample attained the age of 20 years. These values correspond to the chosen time-frame of the study, whereby the oldest

[20] The mean values are not on the firm level but on the observation level. In short, this means that

firms that survived to the year 2019 were founded the year of 2000. Furthermore, the mean of the variable is 6.092, which means that the average age attained of firms was roughly six years. The mean-centered age variable shows a mean value of 0. This result shows that the variable indeed is mean-centered, which in practice means that the mean age of firms should take on values of 0, with the maximum values falling above 0 and the minimum values below 0. This assumption is correct as the variable's minimum value is -9.5, and the maximum value is 9.5.

Finally, distance to urban centers provides a puzzling finding. While the "core-islands" are defined as islands that have a distance of more than 45 minutes to the nearest urban center, the minimum value of the variable is 25 minutes. This value corresponds to the travel distance for the island of Ramsö, which is the only island with a travel distance below 50 minutes. Despite the authorities' definition of core-islands, there is as far as I know no motivation for the island's inclusion.

4.3. Correlation and multicollinearity

An essential consideration for models with multiple predictor variables is multicollinearity. In essence, a strong correlation ($r > 0.8$) between two or more predictors leads to modeling issues and low model fit (Field 2017, p. 533). The correlation matrix in Appendix I, presents the bivariate correlation of all variables in the study. Three sets of variables show a high correlation with each other. First, the age variable highly correlates with the mean-centered age variable ($r = 0.8$). The high correlation is not surprising as the second variable is a transformation of the first. Further, these variables will not enter the logistic regression models together as that would defeat the purpose of mean-centering the age variables. Second, the dummy variables of population density have a perfect negative correlation ($r = -1$). This result is not surprising as the two variables are the exact opposite of each other. Further, only one of the variables will enter the model. Third, the number of winter and summer visitor variables are highly correlated ($r = 0.83$). These variables will thus not enter the multilevel logistic regression. Altogether, beyond the three mentioned exceptions, there are no other issues of collinearity to be detected on the correlation matrix.

Correlation analysis of the outcome variable and the predictor variables can now be conducted. From the matrix, we can see that out of the 23 predictor variables, 15 are significantly correlated with the outcome variable. To start, on the firm-level, firm age is negatively correlated with firm-exit, increasing age is thus more correlated with survival. The cluster-mean centered age variable, while having the strongest correlation to failure, is positively related to firm-exit. As

this variable has been cluster-mean transformed due to the multilevel logistic regression, it essentially means that a firm with an average age, six[21] years, is correlated with firm-exit. It is vital to make a note of this finding for the logistic regression analysis, as the age variable will need further analysis to establish the age-survival hazard better[22].

Moreover, the dummy variable for joint-stock companies has a strong, significant, and negative relation to failure. Firms with joint-stock legal structures, therefore, are correlated with survival. Economic associations, while negatively correlated to failure, the variable is insignificant on the 0.05 level. On the other hand, sole proprietorships and trading and limited partnerships have a positive and significant correlation with failure, and these legal structures are then more correlated with failure. The service industry variable is the second non-significant variable of the firm-level variables. The non-significance could be due to the broad nature of the category compared to trade and especially manufacturing. Indeed, both the trade and manufacturing industries are significant. However, while manufacturing is negatively related to failure, trade has a positive correlation. Thereby, manufacturing firms are correlated with survival.

Moving on to the island level variables, especially the two variables of the number of passengers to the islands, have the weakest and most non-significant correlation to firm-exit. However, as established previously, these variables will not be included due to probable collinearity issues. The ratio (proportion) of high-season passengers to that of the off-season passengers shows a significant and moderately strong relation to failure. In effect, this means that higher proportions of summer guests correlate with firm-exit. Access to primary school and fiber broadband infrastructure on the island is positively and significantly correlated with firm-exit. However, access to kindergarten is slightly weaker, and year-round grocery stores show a non-significant effect. Distance to nearest urban center is significant but negatively correlated to firm-exit. The variables of institutional thickness and financial crisis are positively correlated to failure but non-significant. Finally, regional GDP is significant and negatively correlated with failure.

[21] See the descriptive analysis.
[22] My reasoning is that analyzing the effect of *average* age of firm on survival does not untangle the true relationship between longitudinal firm age and probabilities of survival; examining the average age beyond making it difficult to discuss the findings of the present study to those established in the literature, it further makes it difficult to provide implications for policy.

4.4. Does the multilevel logistic regression model hold?

Designing a multilevel study is not always awarded with a significant model when the time of analysis is due. Instead, researchers must test the hierarchical multilevel model for actual statistical significance as well as reaching the set requirements.

4.4.1. Sample size

The sample size of a study is a critical aspect in statistical analysis as it affects statistical power and is an indicator of a well-performed study (Świątkowski & Dompnier 2017, pp. 112–113). However, in multilevel modeling, of more importance is having a sufficient number of groups at the second level of the model. More precisely, at least 50 level-2 groups are needed to estimate the standard errors correctly (Maas & Hox 2005, pp. 90–91). The present study achieves 10 082 observations at the level-1, 1 341 groups (firms) at the level-2, and 17 higher-level groups (islands) at the level-3. Due to the repeated measurement nature of longitudinal data, even islands with few firms reach a sufficient sample size. Thereby, the study fulfills the required sample size for a multilevel model by a decent margin.

4.4.2. The intercept-only model

Table 5: Intercept-only multilevel model.

y	Odds. Ratio	p-value	Sig.	St.Err	z-value	[95%Conf	Interval]
Fixed-part							
Constant (level-1)	0.046	0.000	***	0.004	-31.95	0.038	0.055
Random-part	(estimate)						
Constant (estimate level-3)	0.043	0.000	***	0.392	.	0.007	0.258
Constant (estimate level-2)	0.595	0.000	***	0.207	.	0.301	1.178
Log-likelihood		-2074.403		Random-part chi²		16.76	
Number of observations		10082		Ran. Prob>Chi²		0.000	
Prob > Chi²		0.000					

**** p<0.001, ** p<0.05* Source: Archipelago Organizations Dataset

Table 6: Intra-class correlation (ICC)

Level	ICC	St. Err.	[95% Conf.	Interval]
Level-3 (island)	0.011	0.009	0.002	0.062
Level-2 (Firm)	0.162	0.045	0.092	0.270

Source: Archipelago Organizations Dataset

As discussed in the methodology chapter, a significant intercept-only model with a moderate to high ICC value means that the model violates the inter-dependency requirement within groups, and thereby warrants a multilevel approach. The results of table 5, show a level-2 ICC of 0.162, which indicates that differences between-firms, explain 16.2% of the firm-exit variance. Further, level-3 ICC of 0.011 indicates that 1.1% of the firm-exit is explained by between island differences. Conversely, this means that within-firm differences explain 82.7% of firm-exit var-

iability. An ICC of 0.162 is moderately high. Thus, the modeling warrants a multilevel approach. As the level-3 ICC is so low, we could omit it. However, as discussed previously, when the neighborhood context is not sufficiently different, for instance, analyzing a core-region against a peripheral-region, ICC values are expected to be lower.

Further, larger sample sizes lower the ICC. Finally, the intercept-only model is significant, with both the level-2 and level-3 effects being significant at less than 0.01 level. Thus, the multilevel model holds, and I can now continue to perform the multilevel logistic regression.

4.5. Multilevel logistic regression analysis

Table 7: Models I-III of multilevel logistic regression with the firm, island, and macroeconomic predictors.

Model	I	II	III
Fixed-part			
Age (cluster-mean centered)	1.839 (0.152) ***	2.307 (0.241) ***	3.011 (0.466) ***
Joint-stock companies (ref)	-	-	-
Sole proprietorship	66.395 (36.108) ***	41.053 (21.024) ***	16.101 (8.136) ***
Trading/limited partnership	111.394 (72.214) ***	61.025 (36.886) ***	14.137 (7.902) ***
Economic association	9.216 (8.834) **	7.458 (7.305) **	5.081 (5.038)
Manufacturing industry (ref)	-	-	-
Trade industry	16.024 (9.677) ***	11.895 (6.897) ***	5.295 (2.920) ***
Service industry	8.087 (3.817) ***	7.712 (3.578) ***	4.048 (1.779) ***
Large population density		5.630 (3.320) ***	3.404 (1.754) **
Distance to urban center		1.013 (0.010)	0.998 (0.008)
Island land area		1.001 (0.000) ***	1.000 (0.000)
Access to fiber		0.420 (0.102) ***	0.747 (0.204)
Access to primary school		2.661 (1.511) *	1.734 (0.744)
Access to kindergarten		0.328 (0.179) **	0.511 (0.265)
Access to a grocery store		0.156 (0.128) **	0.607 (0.447)
Ratio of summer passengers		1.022 (0.008) ***	0.998 (0.007)
Institutional thickness		0.917 (0.016) ***	0.991 (0.014)
Regional GDP			0.407 (0.052) ***
Financial crisis			2.927 (0.804) ***
Constant (level-1)	0.000 (0.000) ***	0.000 (0.000) ***	2.036 (2.667)
Random-part (estimate)			
Constant (estimate level-3)	0.138 (0.176) ***	0.117 (0.195) ***	0.000 (0.000) ***
Constant (estimate level-2)	8.271 (2.167) ***	7.925 (2.156) ***	5.673 (2.044) ***
Log-likelihood (model fit)	-1522.200	-1448	-1178.120
Number of observations	9687	9687	8831
Prob > Chi^2	0.000	0.000	0.000
Random-part Chi^2	125.02	96.37	28.99
Random-part Prob>Chi^2	0.000	0.000	0.000

**** p<0.01, ** p<0.05, * p<0.1* Source: Archipelago Organizations Dataset

Table 7 summarizes the results of multilevel[23] regression modeling. The first model, Model I, includes all the variables on the firm level, firm age, legal structure, and industry affiliation. The number of observations is 9687, which was expected as missing values were detected for the industry variables during descriptive analysis. These values were not imputed using statistical methods as no imputation method can account for the correct industry. The whole model

[23] For an overview of *simple* logistic regression results, see Appendix II.

is significant at less than the 0.001 level. Furthermore, the random variance intercepts are also significant at less than the 0.001 level, both at the variable level and modeling level. The log-likelihood now has a more substantial value than the intercept-only model (table 4), meaning that the variables on the firm-level explain more of the variation between firms that fail and those that survive compared to the intercept-only model.

Keeping in mind that the age variable in the multilevel logistic regression has been mean-centered, an OR value of 1.84 signifies that a firm with an average age, six years in this study, is 1.84 times more likely to exit instead of surviving, of course, this result does not tell us much about the likelihood of exit as the firm ages. To solve this issue, I argued previously that a hazard estimate graph should be done later in the chapter. However, it is still essential to keep in mind that the age of firms is a significant variable. Having joint-stock companies as a reference group, I can now analyze the likelihood of firm-exit depending on the legal structure of the firm. Firms with the legal structure of sole proprietorship are 66.4 times more likely to fail than joint-stock companies if all other variables are the same. For trading and limited partnerships, the likelihood of failure is 111.4 more times than joint-stock companies. Economic associations show the most resilience compared to other forms by being 9.2 times more likely to fail than joint-stock companies.

Similarly, to analyze the likelihood of failure depending on the industry, I set the manufacturing industry as a reference group. Firms in the trade industry are 16 times more likely to exit compared to manufacturing firms. While service firms are 8.1 times more likely to exit compared to manufacturing. All firm-level predictors are significant on less than 0.01 level, except for the economic association dummy variable, which is significant on the 0.05 level.

Model II includes all the previous variables on the firm-level, as well as variables on the island-level: population density, distance to the urban center, island land area, access to fiber broadband, primary school, kindergarten, grocery store, the ratio of summer passengers and the proxy for institutional thickness (number of associations). Starting with the analysis of the model, the number of observations is 9687, which is consistent with those of model I. The whole model is significant on the 0.001 level. Furthermore, the random variance intercepts are also significant at the 0.001 level, both the intercept-level as well as the modeling level. The log-likelihood has increased further, meaning that the model is a better fit. Consequently, model II explains more of the variation between firms that survive than in model I.

Firms with the legal structure of sole proprietorship are now 41 times more likely to fail than joint-stock companies if all other variables are equal. This change is a significant lowering of

the coefficient, indicating that island-level predictors better explain some of the effects that previously was caught by the legal structure variables. For trading and limited partnerships, the likelihood of failure is 61 times more than for joint-stock companies. Economic associations still show the most resilience compared to other forms by being 7.5 times more likely to fail than joint-stock companies. Moreover, firms in the trade and service industries are now 11.9, respectively, 7.7 times more likely to exit compared to manufacturing firms. Most firm-level predictors are still significant on the 0.01 level, with the dummy variable for the economic associations being significant on the 0.05 level.

Moving on to the island variables, distance to the urban center and access to primary school becomes the only insignificant variables of the study. However, access to primary school comes close to reaching a significance level of $p<0.05$. Further, firms in islands with a denser population, specifically more than 100 permanent inhabitants, are 5.6 times more likely to exit than firms in islands with lower population density and significant at the 0.01 level. The land area of the island is significant at the 0.001 level; however, increasing land area by 1 ha leads to only a slightly higher likelihood of exit as the odds ratio is 1.001. Access to fiber broadband is the model's first predictor that *decreases* the likelihood of exit by 2.3 times and significant at the 0.01 level. Further, firms situated in islands with access to kindergartens are three times less likely to exit than islands without a kindergarten (0.05 level significance).

Similarly, firms located on islands with access to a year-round grocery store are 6.4 times less likely to exit all other things being equal (0.05 level significance). Finally, if the proportion of summer visitors increases by 1, firms become 1.02 times more likely to exit, meaning slightly more likely to exit than before increase. The proxy for institutional thickness, number of active associations on the island, decreases the likelihood of exit by 1.1 times for *each* new association created.

Finally, Model III merges the macroeconomic variables with those of the firm-level and island contextual variables. Specifically, the variable of regional GDP (in 100 billion) and dummy variables for the financial crisis of 2008-2009. Starting with the analysis of the model, the number of observations has decreased to 8831. The decrease is due to missing values on the regional GDP variable, as previously explained. The whole model is significant on the 0.001 level, as previous models. Similarly, the log-likelihood has continued to increase, indicating a better fit of model III compared to previous models.

All in all, it is a significant and better fit model. Interestingly, when adding macroeconomic variables, sole proprietorships have now a higher likelihood of exit than trading and limited

partnerships. Moving on to the macroeconomic variables, an increase of 100 billion in the regional GDP of the previous year leads to 2.5 times decreased likelihood of firm-exit. Further, all other variables held constant, during the financial crisis years of 2008-2009 firms became 2.9 times more likely to exit than during other years.

4.6. What about age?

In the methodology chapter, I established that repeated measurements need to be mean-centered in multilevel models, which makes them particularly tricky to interpret. Thus, I argued that a better way of analyzing the effects of age is by plotting the hazard function[24], a standard procedure in survival analysis (e.g., Geroski, Mata & Portugal 2010, pp. 523–524; Cefis & Marsili 2019, p. 574).

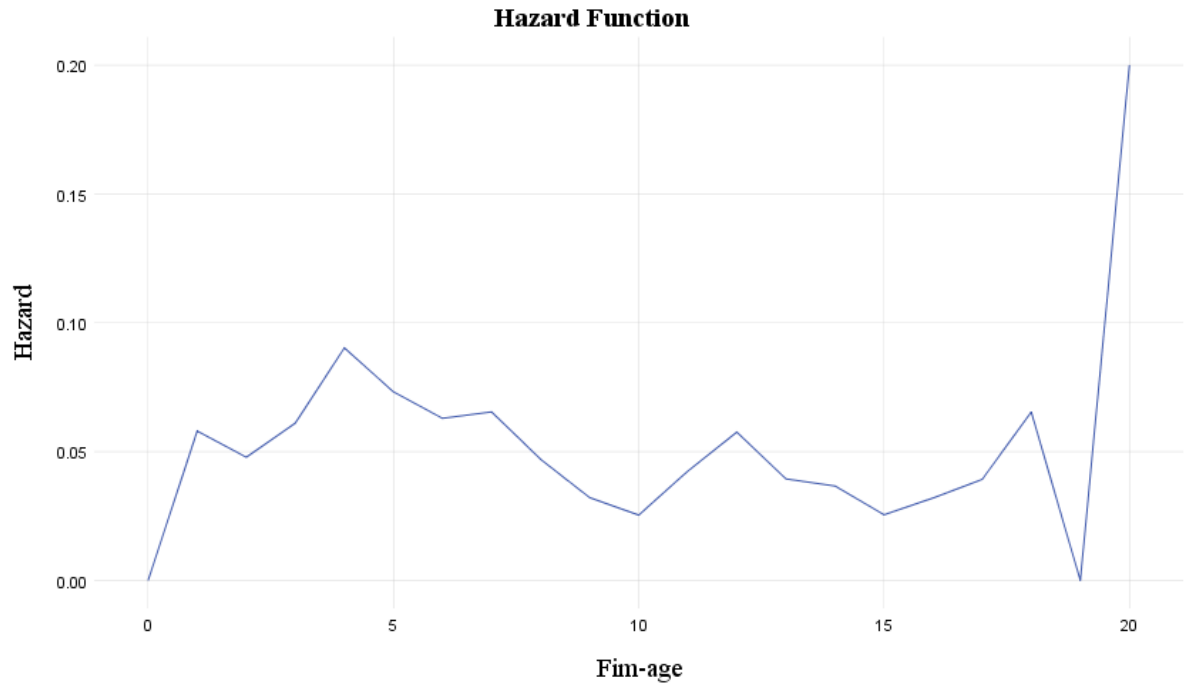

Figure 5: Hazard function graph.

It is important to recall that the significance of the age variable has been confirmed by multilevel logistic regression. The failure estimates on the graph above, figure 5, shows that between age 0 and 5, there is an increase in hazard (risk of exit) reaching the highest point shortly before age 5 (ignoring the age 19-20 jump). A plateau is reached between the fifth and seventh years. After year 7, the hazard function continuously declines until age 10. Between ages 10 and 13, there is a continuous increase. After that, there is yet again a decrease in hazard to year 15 and then

[24] See 3.2.1. Survival analysis

an increase again. The high fluctuation from year 19 to 20 is explained by the statistical power of the few remaining firms in the sample. In short, one exit by the end of the analysis time has much stronger statistical power than early on when the sample is more extensive. Thus, this last fluctuation is ignored.

4.7. Hypothesis testing and summary of results

4.7.1. Hypothesis testing

Island-level (contextual) hypotheses

The first set of hypotheses generated in the research framework regards the contextual variables of the islands and their potential effect on firm-exit. The multilevel logistic model II provides the odds ratio values and significance levels necessary to falsify these hypotheses. Thus, with the results in mind, *H1b is rejected* (null hypothesis accepted) since the distance to the nearest market is insignificant. Further, while the land area of islands is significant, the variable very slightly increases the likelihood of firm-exit. Thus, *H1e is rejected*.

Moreover, *H1a and H1d are accepted* as firms in densely populated islands, and firms in islands with higher proportions of visitors have a decreased likelihood of survival. Institutional thickness is found to increase the likelihood of survival. Therefore, *H1f is accepted*. I did not assume a direction for H1c due to a lack of previous studies. To answer this hypothesis, I constructed several variables that measure different services and infrastructures: access to fiber broadband (increase survival), primary schools (insignificant), kindergartens (increase survival), and year-round grocery shops (increase survival). All variables were significant, with the sole exception of primary school access, and thus *H1c is accepted*.

Macroeconomic hypothesis

The second hypothesis makes assumptions on the effects of macroeconomic conditions on the survivability of firms. The multilevel logistic model III provides the odds ratio values and significance levels necessary to falsify this hypothesis. The first variable, regional GDP, is consistent with the decreased likelihood of exit as the GDP increases. The second variable, the financial crisis, shows an increased likelihood of exit during 2008-2009. Furthermore, both variables are significant. *Thus, H2 is accepted*.

Structural factors hypotheses

The last set of hypotheses regards the effects of firms' structural factors on the likelihood of survival. Hypotheses 3a-d make assumptions on the relation between firm-age and probability of exit. These hypotheses were analyzed using a Kaplan-Meier hazard graph. Age was shown to have a non-linear relation to the likelihood of survival. New firms have a higher likelihood of exit than older firms. Thus, *H3a is accepted*. The exit likelihood increases to reach an all-time high shortly before the fifth year and then decreases; therefore, *H3b-c are also accepted*. Moreover, exit likelihood increases for older firms after the first decade, and thus *H3d* is accepted.

Hypotheses H4-H5, correspond to the legal structure of firms and the industry affiliation, respectively. Joint-stock companies and economic associations show consistently and significantly lower probabilities of exit than sole proprietors, and trading and limited partnerships (model I, II). Thus, *H4 is accepted*. Finally, firms with manufacturing activities have a significantly lower probability of exit compared to trade and service firms. Therefore, *H5 is supported*.

4.7.2. Summary of results

Accepted hypotheses:

Hypothesis 1a: Lower island population density decreases the likelihood of firm-exit.

Hypothesis 1c: The island level of service significantly affects the likelihood of firm-exit.

Hypothesis 1d: Increasing activity, in the form of visitors, increases the likelihood of firm-exit.

Hypothesis 1f: Institutional thickness decreases the likelihood of firm-exit.

Hypothesis 2: The likelihood of firm-exit increases during worsened economic conditions

Hypothesis 3a: Younger firms have a higher likelihood of exit than older firms due to the liability of newness.

Hypothesis 3b: The likelihood of new-firm exit increases until approximately year 5.

Hypothesis 3c: The likelihood of firm-exit decreases after reaching a peak after approximately year 5.

Hypothesis 3d: In the long term, the likelihood of firm-exit increases due to the liabilities of obsolescence and senescence.

Hypothesis 4: Firms with limited liability have a lower likelihood of exit compared to firms with other legal forms.

Hypothesis 5: Firms in manufacturing have a lower likelihood of exit than firms in service and trade.

Rejected hypotheses:

Hypothesis 1b: Longer distance to the nearest urban center (market) increases the likelihood of firm-exit.

Hypothesis 1e: Increasing island land area decreases the likelihood of firm-exit.

V. Discussion: who will survive?

The purpose of this chapter is to discuss the findings of this study in relation to previous literature. Initially, I will discuss the structural and macroeconomic hypotheses of the study, together with prior empirical findings. After that, a discussion will be held on the contextual hypotheses with a focus on the organizational ecology perspective.

In the introductory chapter of this thesis, I formulated the following question:

- What are the challenges and conditions for firm-survival in a peripheral region?

To answer this research question, I set up several hypotheses to test propositions in previous empirical studies concerning structural factors of the firm as well as macroeconomic variations and firm-survival. Further, while there is a lack of research regarding entrepreneurship in peripheral regions, prior studies have discussed some of the challenges prevalent in peripheral and island entrepreneurship. Thus, additional hypotheses were generated concerning contextual factors that may affect survival. The research question and the hypotheses are in line with the aim of the study: to analyze firm-survival across *time*, on *multiple levels* of analysis, within the *micro-spatial* environment in a peripheral region. Furthermore, this analysis is argued to shed light on the conditions and challenges for firm-survival *within* the regional context of Stockholm Archipelago, specifically the 17 core-islands. Finally, it is essential to note that all results can only be said to apply to firms that have survived the initial *nascent*[25] phase.

5.1. Structural and macroeconomic factors

This thesis confirms previous results linking structural factors of firms and their probabilities of survival. Age has a non-linear relationship with survival. The hazard of exit is generally lower for older firms. More specifically, the likelihood of exit increase during the preceding years to an all-time high and then decrease. However, there is a slight increase again shortly after reaching a bit more than a decade of maturity. Thus, founders face increasing challenges during the initial years of the venture, consistent with the stylized fact of liability of newness (e.g., Holmes, Hunt & Stone 2010, p. 186; Axtell 2016, p. 810; Box, Gratzer & Lin 2017, pp. 450–452). However, older firms face new challenges after more than a decade in the market,

[25] For more detail, see discussion on left-truncation in section 3.2.2.

characterized by the increased hazard of exit. The findings are consistent with previous empirical findings stylized as the liability of obsolescence and senescence (Coad 2018, pp. 17–25). These findings could be due to increased competition from new entrants, changes in technology, and external markets.

In previous studies, the legal structure of firms was found to be linked to firm-survival. Specifically, firms with unlimited liability were found to have increased hazard rates compared to firms with limited liability (e.g., Mata & Portugal 1994, p. 242; Esteve-Pérez & Mañez-Castillejo 2008, p. 244; Türkcan & Erkuş-Öztürk 2019, p. 16; Baumöhl, Iwasaki & Kočenda 2020, p. 14; Iwasaki & Kim 2020, p. 23). In the present study, joint-stock companies, and economic associations (limited liability) were found to have a decreased likelihood of exit compared to sole proprietorships and limited and trading partnerships (unlimited liability). Thus, sole proprietorships, as well as limited and trading partnerships, face increasing challenges. However, it is essential to recall that these legal forms have different requirements and purposes. Particularly, joint-stock companies have relatively high initial capital requirements compared to all other forms (Hemström & Giertz 2014, p. 58; Swedish Companies Registration Office 2020b). Further, economic associations have more members and are found to play a crucial role in terms of social causes (Lithander 2005, p. 36).

Moreover, in the theoretical framework, I mentioned several studies that linked larger firm size with a lower likelihood of exit (e.g., Wagner 1999, p. 77; Box 2008, pp. 388–389; Geroski, Mata & Portugal 2010, p. 523). Thus, there is a possibility that the results of the legal structure may be due to the effects of size, whereby legal structures requiring more considerable capital or plural membership are larger and therefore survive longer. Further, economic associations could have better survivability due to being crucial for community needs. Nonetheless, the results of the study are consistent with previous findings concerning legal forms and survival probabilities.

The firm's activities, specifically their industry affiliation, is also linked to different probabilities of exit. Mainly, firms with manufacturing activities have been found to have lower exit rates compared to service and trading firms (e.g., Box, Gratzer & Lin 2017, p. 450; Iwasaki & Kim 2020, p. 31). However, within the manufacturing industry, not all activities have necessarily lower hazard risk (Ferragina, Pittiglio & Reganati 2012, pp. 369–371). In the present study, the sample of firms is too small to warrant further division than the broader categories of manufacturing, trade, and service. Further, the results of the study confirm that manufacturing firms have a lower likelihood of exit. Thus, firms with trading and service activities face

different challenges than those in the manufacturing sector. However, manufacturing activities often have higher capital requirements due to expensive machinery and specialized know-how. Thus, similarly to the legal form, the industry variables may be, to an extent, indicators of size.

Finally, the results of the study show that during the years of 2008-2009, global financial crisis years, probabilities of exit were multiple times higher compared to other years. Further, when the regional GDP increased, the likelihood of exit substantially decreased the following year. Thus, these findings are consistent with previous studies linking favorable economic conditions to better firm-survival (e.g., Geroski, Mata & Portugal 2010, p. 526; Box, Gratzer & Lin 2017, p. 434; Cefis & Marsili 2019, p. 573; Türkcan & Erkuş-Öztürk 2019, p. 15). Therefore, a conclusion is that, despite being on the periphery, Stockholm Archipelago firms still face challenges of regional and global macroeconomic downturns.

5.2. Context of the island and organizational ecology

As mentioned in the theoretical framework, many variables may influence survival rates beyond those covered in the present study. For instance, the results of the intercept-only model show that within-firm differences explain 82.7% of firm-exit variability. In other words, a large portion of why some of the firms survive and others perish, may be found inside the firms. There are two, not necessarily exclusive, explanations for these results.

First, the effect could simply be due to the study's *sample* of islands and firms. The islands of the study have all been designated "core-islands" by the authorities, as they are argued to be suitable for long-term investment (County Board of Stockholm 2019, p. 1). These islands can then be assumed to have more things in common than not, which would explain the results. However, higher regional variability would likely have been found in the study if core regions were included. The difference between regions would be what Clausen (2020, p. 115) labeled "liability of rurality." The particular challenges faced by entrepreneurs in peripheral regions when all other factors (e.g., legal form, size, and age) are equal.

Second, the results are due to the internal factors of the firms. For instance, firms with better competence, more capital, and a more significant share of the market have higher probabilities of survival. Thereby, the type of entrepreneurship may be a significant factor for survival. For instance, necessity driven entrepreneurs are usually not innovative and are less likely to perceive opportunities in the market (Stephens, Partridge & Faggian 2013, p. 780). Thus, it would stand to reason that this type of firm has difficulties exploiting enough resources and attracting

the right competence for growth. I argue that it is highly probable that together, these two explanations (sample and type of entrepreneurship) explain the 82.7% of firm-exit variability.

Beyond the above results, the islands of the study have different contexts, which in the previous chapter were shown to have substantial effects on firm-survival. Applying the perspective of organizational ecology could explain some of the findings on the island-level. The organizational ecology perspective proposes that organizations within the same population are subject to environmental forces and dependent on specific resources niches, leading to competition and cooperation both within the same population and other populations. Further, the competition and cooperation result in particular organizational populations that are adapted to their specific environments. Notably, the evolutionary process of variation, selection, and retention is proposed to explain the results (Aldrich & Ruef 2006, p. 35). Aldrich and Wiedenmayer (2019, p. 64), further emphasize the *diffusion* and *competitive struggle* of organizational forms. The competence held by the individuals inside organizations is diffused through networks. Finally, a competitive struggle over resources occurs once an organizational form thrives, again leading to variations and the selection of those forms that best fit the environment.

In core regions, increasing population density is linked to agglomeration economies, whereby increased competition leads to better outcomes for the region (Audretsch, Hülsbeck & Lehmann 2012, pp. 498–499; Koster & van Stel 2014, p. 215). On the other hand, firms situated in larger islands (surface area) with higher population density, face increased probabilities of exit. I can assume that higher population density, above a certain threshold, harms firm-survival. Thus, using the organizational ecology perspective, I argue that in more densely populated islands, there may be a harsher selection process. Specifically, densely populated islands have higher numbers of firms and thus possibly more considerable variations and higher competition for resources. Conversely, firms located on islands with lower population density thus have lower variations, with lower competition and sufficient resources. This argument is supported by prior empirical findings that excessive market crowding and competition reduce the survival prospects of firms (Geroski, Mata & Portugal 2010, p. 526).

Therefore, contrary to the argument of Delfmann et al. (2014, p. 1048), that at first, population decline captures the attention, and as the decline continues, the regions adjust to the shock by increasing entrepreneurship. I propose an alternative explanation. Regions do not necessarily adjust to the shock by increasing entrepreneurship; instead, competition is decreased, and perhaps more resources become available. Moreover, it is possible that the remaining entrepre-

neurs, due to *necessity*, seek to take on more activities by providing lacking services. Furthermore, contrary to the core-regions, due to the limitations of necessity entrepreneurship, islands are less likely to reap the benefits of "simple" competition (Low, Henderson & Weiler 2005, pp. 68–69; Acs 2006, pp. 98–101). Archipelago firms, then, indeed face particular challenges related to *islandness*, whereby the very context work to their disadvantage (Baldacchino & Fairbairn 2006, p. 331).

Prior studies have found that the Stockholm Archipelago islands have substantial differences regarding institutions (Rytkönen et al. 2019, pp. 71–82). In the present study, institutional thickness, operationalized as the number of associations on the island, is linked with an increased likelihood of survival. Therefore, on islands with higher numbers of associations, the community can be assumed to be closer. Further, the diffusion of competence is dependent on the external network of individuals inside organizations. I assume that islands with increasing institutional thickness have increased community dynamics with more knowledge spillovers, community-level learning, and entrepreneurial ecosystems and clusters (Clausen 2020, p. 116). These ecosystems and clusters may then be the very networks used for the diffusion of competence.

The present study found that firms located on islands with access to either year-round grocery stores, kindergartens, or fiber broadband have decreased probabilities of exit. Local access to primary schools, although coming close to being significant, did not reach the specified significance level. As only islands above a certain population threshold have local school access, I assume that the variable of local school access is also an indicator of population density. Thus, it would explain why the variable became insignificant in the analysis, as it measures the same effect as that of the population density variable. This argument is further motivated by the increased, but insignificant, likelihood of firm-exit; which is consistent with that of higher population density.

Time distance to urban center was found to be insignificant, while islands with higher ratios of summer visitors were found to have a significantly higher likelihood of exit. These findings have important policy implications, especially for the RUFS 2050 agenda. According to the development plan, the highest priority is to invest in transport-related infrastructure to reach "better accessibility" (Tillväxt- och regionplaneförvaltningen 2018, pp. 10–11). First, many islands on the archipelago still lack essential services and communications technology that are not only important for the inhabitants but also the local firms. Second, firms on the islands have probably adapted to the particular issue of distance to the mainland. Thus, investments should

prioritize reaching an acceptable level of infrastructure on the islands. Investment in transport, although undoubtedly beneficial, may not have the same level of impact, as other problems need to be prioritized.

Further, most islands have substantial differences between the number of visitors during the off-season and high-season. More summer visitors do not seem to translate to a higher likelihood of firm-survival. Thus, one strategy could be for the archipelago islands to market themselves year-round and for firm owners to diversify away from season-dependent activities. However, the former is particularly tricky due to the persistent fiction of the archipelago as a summer destination (Widholm 2019, p. 59).

VI. Conclusion

The concluding chapter of the thesis provides a summary of the findings, followed by a short discussion on the study's contributions and implications for policy. Lastly, I discuss the reach and limitations of the study, as well as suggest avenues for future research.

6.1. Summary of the findings

This study aimed to longitudinally analyze firm-survival in peripherality on multiple levels of analysis, by answering a research question: What are the challenges and conditions for firm-survival in a peripheral region? Earlier studies have not taken into consideration the micro-spatial and multilevel context of organizations. For this reason, the study was designed to include aspects of change in study subjects (firms) and context (islands). It is essential to note that all results can only be said to apply to firms that have survived the initial *nascent*[26] phase. Thus, using triangulation methods, I constructed a longitudinal dataset, encompassing changes, and characteristics of the islands and firms from the year 2000 through 2019. One of the most striking results of the present study is the substantial effect of the within-firm level on firm-survival, which amounts to 82.7% of the firm-exit variability. Whereas the sample of the study could partially explain some of this variability, I propose that the type of entrepreneurship – necessity or opportunity-driven – may also explain these differences and thus play a significant role in improving firm-outcomes

Further, on the firm-level, younger firms in peripherality face challenges, or liability, of newness as the hazard of exit increase from entry and reach a maximum roughly around the fifth year. Whereas older peripheral firms, in general, have a lower risk of exit, they face increasing hazard rates a decade after founding (liabilities of senescence and obsolescence). Firms with manufacturing activities have the lowest probability of exit. Thus, firms with trade and service activities face significant challenges compared to manufacturing firms. Furthermore, differences in the legal structures may bring significant challenges. Expressly, trading and limited partnerships, and sole proprietorships have higher probabilities of exit than joint-stock companies and economic associations. Considering prior findings concerning the size of firms, it may be possible that legal form and industry could be indicators of size. Thus, firms with more considerable initial capital or plural membership are larger and therefore survive longer.

[26] For more detail, see discussion on left-truncation in section 3.2.2.

On the micro-spatial level, the results suggest that differences in local contexts lead to different survival outcomes. Firms situated in islands with higher population density face increased probabilities of exit. Thus, analyzing the results through an organizational ecology perspective, it could be that firms on islands above a certain population threshold have increased organizational variations and a harsher selection process. Further, this assumption falls in line with previous findings that overcrowding of markets leads to decreased survival rates. Time distance to markets (urban centers) has no significant effect on the survival of firms. However, access to fiber broadband, kindergartens, and year-round grocery stores are significant conditions for firm-survival. Similarly, a thicker institutional context, in the form of local associations, is a condition for firm-survival.

Firms situated in islands with an increasing proportion of summer visitors, independent of island size, have an increased likelihood of exit. Finally, favorable regional macroeconomic conditions, in the form of increasing regional GDP, are a condition for increased probabilities of survival. Conversely, worsened economic conditions, such as during the global financial crisis of 2008-2009, substantially decrease the likelihood of survival despite their peripheral location. Thus, in general, peripheral firms can be said to face the same challenges as core-region firms, as well as additional challenges unique to their local context.

6.2. Contributions

This study thus validates previous knowledge about the relation between the structural factors of firms, macroeconomic conditions and firm-survival (e.g., Geroski, Mata & Portugal 2010; Box, Gratzer & Lin 2017; Türkcan & Erkuş-Öztürk 2019; Baumöhl, Iwasaki & Kočenda 2020). The study also took into consideration the context of entrepreneurship, and thus contributed with new data and knowledge regarding factors on a micro-spatial and peripheral context (e.g., Korsgaard, Ferguson & Gaddefors 2015; Huggins, Prokop & Thompson 2017; Clausen 2020).

More specifically, this study adds to recent discussion in entrepreneurship and economic geography regarding *islandness* and peripherality (e.g., Baldacchino & Fairbairn 2006; Stratford et al. 2011; Burnett & Danson 2017; Rytkönen et al. 2019). Furthermore, the study did not neglect to consider possible effects of the institutional thickness on firm-survival (e.g., Rodríguez-Pose 2013; Zukauskaite, Trippl & Plechero 2017). The study also incorporated time and space perspective (e.g., Coviello & Jones 2004; Aldrich 2009) as well as contributed to the understanding of the multilevel nature of organizational events (Morgeson, Mitchell & Liu 2015).

6.3. Policy implications

Beyond the effects of firms' structural factors, the study has sought to investigate the effects of contextual factors through time. The results show significant differences in terms of organizational outcomes. Thus, policymakers should not only take into consideration the uniqueness of Stockholm Archipelago as a whole but also the mixture of contextual variables present on each island is vital. Strategies that may show results in some of the islands may, therefore, not have the same effects on other islands with a different context. However, in general, supporting local associations as well as improving local access to necessary infrastructure and services should be prioritized.

6.4. Reach, limitations, and avenues for future research

The discussion on necessity and opportunity entrepreneurship provided useful guidance in the construction of the study's variables relevant to firm-survival in peripherality. Whereas the purpose of the study was not to explore these variations of entrepreneurship and their effect, I proposed that the type of entrepreneurship could be a major factor of survival. Thus, future research should take into consideration these crucial differences in entrepreneurship when investigating firm outcomes in peripherality.

Through an extensive collection of data and statistical analysis, the present study has managed to bridge an apparent knowledge gap in regards to firm-survival and peripherality. Furthermore, several new determinants of firm-survival on a micro-spatial level were found. However, a limitation of the study has been the dependency on registration data, making it challenging to reach firms that did not survive the nascent phase. For future research concerning the survival of firms in Sweden, I would recommend research designs that seek to reach these nascent firms.

The present study has managed to make use of imprinting theory and organizational ecology perspective to discuss these findings. Future studies may thus make use of different perspectives and theories to analyze firm-survival. Furthermore, while the present study did not problematize the particular closeness of the Stockholm Archipelago to the core region, such a problematization would create opportunities for comparative research. Moreover, taking into consideration competing risks of exit may lead to further nuanced results and analysis. However, researchers need to collect sufficient data covering different types of exits. The empirical data of the present study, for instance, showed very few bankruptcies relative to liquidation.

The empirical results of this study have shown that internal factors substantially predict firm-survival and firm-exit probabilities. Thus, an avenue of future research is to analyze these through a mixed-method approach using both statistics and phenomenography through participant and group interviews. Finally, due to their particular importance for regional development, I suggest the use of qualitative methods to study innovative firms in peripheral regions.

References

Acs, Z. (2006). How Is Entrepreneurship Good for Economic Growth? *Innovations: Technology, Governance, Globalization*, 1(1), pp. 97–107, doi:10.1162/itgg.2006.1.1.97.

Agarwal, R. & Audretsch, D. B. (2001). Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival. *Journal of Industrial Economics*, 49(1), pp. 21–43, doi:10.1111/1467-6451.00136.

Al-Turk, A. & Aldrich, H. E. (2019). Revisiting "Traits to Rates" After 25 Years: Organizational Ecology's Limited Impact on Entrepreneurship Research. In: Katz, J. A. & Corbet, A. C. (eds.) *Seminal Ideas for the Next Twenty-Five Years of Advances*. Emerald Publishing Limited, pp. 99–114.

Aldrich, H. E. (2009). Lost in space, out of time: Why and how we should study organizations comparatively. In: King, B. G., Felin, T., & Whetten, D. A. (eds.) *Studying Differences between Organizations: Comparative Approaches to Organizational Research*. Emerald Group Publishing Limited, pp. 21–44.

Aldrich, H. E. & Ruef, M. (2006). *Organizations Evolving*. 2nd ed. London: SAGE Publications Ltd.

Aldrich, H. E. & Wiedenmayer, G. (2019). From Traits to Rates: An Ecological Perspective on Organizational Foundings. In: Katz, J. A. & Corbet, A. C. (eds.) *Seminal Ideas for the Next Twenty-Five Years of Advances*. Emerald Publishing Limited, pp. 61–97.

Aldrich, H. E. & Yang, T. (2012). What Did Stinchcombe Really Mean? Designing Research to Test the Liability of Newness among New Ventures. *Entrepreneurship Research Journal*, 2(3), doi:10.1515/2157-5665.1077.

Amin, A. & Thrift, N. (1994). Living in the global. In: Amin, A. & Thrift, N. (eds.) *Globalization, institutions, and regional development in Europe*. Oxford: Oxford University Press.

Anyadike-Danes, M. & Hart, M. (2018). All grown up? The fate after 15 years of a quarter of a million UK firms born in 1998. *Journal of Evolutionary Economics*, 28(1), pp. 45–76, doi:10.1007/s00191-017-0549-x.

Audretsch, D. B., Hülsbeck, M. & Lehmann, E. E. (2012). Regional competitiveness, university spillovers, and entrepreneurial activity. *Small Business Economics*, 39(3), pp. 587–601, doi:10.1007/s11187-011-9332-9.

Axtell, R. L. (2016). 120 Million Agents Self-Organize into 6 Million Firms: A Model of the U.S. Private Sector. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, pp. 806–816.

Baldacchino, G. & Fairbairn, T. I. J. (2006). Editorial: Entrepreneurship and Small Business Development in Small Islands. *Journal of Small Business & Entrepreneurship*, 19(4), pp. 331–340, doi:10.1080/08276331.2006.10593374.

Baumöhl, E., Iwasaki, I. & Kočenda, E. (2020). Firm survival in new EU member states. *Economic Systems*, 44(1), p. 100743, doi:10.1016/j.ecosys.2020.100743.

Baumol, W. J. (1990). Entrepreneurship: Productive, Unproductive, and Destructive. *Journal of Political Economy*, 98(5), pp. 893–921.

Bisnode Infororg (2016). *Informationsmaterial: InfoTorg Person & Företag*. Stockholm.

Box-Steffensmeier, J. M. & Jones, B. S. (2004). *Event History Modeling*. Analytical Methods for Social Research. Cambridge: Cambridge University Press.

Box, M. (2008). The death of firms: exploring the effects of environment and birth cohort on firm survival in Sweden. *Small Business Economics*, 31(4), pp. 379–393, doi:10.1007/s11187-007-9061-2.

Box, M., Gratzer, K. & Lin, X. (2017). New-Firm Survival in Sweden: New Methods and Results. *International Review of Entrepreneurship*, (4), pp. 431–464.

Burnett, K. A. & Danson, M. (2017). Enterprise and Entrepreneurship on islands and remote rural environments. *The International Journal of Entrepreneurship and Innovation*, 18(1), pp. 25–35, doi:10.1177/1465750316686237.

Carroll, G. R. & Hannan, M. T. (2000). *The Demography of Corporations and Industries*. Princeton, N.J.: Princeton University Press.

Cefis, E. & Marsili, O. (2019). Good times, bad times: innovation and survival over the business cycle. *Industrial and Corporate Change*, 28(3), pp. 565–587, doi:10.1093/icc/dty072.

Chakraborty, S. (2018). A step-wise guide to performing survival analysis. *Cancer Res Stat Treat*, 1(1), doi:10.4103/CRST.CRST_5_18.

Clausen, T. H. (2020). The liability of rurality and new venture viability. *Journal of Rural Studies*, 73, pp. 114–121, doi:10.1016/j.jrurstud.2019.12.005.

Coad, A. (2018). Firm age: a survey. *Journal of Evolutionary Economics*, 28(1), pp. 13–43, doi:10.1007/s00191-016-0486-0.

Congregado, E., Golpe, A. A. & Parker, S. C. (2012). The dynamics of entrepreneurship: hysteresis, business cycles, and government policy. *Empirical Economics*, 43(3), pp. 1239–1261, doi:10.1007/s00181-011-0516-6.

Copus, A. K. (2001). From Core-periphery to polycentric development: Concepts of spatial and aspatial peripherality. *European Planning Studies*, 9(4), pp. 539–552, doi:10.1080/713666491.

Coulson, A. & Ferrario, C. (2007). 'Institutional Thickness': Local Governance and Economic Development in Birmingham, England. *International Journal of Urban and Regional Research*, 31(3), pp. 591–615, doi:10.1111/j.1468-2427.2007.00739.x.

County Board of Stockholm (2012). *Skärgårdsfakta – Grafiska kartor 2012*. Stockholm.

County Board of Stockholm (2014). *Skärgårdsfakta – Grafiska kartor 2014*. Stockholm.

County Board of Stockholm (2015). *Skärgårdsfakta – Grafiska kartor 2015*. Stockholm.

County Board of Stockholm (2016). *Skärgårdens utveckling i siffror: Rapport 2016:01*. Stockholm.

County Board of Stockholm (2017). *Skärgårdsfakta – Grafiska kartor 2017*. Stockholm.

County Board of Stockholm (2019). *Skärgårdsfakta – Grafiska kartor 2019*. Stockholm.

Coviello, N. E. & Jones, M. V. (2004). Methodological issues in international entrepreneurship research. *Journal of Business Venturing*, 19(4), pp. 485–508, doi:10.1016/j.jbusvent.2003.06.001.

Danson, M. & De Souza, P. (2011). Periphery and marginality Definitions, theories, methods and practice. In: Danson, M. & De Souza, P. (eds.) *Regional development in Northern Europe: peripherality, marginality and border issues*. New York: Routledge.

Delfmann, H., Koster, S., McCann, P. & Van Dijk, J. (2014). Population Change and New Firm Formation in Urban and Rural Regions. *Regional Studies*, 48(6), pp. 1034–1050, doi:10.1080/00343404.2013.867430.

Delmar, F. & Shane, S. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing*, 19(3), pp. 385–410, doi:10.1016/S0883-9026(03)00037-5.

Edmondson, A. C. & Mcmanus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review*, 32(4), pp. 1155–1179, doi:10.5465/AMR.2007.26586086.

Eriksson, M. (2008). (re)producing a "peripheral" region – northern sweden in the news. *Geografiska Annaler: Series B, Human Geography*, 90(4), pp. 369–388, doi:10.1111/j.1468-0467.2008.00299.x.

Ersoy, A. (2018). *Turkey - an economic geography*. I.b.tauris & Co. Ltd.

Esteve-Pérez, S. & Mañez-Castillejo, J. A. (2008). The Resource-Based Theory of the Firm and Firm Survival. *Small Business Economics*, 30(3), pp. 231–249, doi:10.1007/s11187-006-9011-4.

Etikan, I. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), p. 1, doi:10.11648/j.ajtas.20160501.11.

European Commission (2020). *European Regional Development Fund - Regional Policy*. https://ec.europa.eu/regional_policy/en/funding/erdf/ [2020-02-14].

Eurostat (2008). *NACE Rev. 2 - Statistical classification of economic activities*. Luxembourg:

Ferragina, A., Pittiglio, R. & Reganati, F. (2012). Multinational status and firm exit in the Italian manufacturing and service sectors. *Structural Change and Economic Dynamics*,

23(4), pp. 363–372, doi:10.1016/j.strueco.2011.10.002.

Field, A. (2017). *Discovering statistics using IBM SPSS statistics*. 5th ed. London: SAGE Publications Ltd.

Fritsch, M. & Storey, D. J. (2014). Entrepreneurship in a Regional Context: Historical Roots, Recent Developments and Future Challenges. *Regional Studies*, 48(6), pp. 939–954, doi:10.1080/00343404.2014.892574.

Geroski, P. A., Mata, J. & Portugal, P. (2010). Founding conditions and the survival of new firms. *Strategic Management Journal*, 31(5), pp. 510–529, doi:10.1002/smj.823.

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, 8, pp. 597–607.

Griffith, D. A. (2008). Ethical Considerations in Geographic Research: What Especially Graduate Students Need to Know. *Ethics, Place & Environment*, 11(3), pp. 237–252, doi:10.1080/13668790802559650.

Gyllenberg, E.-K. & Maluszynski, C. (2004). Äntligen invigdes nya skolan på Utö. *Dagens Nyheter.*, October 1, 2004.

Harhoff, D., Stahl, K. & Woywode, M. (1998). Legal form, growth and exit of west German firms - Empirical results for manufacturing, construction, trade and service industries. *Journal of Industrial Economics*, 46(4), pp. 453–488, doi:10.1111/1467-6451.00083.

Hay, P. (2006). A Phenomenology of Islands. *Island Studies Journal*, 1(1), pp. 19–42.

Hedström, P. & Ylikoski, P. (2010). Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, 36(1), pp. 49–67, doi:10.1146/annurev.soc.012809.102632.

Hemström, C. & Giertz, M. (2014). *Bolag - föreningar - stiftelser*. 8., [rev.]. Stockholm: Norstedts juridik.

Hill, C. & Wallström, K. (2008). The Stockholm Archipelago. In: Schiewer, U. (ed.) *Ecology of Baltic Coastal Waters*. Berlin: Springer Berlin Heidelberg, pp. 309–334.

Holmes, P., Hunt, A. & Stone, I. (2010). An analysis of new firm survival using a hazard function. *Applied Economics*, 42(2), pp. 185–195, doi:10.1080/00036840701579234.

Hox, J., Moerbeek, M. & Schoot, R. van de (2018). *Multilevel analysis : techniques and applications*. Third edit. Taylor & Francis Ltd.

Huggins, R., Prokop, D. & Thompson, P. (2017). Entrepreneurship and the determinants of firm survival within regions: human capital, growth motivation and locational conditions. *Entrepreneurship & Regional Development*, 29(3–4), pp. 357–389, doi:10.1080/08985626.2016.1271830.

Huynh, K. P., Petrunia, R. J. & Voia, M. (2010). THE IMPACT OF INITIAL FINANCIAL STATE ON FIRM DURATION ACROSS ENTRY COHORTS*. *The Journal of Industrial Economics*, 58(3), pp. 661–689, doi:10.1111/j.1467-6451.2010.00429.x.

Island by Island (2020). *Ö för Ö*. http://www.oforo.se/ [2020-01-13].

Iwasaki, I. & Kim, B.-Y. (2020). Legal forms, organizational architecture, and firm failure: a large survival analysis of Russian corporations. *European Journal of Law and Economics*, doi:10.1007/s10657-020-09644-8.

Kinossian, N. (2019). Agents of change in peripheral regions. *Baltic Worlds*, XII(2), pp. 61–66.

Korsgaard, S., Ferguson, R. & Gaddefors, J. (2015). The best of both worlds: how rural entrepreneurs use placial embeddedness and strategic networks to create opportunities. *Entrepreneurship & Regional Development*, 27(9–10), pp. 574–598, doi:10.1080/08985626.2015.1085100.

Koster, S. & van Stel, A. (2014). The relationship between start-ups, market mobility and employment growth: An empirical analysis for Dutch regions. *Papers in Regional Science*, 93(1), pp. 203–217, doi:10.1111/pirs.12000.

Lidström, S. (2019). 3 islands— 3 entrepreneurs. *Baltic Worlds*, XII(2).

Lithander, J. (2005). *Gränslandets ekonomi och entreprenörskap*. Växjö universitet.

Liu, Z. (2020). Unraveling the complex relationship between environmental and financial performance ——— A multilevel longitudinal analysis. *International Journal of Production Economics*, 219(July 2019), pp. 328–340, doi:10.1016/j.ijpe.2019.07.005.

Low, S., Henderson, J. & Weiler, S. (2005). Gauging a Region's Entrepreneurial Potential. *Economic Review - Federal Reserve Bank of Kansas City*, 90(3), pp. 3,61-89.

Lundström, A., Nordström, C. & von Friedrichs, Y. (2014). *Hur kan mer effektiva stöd utformas för att få fram livskraftiga nya och unga företag?*

Maas, C. J. M. & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1(3), pp. 86–92, doi:10.1027/1614-2241.1.3.86.

Manjón-Antolín, M. C. & Arauzo-Carod, J.-M. (2008). Firm survival: methods and evidence. *Empirica*, 35(1), pp. 1–24, doi:10.1007/s10663-007-9048-x.

Marcus Aurelius & Hays, G. (2003). *Meditations : a new translation*. New York: Modern Library.

Mata, J. & Portugal, P. (1994). Life Duration of New Firms. *The Journal of Industrial Economics*, 42(3), p. 227, doi:10.2307/2950567.

McAdam, R., Reid, R. & Shevlin, M. (2014). Determinants for innovation implementation at SME and inter SME levels within peripheral regions. *International Journal of Entrepreneurial Behavior & Research*, 20(1), pp. 66–90, doi:10.1108/IJEBR-02-2012-0025.

Merlo, J., Wagner, P., Austin, P. C., Subramanian, S. & Leckie, G. (2018). General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial. *SSM - Population Health*, 5, pp. 33–37,

doi:10.1016/j.ssmph.2018.05.006.

Millán, J. M., Román, C. & Congregado, E. Entrepreneurship Analysis from a Human Population Surveys' Perspective. In: *Measuring Entrepreneurship*. Boston, MA: Springer US, pp. 85–112.

Molina-Azorín, J. F., Pereira-Moliner, J., López-Gamero, M. D., Pertusa-Ortega, E. M. & Tarí, J. J. (2019). Multilevel research: Foundations and opportunities in management. *BRQ Business Research Quarterly*, doi:10.1016/j.brq.2019.03.004.

Morgeson, F. P., Mitchell, T. R. & Liu, D. (2015). Event System Theory: An Event-Oriented Approach to the Organizational Sciences. *Academy of Management Review*, 40(4), pp. 515–537, doi:10.5465/amr.2012.0099.

Nämdö Fiber (2020). *Nämdö Fiber*. http://www.namdofiber.se/ [2020-03-28].

Nämdö Library (2020). *Om Nämdö skola*. http://namdo.dinstudio.se/text1_6.html [2020-04-2].

North, D. C. (1990). *Institutions, institutional change, and economic performance (Political Economy of Institutions and Decisions)*. Cambridge University Press.

OECD (2019). *About the OECD*. https://www.oecd.org/about/ [2020-05-6].

Ornö Fiber EF (2020). *Ornö Fiber Ekonomisk förening*. http://ornofiber.se/ [2020-03-28].

Rabe-Hesketh, S. & Skrondal, A. (2012). *Multilevel and longitudinal modeling using stata. Volumes I and II*. 3rd ed. Texas: Stata Press.

Region Stockholm (2020). *Public transport*. https://www.sll.se/om-regionstockholm/Information-in-English1/Public-transport/ [2020-05-7].

Retriever Business *Sökmanual för Retriever Business*. Stockholm.

Rodríguez-Pose, A. (2013). Do Institutions Matter for Regional Development? *Regional Studies*, 47(7), pp. 1034–1047, doi:10.1080/00343404.2012.748978.

Rytkönen, P. & Kinossian, N. (2019). The Baltic Sea Region Archipelagos and Islands: Conditions and challenges. *Baltic Worlds*, XII(2), pp. 37–39.

Rytkönen, P., Larsson-segerlind, T., Onn, G., Degerstedt, L. & Kaipainen, M. (2019). Facing business challenges with the Stockholm Archipelago as a context. *Baltic Worlds*, XII(2), pp. 74–86.

Schumpeter, J. (1983). *The theory of economic development : an inquiry into profits, capital, credit, interest, and the business cycle*. New Brunswick, (U.S.A.): Transaction Publishers.

Schwarz, A. & Jax, K. (2011). Etymology and Original Sources of the Term "Ecology". In: *Ecology Revisited: Reflecting on Concepts, Advancing Science*. Dordrecht: Springer Netherlands, pp. 145–147.

Seltzer, M. The Use of Hierarchical Models in Analyzing Data From Experiments and Quasi-Experiments Conducted in Field Settings. In: Kaplan, D. (ed.) *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks: SAGE Publications, Inc., pp. 260–281.

Simsek, Z., Fox, B. C. & Heavey, C. (2015). "What's Past Is Prologue." *Journal of Management*, 41(1), pp. 288–317, doi:10.1177/0149206314553276.

Sinek, S. (2009). *Start with why : how great leaders inspire everyone to take action*. London: Portfolio Penguin.

Smelser, N. J. & Baltes, P. B. (2004a). Longitudinal Data. *International Encyclopedia of the Social & Behavioral Sciences*.

Smelser, N. J. & Baltes, P. B. (2004b). Panel Surveys: Uses and Applications. *International Encyclopedia of the Social & Behavioral Sciences*.

Sommet, N. & Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology*, 30(1), pp. 203–218, doi:10.5334/irsp.90.

Statistics Sweden (2007a). *SNI 2007 Swedish Standard Industrial Classification 2007*. Stockholm: SCB.

Statistics Sweden (2007b). *Sökning efter SNI-kod*. http://www.sni2007.scb.se/default.asp [2020-03-22].

Statistics Sweden (2014). *Öar i Sverige 2013*. Stockholm: Statistiska centralbyrån.

Statistics Sweden (2019a). *Strandnära markanvändning*. Statistics Sweden. https://www.scb.se/hitta-statistik/statistik-efter-amne/miljo/markanvandning/strandnara-markanvandning/ [2020-02-15].

Statistics Sweden (2019b). *Regionalräkenskaper*. https://www.scb.se/hitta-statistik/statistik-efter-amne/nationalrakenskaper/nationalrakenskaper/regionalrakenskaper/ [2020-05-19].

Statistics Sweden (2020). *Indelningar Regina SCB*. http://regina.scb.se/indelningar [2020-04-1].

Stephens, H. M., Partridge, M. D. & Faggian, A. (2013). INNOVATION, ENTREPRENEURSHIP AND ECONOMIC GROWTH IN LAGGING REGIONS. *Journal of Regional Science*, 53(5), pp. 778–812, doi:10.1111/jors.12019.

Stinchcombe, A. (1965). Social structure and organizations. In: March, J. P. (ed.) *Handbook of organizations*. Chicago, pp. 142–193.

Stratford, E., Baldacchino, G., McMahon, E., Farbotko, C. & Harwood, A. (2011). Envisioning the Archipelago. *Island Studies Journal*, 6(2), pp. 113–130.

Swedish Companies Registration Office (2015). *Driver företaget fortfarande verksamhet?* https://bolagsverket.se/ff/foretagsformer/driver-verksamhet-1.2089 [2020-03-1].

Swedish Companies Registration Office (2018a). *Vad är ett handelsbolag?* https://bolagsverket.se/ff/foretagsformer/handelsbolag/vad-1.3453 [2020-05-8].

Swedish Companies Registration Office (2018b). *Vad är en enskild näringsidkare?* https://bolagsverket.se/ff/foretagsformer/enskild/vad-1.1971 [2020-05-8].

Swedish Companies Registration Office (2019). *Vad är ett kommanditbolag?* https://bolagsverket.se/ff/foretagsformer/kommanditbolag/vad-1.3241 [2020-05-8].

Swedish Companies Registration Office (2020a). *Vad är en ekonomisk förening?* https://bolagsverket.se/fo/foreningsformer/ekonomisk/vad-1.1700 [2020-05-8].

Swedish Companies Registration Office (2020b). *Vad är ett aktiebolag?* https://bolagsverket.se/ff/foretagsformer/aktiebolag/vad-1.3068 [2020-05-8].

Świątkowski, W. & Dompnier, B. (2017). Replicability Crisis in Social Psychology: Looking at the Past to Find New Pathways for the Future. *International Review of Social Psychology*, 30(1), p. 111, doi:10.5334/irsp.66.

Syrett, S. (2012). Conceptualising marginalisation in cities and regions. In: Danson, M. & De Souza, P. (eds.) *Regional development in Northern Europe: peripherality, marginality and border issues*. New York: Routledge.

Taris, T. (2000). *A Primer in Longitudinal Data Analysis*. London: SAGE Publications.

Thurén, T. (2003). *Sant eller falskt? : metoder i källkritik*. Stockholm: Krisberedskapsmyndigheten.

Tillväxt- och regionplaneförvaltningen (2018). *Landsbygds- och skärgårdsstrategi för Stockholmsregionen*. Stockholm.

Tödtling, F. & Trippl, M. (2005). One size fits all? *Research Policy*, 34(8), pp. 1203–1219, doi:10.1016/j.respol.2005.01.018.

Tunón, H., Kvarnström, M., Boström, J. & Utbult Almkvist, A.-K. (2019). Continued use of ecosystems: Challenges for fishing and farming communities. *Baltic Worlds*, XII(2), pp. 40–49.

Türkcan, K. & Erkuş-Öztürk, H. (2019). The impact of economic and political crises on the survival of tourism-related firms: Evidence from Antalya. *Tourism Economics*, doi:10.1177/1354816619868614.

Värmdö Municipality (2018). *Om skolan - Utö skola*. https://utoskola.haninge.se/var-skola/om-skolan/ [2020-03-27].

Värmdö Municipality (2020). *Befolkning*. https://www.varmdo.se/kommunochpolitik/kartorochkommunfakta/befolkning.4.3251048d16e2a7e784836fd8.html [2020-05-21].

Van De Ven, A. H. & Poole, M. S. (1995). Explaining Development and Change in Organizations. *Academy of Management Review*, 20(3), pp. 510–540, doi:10.5465/amr.1995.9508080329.

Wagner, J. (1999). The Life History of Cohorts of Exits from German Manufacturing. *Small Business Economics*, 13(1), pp. 71–79, doi:10.1023/A:1008197109549.

Widholm, C. (2019). Entrepreneurship in the Stockholm Archipelago: A historic perspective. *Baltic Worlds*, XII(2), pp. 57–60.

Xi, G., Block, J., Lasch, F., Robert, F. & Thurik, R. (2020). The survival of business takeovers and new venture start-ups. *Industrial and Corporate Change*, doi:10.1093/icc/dtz076.

Yang, T. & Aldrich, H. E. (2012). Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing*, 27(4), pp. 477–492, doi:10.1016/j.jbusvent.2012.01.001.

Zahra, S. A. (2008). The virtuous cycle of discovery and creation of entrepreneurial opportunities. *Strategic Entrepreneurship Journal*, 2(3), pp. 243–257, doi:10.1002/sej.47.

Zukauskaite, E., Trippl, M. & Plechero, M. (2017). Institutional Thickness Revisited. *Economic Geography*, 93(4), pp. 325–345, doi:10.1080/00130095.2017.1331703.

Appendix I: Correlations Matrix

-	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	1.000																							
2	-0.037***	1.000																						
3	0.146***	**0.803*** **	1.000																					
4	-0.126***	-0.004	-0.000	1.000																				
5	0.092***	0.012	-0.000	-0.769***	1.000																			
6	0.049***	-0.005	-0.000	-0.218***	-0.392***	1.000																		
7	-0.015	-0.019*	0.000	-0.093***	-0.168***	-0.048***	1.000																	
8	0.009	-0.078***	0.000	0.105***	-0.065***	-0.077***	0.047***	1.000																
9	0.027***	0.034***	-0.000	0.020*	-0.087***	0.114***	-0.004	-0.617***	1.000															
10	-0.035***	0.067***	-0.000	-0.150***	0.160***	-0.006	-0.055***	-0.705***	-0.124***	1.000														
11	-0.036***	0.039***	-0.015	-0.031***	0.058***	-0.035***	-0.029***	-0.019*	0.004	0.020**	1.000													
12	0.036***	-0.039***	0.015	0.031***	-0.058***	0.035***	0.029***	0.019*	-0.004	-0.020**	-1.000	1.000												
13	-0.029***	0.038***	-0.000	0.039***	-0.057***	0.019*	0.035***	-0.029***	0.019*	0.019*	0.111***	-0.111***	1.000											
14	0.041***	-0.043***	0.000	-0.027***	0.037***	-0.003	-0.034***	0.028***	-0.019*	-0.018*	-0.515***	0.515***	-0.269***	1.000										
15	0.030***	0.234***	0.428***	0.081***	-0.039***	-0.063***	0.008	0.044***	-0.065***	0.003	-0.189***	0.189***	-0.140***	0.250***	1.000									
16	0.031***	-0.078***	-0.035***	-0.067***	0.072***	-0.007	-0.020**	-0.042***	0.000	0.053***	-0.468***	0.468***	-0.544***	0.520***	0.189***	1.000								
17	0.009	0.015	0.093***	0.039***	-0.031***	-0.014	0.012	0.002	0.002	-0.005	-0.617***	0.617***	-0.421***	0.407***	0.172***	0.686***	1.000							
18	0.006	-0.021**	-0.000	-0.025**	-0.002	0.034***	0.015	-0.003	0.006	-0.001	-0.550***	0.550***	-0.173***	0.308***	0.044***	0.398***	0.580***	1.000						
19	0.001	0.076***	0.099***	0.038***	-0.024**	-0.034***	0.034***	-0.018*	0.061***	-0.032***	-0.024**	0.024**	-0.345***	-0.255***	0.094***	0.220***	0.237***	0.092***	1.000					
20	0.001	0.050***	0.041***	0.060***	-0.048***	-0.025**	0.025**	-0.042***	0.080***	-0.019*	-0.064***	0.064***	-0.084***	-0.176***	0.062***	0.145***	0.227***	0.134***	**0.832*** **	1.000				
21	0.032***	-0.139***	-0.168***	-0.000	-0.054***	0.073***	0.035***	-0.022**	0.049***	-0.016	-0.141***	0.141***	0.175***	-0.130***	-0.154***	-0.194***	-0.188***	0.069***	-0.335***	-0.256***	1.000			
22	0.015	0.161***	0.310***	0.074***	-0.040***	-0.038***	-0.022**	0.060***	-0.049***	-0.032***	-0.441***	0.441***	-0.117***	0.734***	0.530***	0.258***	0.369***	0.254***	-0.206***	-0.164***	-0.127***	1.000		
23	-0.035***	0.431***	0.698***	0.170***	-0.087***	-0.119***	0.013	0.050***	-0.064***	-0.005	-0.032***	0.032***	0.000	0.023**	0.535***	-0.008	0.161***	0.037***	0.136***	0.057***	-0.264***	0.460***	1.000	
24	0.003	-0.092***	-0.198***	-0.030***	-0.003	0.054***	-0.010	-0.026**	0.031***	0.005	0.027***	-0.027***	-0.006	-0.002	-0.169***	-0.020**	-0.088***	-0.004	-0.020**	0.012	0.018*	-0.124***	-0.203***	1.000

*** p<0.01, ** p<0.05, * p<0.1*

1. Failure (exit)
2. Age
3. Age (cluster-mean centered)
4. Joint-stock companies
5. Sole proprietorship
6. Trading/limited partnership
7. Economic association
8. Service industry
9. Trade industry
10. Manufacturing industry
11. Low population density
12. High population density
13. Distance to urban center
14. The land area of islands
15. Fiber broadband infrastructure
16. Access to Primary School
17. Access to kindergarten
18. Access to a year-round grocery store
19. Off-season Passengers
20. High-season Passengers
21. Ratio of high-season Passengers
22. Number of associations
23. Regional GDP
24. Financial crisis

Appendix II: Simple logistic regression table

Model	I	II	III
Age (cluster-mean centered)	1.221 (0.018) ***	1.362 (0.025) ***	1.786 (0.052) ***
Joint-stock companies (ref)	-		
Sole proprietorship	7.410 (1.479) ***	6.257 (1.269) ***	3.985 (0.928) ***
Trading/limited partnership	10.420 (2.340) ***	8.085 (1.844) ***	4.008 (1.046) ***
Economic association	3.328 (1.528) ***	2.949 (1.369) **	2.139 (1.067)
Manufacturing industry (ref)	-		
Trade industry	3.315 (0.773) ***	3.114 (0.737) ***	2.034 (0.524) ***
Service industry	2.615 (0.506) ***	2.603 (0.513) ***	2.046 (0.434) ***
Large population density		2.385 (0.650) ***	1.901 (0.501) **
Distance to urban center		1.006 (0.004)	0.998 (0.004)
Island land area		1.000 (0.000) ***	1.000 (0.000)
Access to fiber		0.645 (0.103) ***	0.880 (0.177)
Access to primary school		1.340 (0.297)	1.257 (0.305)
Access to kindergarten		0.613 (0.183)	0.777 (0.234)
Access to grocery store		0.370 (0.123) ***	0.754 (0.279)
Ratio of summer passengers		1.018 (0.004) ***	1.001 (0.004)
Institutional thickness		0.960 (0.005) ***	0.997 (0.008)
Regional GDP			0.536 (0.026) ***
Financial crisis			1.997 (0.369) ***
Constant	0.003 (0.001) ***	0.002 (0.001) ***	3.348 (2.347) *
Log-likelihood	-1584.802	-1496.185	-1192.617
Pseudo R-square	0.116	0.166	0.254
LR chi-square	416.94	594.18	810.79
Number of observations	9687	9687	8831
Prob > Chi^2	0.000	0.000	0.000

**** p<0.01, ** p<0.05*

Source: Archipelago Organizations Dataset

Appendix III: Map over the core-islands

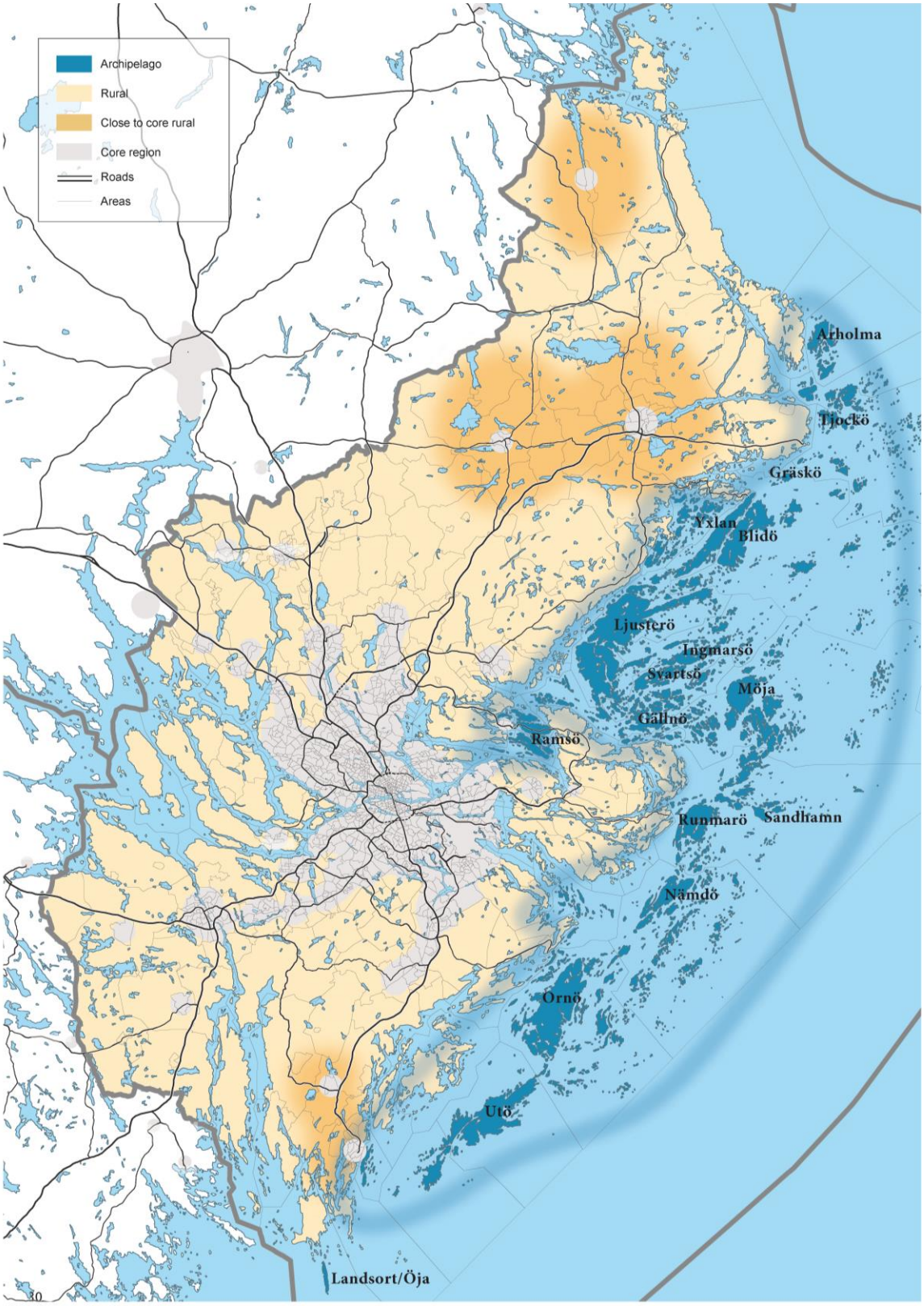

Figure 6: Map over the core-islands colored in blue (adapted from: Tillväxt- och regionplaneförvaltningen 2018, p. 30).