NORWEGIAN SCHOOL OF ECONOMICS
Bergen, Fall of 2023

Decoding the Coded

Understanding the characteristics of the Norwegian software development sector and market

Magnus Nadheim
Halvard Haugse

Supervised by Aline Bütikofer

Master's thesis in Economics and Business administration with a major in Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This thesis provides an in-depth analysis of the dynamic Norwegian software development market, which has undergone significant growth and structural evolution in recent years. It offers a comprehensive view of the market's development, key characteristics, and future directions. The research encompasses a detailed review of the Norwegian software development sector, followed by a robust multi-method quantitative approach for an exploratory and inductive analysis, including descriptive statistics, regression analyses, and survey data. Key findings highlight the market's substantial expansion, fragmentation, and the correlation between higher grossing, more profitable firms and their likelihood for market share growth and productivity. Interestingly, the relationship between R&D investment and expansion is complex and ambiguous. Survey results indicate trends towards diverse acquisition methods and the importance of innovation and emerging technologies in shaping the market's future. This synthesis of insights offers a foundation for future research, emphasizing segment- and technology-specific studies and innovation strategies within the market. The thesis provides crucial insights into the Norwegian software development market's growth, fragmentation, and the interplay between company characteristics and market dynamics, serving as a valuable resource for industry professionals and academic researchers in understanding evolving markets in a technologically advancing era.

Table of Contents

Table of Figures

Table of Tables

Chapter 1 - Introduction

In this introductory chapter of our master thesis, we aim to establish a comprehensive understanding of our study. We will delve into the background and rationale for the analysis, outline the problem statement, articulate the objectives of our research, and delineate the scope of the study. Additionally, we will present the structure of the thesis.

1.1 Background and Rationale

The Norwegian Information Technology (IT) industry has undergone significant transformations over the past four decades. Until the late 1980s, the industry was renowned for its prowess in hardware technology (Lundh, 2018). However, unable to adapt to the transformative and disruptive technological innovations, several Norwegian hardware companies lost their market positions (Steine, 2010).

In later decades, the advancements in hardware have facilitated an increased demand for new software solutions and services. Software is commonly understood as computer code executed on a computer, but it also encompasses a broader range of elements such as specifications, designs, testing regimes, and results, challenging the narrow perception of software as merely computer code (Osterweil, 2018). IBM (n.d.) describes the development of software as *" (...) a set of computer science activities dedicated to the process of creating, designing, deploying, and supporting software"*.

The evolution of smart devices, operating systems like iOS and Android, mobile applications, games, and social networking applications has had a disruptive effect on the global IT industry (Akbar et al., 2015). To keep up with the increased quality standards opposed by users of technology products, several prominent software development methods have emerged (Diansyah et al., 2023). However, despite these advancements in processes, tools and techniques, many software development companies are struggling to meet customer needs (Javeed, et al., 2022).

As digital elements are increasingly incorporated into traditional industries, the boundaries between IT and non-IT sectors are getting blurry (Laato et al., 2022). Many companies decide to recruit in-house software development expertise, and while it excels in control and adaptability, it involves challenges in financial and recruitment aspects, striking a balance between internal cohesion and resource demands (Sadowski & Naborshchikov, n.d.). Thus, many companies seeking to digitalize their organization are faced with the “build vs. buy” decision (Javeed, et al., 2022). Javeed et al. (2022) further highlight that while developing software inhouse may be right in some cases, buying is often more applicable, specifically for small and medium-sized businesses. This has given rise to a booming service sector within the software development industry (Quirt, 2022).

IBM´s (n.d.) definition of software development mainly focuses on the product side of software development. However, it doesn’t entirely capture the increasingly important service aspect of the industry, involving not just the creation of software but also its ongoing support and adaptation to meet evolving client needs and market demand (AxiomQ, 2022). While software products like Microsoft Office or Adobe Creative Cloud offer specific functionalities, software services focus on assisting efficient software use, including installation, integration, support, and maintenance. These services are generally sold separately from software products and aim to enhance customer experience in using software more effectively. As the industry evolves, many companies are moving towards a hybrid model that incorporates both products and services to meet diverse customer needs (Schutz, 2018).

The need for software products and services in business has given rise to IT consulting as a key solution for organizations to address their business challenges through technology (Kumar et al., 2017). Additionally, emerging offerings such as Software as a Service (SaaS) are gaining more traction within the industry. Kulkarni et al. (2012) describe SaaS as a model where software applications are hosted online for business use, providing cost-effective and scalable access with reduced maintenance, but limiting user control over software versions and requirements, thereby eliminating the need for on-premises hardware and software management. Another emerging service offering is white label (off-the-shelf) solutions, offering a re-brandable, standardized product suitable for various industries, a time- and cost-efficient solution, but with limited customization options in terms of specific features and updates (Silva et al., 2020). Outsourcing of software development is also a common practice for cost reduction and accessing skilled labor but involves risks such as uncertainty of technical skills, poor communication, and unclear requirements, as identified in a systematic literature review by Wahab and San (2018).

Adapting to the structural changes in the global trends, a robust software development market has emerged within the Norwegian IT industry (Irascu, 2023). Evolving at a rapid pace, this growing sector has become an increasingly more important part of the overall Norwegian IT industry, impacting economic activity and everyday lives. Characterized by rapid technological advancements and innovations that render older solutions obsolete more quickly than in other industries, the importance for individual companies to innovate and progress is profound. The sector's socio-economic significance and fast-paced environment makes for a fascinating subject to explore.

Our study aims to provide a holistic analysis of the trends and characteristics of the software development sector in Norway. Although the established diversity in sector makes it somewhat opaque and fragmented, by aggregation we aim to shed light on its complexities, composition, and overall trends.

1.2 Problem Statement and Research Questions

This thesis seeks to explore the intricate landscape of the Norwegian software development sector and market. We will conduct the study by investigating trends and relationships while we examine how emerging technologies and market dynamics shape this sector. The underlying central question is:

What are the recent and expected characteristics of the Norwegian software development sector and market?

To address the problem statement, the thesis is structured in a way that targets the following three research questions:

1. *What are the recent trends in the Norwegian software development sector and market?*
2. *What are the characteristics of companies that gain market shares, and which aspects indicate current market position?*
3. *What are the most common offerings and preferences regarding software solutions, and what are the perceptions and anticipated future trends in this sector and market?*

1.3 Scope of the Study

There is no straight forward way of isolating and defining the Norwegian software development sector. The rapid digitalization of economics has sparked a wide range of IT-solutions, often intertwined. Brønnøysund Register Centre (BRC), however, classifies Norwegian companies into industry groups based on their core activity, assigning each group a NACE code (Nomenclature of Economic Activities). We utilize these definitions to get the most representative scope of the software development sector.

The IT service industry (NACE code 62) consists of four groups: (1) Programming services, (2) IT consulting services, (3) Management and operation of IT-systems, and (4) Other services associated with IT. Programming services (62.010) includes the development, modification, testing, and support of customized software to meet specific client needs. This encompasses entities dedicated to creating software solutions uniquely aligned with individual customer requirements, while explicitly excluding activities such as software package publishing and comprehensive system designs integrating with various IT components (SSB, 2009). This group therefore encapsulates companies with core activities regarding software development.

IT consulting services (62.020), on the other hand, covers a broader spectrum of consultancy services related to both computer hardware, software and information technology (SSB, 2009). This category extends its reach beyond mere software development, encapsulating a wider range of services. However, it is important to recognize that many IT consulting firms typically do not restrict their expertise exclusively to software development, although it is a central service offering (Kumar et al. 2017). Including such companies is therefore considered to be essential to encapsulate the full diversity of the software development sector, but presents certain challenges in controlling for other activities not directly related to software development.

Groups of management and operation of IT-systems (62.030), and other services associated with IT (62.090), both appear to be outside the definition of the software development sector (SSB, 2009). While they encompass aspects like data system management and IT support services, they do not specifically include the core activities of software development such as programming, IT consultancy, and software innovation, which are central to the nature of software development companies. Therefore, we only consider programming services (62.010) and IT consulting services (62.020) as the software development sector.

Furthermore, we consider only limited liability companies (AS) and publicly limited companies (ASA), as they best represent the most relevant landscape of the industry. Other ownership structures such as sole proprietorship, while interesting in and of itself, capture an even broader and more diverse range of dynamics, and generally cannot compete with the service offerings of limited liability companies and publicly limited companies (Sparebank 1, n.d.).

In concrete terms, the software development sector therefore encompasses a range of activities and types of companies, either defined as *custom software development companies* or *IT consulting firms* and classified as *limited liability companies* or *publicly limited companies*.

1.4 Objectives of the Study

Our study is guided by several objectives, closely related to our problem statement and research questions. To understand the evolution of the software development sector, we will analyze data from the last two decades, providing a baseline for current and future market assessments. Conducting statistical analyses, we seek to identify relationships between indicators such as market share, market share growth, innovative activities, and profitability, within different company size segments. This analysis aims to provide an overview of key market characteristics. Gathering market insights from the software development market, we seek to understand perspectives, preferences, expectations, and attitudes. We will not consider the functional aspects of the technologies being discussed, but rather their relevance in industry and market. Through these objectives, our study aims to offer a detailed and nuanced understanding of the Norwegian software development sector and market and its trajectory.

1.5 Structure of the Thesis

In Chapter 1, we introduce the problem statement and objectives of our research for analyzing the software development industry and markets in Norway. Chapter 2 delves into the previous literature on the Norwegian IT-sector. In Chapter 3, we describe our methodology, detailing the analytical tools and techniques used in our study. Chapter 4 presents descriptive insights of the software development market. In Chapter 5 and Chapter 6, we present and briefly discuss our statistical findings and market insights. Chapter 7 encompasses a broader discussion of the findings considering our research questions. Finally, Chapter 8 concludes our thesis, summarizing the key findings, insights, and implications, while also reflecting on the study's limitations and suggesting areas for future research.

Chapter 2 - Literature on the Norwegian IT-Sector

There is very limited research on the Norwegian software development market. Nevertheless, a master's thesis from 2020, conducted by Berli and Hundhammer from the Norwegian School of Economics provides amongst other things a strategic analysis of profitability in the Norwegian IT consulting sector (Berli & Hundhammer, 2020, p. 24). Berli and Hundhammer use the Pestel framework as presented by Peterdy (n.d.) and Porter's five forces as theoretical models for their strategic analysis (Porter, 2008).

Berli and Hundhammer's analysis of Norway's IT consulting sector uses a definition that in principle encapsulates the entire industry for IT services (Berli & Hundhammer, 2020, p. 25). This expansive scope includes the software development sector as we define it, in addition to other IT-related sectors. Thus, we find it more appropriate to describe their scope as the IT service industry, instead of merely as the IT consulting sector. Nevertheless, their broad definition increases the transferability of their findings to our study. Our scope, however, targeting a substantial portion of the IT service industry (custom software development companies and IT consulting firms) offers a more focused analysis, while building upon Berli and Hundhammer's foundational insights.

The thesis' PESTEL analysis reveals several key factors influencing the IT service industry (Berli & Hundhammer, 2020, p. 29-32). The political and regulatory environment in Norway, characterized by a favorable political climate and moderate tax levels, creates a secure business setting. This is further bolstered by the government's significant investment in digitalization, particularly in information and communications technology (ICT) and research and development (R&D). Economically, the industry's sensitivity to currency fluctuations, notably the value of the Norwegian Krone, and the dependency on imported hardware, are crucial factors. Additionally, oil prices and production have a significant impact, as evidenced by the effects of the 2014 oil price drop on the market. Social and cultural trends show a growing demand for sustainable development and digital solutions, which in turn boosts the demand for IT services. This trend is accompanied by an evolution in consumer behavior, with an increased reliance on technology. Technological advancements are central to the sector's growth, with firms continuously innovating and either developing proprietary solutions or marketing existing products. Finally, the legal and regulatory landscape, especially stringent data protection and privacy laws in Norway and Europe, plays a significant role in shaping the industry’s offerings, underscoring the diverse factors that collectively influence the IT service industry.

Berli and Hundhammer’s implementation of Porter’s Five Forces highlights profitability (Berli & Hundhammer, 2020, p. 32). They identified a moderate threat of new entrants, citing that the market is not fully saturated and having low entry barriers due to minimal capital requirements and remote work feasibility. Yet, they emphasize the importance of networking and project agreements, particularly for larger firms. Establishing an IT consulting firm is relatively easy compared to sectors like the oil industry, with the IT service industry marked by diverse company sizes and frequent mergers and acquisitions.

In assessing the threat of substitutes, they consider various alternatives like regular consulting firms, in-house IT departments, hardware companies, and other service providers (Berli & Hundhammer, 2020, p. 33). The competitiveness of these substitutes depends on client expertise, pricing strategies, and service quality. The industry’s competition is relatively high, involving a mix of small startups and large corporations. However, the level of competition varies based on specialization, with niche firms often facing less competition in less saturated domains, unlike generalists in highly competitive areas like Oslo.

The bargaining power of customers varies with the complexity of services, being higher for standard services due to easy comparisons, and lower for unique, customized projects because of high switching costs (Berli & Hundhammer, 2020, p. 33-34). Suppliers strengthen their bargaining power by integrating their unique solutions or software into customer operations and holding exclusive rights or partnerships. Yet, in the IT services industry, human capital is the primary resource, influencing supplier power based on employee availability and the number of firms in the market.

Industry rivalry, particularly high in densely populated areas with many competitors like Oslo, can potentially lead to price wars, though the potential for remote service delivery can mitigate direct competition (Berli & Hundhammer, 2020, p. 35). Additionally, the competition for skilled IT professionals is a notable factor, influencing wages and long-term costs. The sector’s reliance on human capital could balance these increased labor expenses by reducing the need for significant physical infrastructure investment.

Berli and Hundhammer's also propose a statistical analysis of the 12 largest IT consultancy firms (Berli & Hundhammer, 2020, Chapter 7). We raise some concerns about their models' high explanatory power (74%, increasing to 85-86% with fixed effects) and potential model robustness issues. Furthermore, the study's small sample size of 12 companies over a 10-year period risks endogeneity. Focusing on a few large companies could introduce systematic bias, with larger companies having systematically different characteristics than the rest of the market. These factors challenge the validity and applicability of their paper and overall conclusions.

The strategic analysis from Berli and Hundhammer's (2020) thesis serves as a solid foundation for our analysis. However, it's crucial to address a key limitation in directly applying their results. The broad definition of the IT consulting sector is somewhat imprecise. The complexity of the IT service industry makes it challenging to apply broad strategic frameworks uniformly across the entire selection. Each sector within the industry could potentially yield distinct findings if analyzed separately. Therefore, while considering their analysis as a contextual and strategic foundation, we must exercise caution in generalizing their findings across the entire industry, as it does not provide an adequate consideration of its nuances.

Chapter 3 - Methodology

3.1 Brief Overview

This chapter provides a comprehensive overview of the methodological framework employed to examine the research questions. The methodological choices are anchored in the structured approach as outlined by Saunders et al. (2019).

Initially, we introduce the study object, highlighting the specific focus of our analysis. This is followed by a detailed exposition of the research design, which lays the foundation for our methodological choices. Subsequently, we delve into the specifics of data collection, describing both the sources of our data and the methodologies employed in gathering it. The chapter then transitions into a discussion of the ethical considerations and measures taken to ensure the integrity of our research process.

Concluding the methodological overview, we reflect on the quality of the research by assessing the reliability and validity, as well as the constraints of the study. This includes an evaluation of the rigor and reliability of our methods, ensuring that our findings are both robust and credible. Through this structured approach, the chapter aims to provide a clear and thorough understanding of the methods and principles guiding our research.

3.2 The Study Object

The study object for most of the analysis is companies operating within the software development sector. As outlined in Section 1.3, NACE codes 62.010 (Programming services) and 62.020 (IT consulting services) appear to be most relevant. Based on Brønnøysund Centre Registers definitions, we therefore primarily include *custom software development companies* and *IT consulting firms*. Additionally, for parts of the analysis we will narrow our focus to Oslo, to capture indications of possible future trends, insights that might be possible to transfer to the overall industry and market. This is further discussed in Section 3.4.1.

3.3 Research Design

3.3.1 Research Approach

There are three main research approaches: *deductive*, *inductive*, and *abductive* (Saunders et al., 2019, p. 153-156). A deductive research approach sets out to either verify or falsify a hypothesis grounded in existing literature. In an inductive approach, however, the researcher uses data to generate and build new theories. Abductive reasoning seeks to develop further existing theories where data breaks the theoretical expectations. Due to the unexplored topics of our research questions, this study has an inductive research approach. Rather than explicitly proposing new theories, we will reflect on the essence and possible reasons for the findings and propose the implications for future trends.

This study is primarily a *multi-method quantitative study*, leveraging both numerical and categorical data to draw its conclusions (Saunders et al., 2019, p. 178). As will be elaborated in Section 3.4.1, our research also incorporates a minor use of qualitative data. The inclusion of qualitative elements suggests a *mixed-method* approach. However, we advise caution in fully labeling it as such, due to the predominant emphasis on quantitative techniques. While the qualitative component adds depth and context to our findings, the core of our analysis is grounded in quantitative data analysis.

3.3.2 Nature of the Research Design

In general, research is often oriented to be *exploratory, descriptive, explanatory, evaluative*, or a combination of these (Saunders et al., 2019, p. 186). Exploratory studies aim to investigate topics not previously thoroughly understood, to discover new insights and generate hypotheses for future research. Descriptive studies focus on accurately portraying events or phenomena, providing a clear picture of the current aspects. Explanatory studies delve into understanding the causality of occurrences or relationships. Evaluative studies assess the impact of various interventions, often aiming to provide actionable recommendations based on the findings. This research primarily aims to (1) describe the recent and current market landscape for software development companies and (2) explore the relationships between various quantitative metrics such as growth, market size, and other relevant variables, as well as the intricate dynamics shaping the market and how these may influence future trends. This study should therefore be considered both descriptive and exploratory.

3.3.3 Time Horizon

In this research, the time horizon encompasses both a *longitudinal* and *cross-sectional* perspective. As Saunders et al. (2019, p. 212) elaborates, a longitudinal perspective captures a series of observations over an extended period. In contrast, a cross-sectional perspective focuses on a specific point in time, providing a snapshot of the datapoints.

From a longitudinal viewpoint, our study leverages panel data spanning the years from 2000 to 2022. Specifically, in the descriptive part of Chapter 4**,** we examine the recent trends between 2000 to 2020, with some metrics extending up to 2022. The decision to limit certain analyses to data up to 2020 is driven by data quality considerations, as further detailed in Section 3.4.2. Our statistical analysis, however, has more of an exploratory character and examines the decade from 2010 to 2020. This decade is marked by significant technological advancements, including the rise of social media and smartphones. To best capture the most relevant and recent characteristics of the market, we find it appropriate to exclude earlier years from the statistical analysis.

We additionally adopt a cross-sectional perspective, through capturing and describing certain incidence and characteristics at a given point in time. The cross-sectional perspective predominantly considers the primary data, in a descriptive and exploratory manner (Section 3.4.1). Although the panel data is inherently longitudinal, it is initially treated as cross-sectional in the statistical analysis to provide a comparative analysis. This dual approach, integrating both longitudinal and cross-sectional perspectives, enables a multi-layered understanding of the sector, capturing both its evolution over time and its condition at specific junctures.

3.4 Collecting Data

This paper utilizes both *primary* and *secondary* data. Primary data consists of gathering new data, while secondary data is initially collected for some other purpose (Saunders et al., 2019, p. 338). Incorporating primary and secondary data gives the study an analytical depth, ensuring a robust examination of the software development sector and market. However, it should be noted that the secondary data undergoes more extensive statistical analysis compared to the primary data, as its numeric nature is more amenable to statistical analysis.

3.4.1 Primary Data

The primary data consists of cross-sectional categorical data and a small amount of qualitative data, collected through a *survey strategy* utilizing *internet questionnaires*. As stated by Saunders et al. (2019, p. 505) questionnaires are a good complement to other data collection methods and are often used for exploratory and descriptive research. Our questionnaires seek to complement the inductive approach of the thesis, by providing primary data based on our own unique questions.

The goal of the questionnaires is to map the software development market's supply and demand dynamics. Suppliers are identified as software development firms within the previously specified NACE-codes. Reflecting the insight of Kuiken (2022), it's anticipated that all companies will increasingly demand digital solutions and services. Thus, demand is defined as companies outside the specified NACE-codes for software development companies. This broad categorization aids in understanding market potential, yet caution is advised against rigid interpretations. While this approach simplifies, it offers a foundational exploration of the market landscape amidst evolving digital needs.

To define the population as something more manageable, we focus on a subset of each population, namely two *target populations.* The target populations for both the supply and demand are companies located in Oslo. The rationale for this subset is that metropolitan areas tend to be the first movers in digital trends (Golding, 2023). This selection could thereby provide a snapshot of early market changes and reveal trends that later could be applicable for the rest of the country. On the other hand, we risk applying Oslo specific biases to the general market. This will be discussed in greater detail in Section 3.7. For ease of referencing, we will hereby address the suppliers of software development engaged in the questionnaires as *vendors*. The potential customers of software development services, i.e. the demand, will be referred to as *clientele*.

As it would be impracticable to collect data from the entire target populations, we utilize *random sampling* to ensure representative samples. Random sampling provides an unbiased sampling method. Since the purpose is to gain insights and explore trends rather than to make generalizable inferences, the sample size in relation to population size is less strict (Saunders et al., 2019, p. 155). However, to reduce the impact of potential outliers, and increase credibility and robustness, we use a sample size of 300 for both the *vendors* and the *clientele*. Assuming conservative response-rates of 15-20%, we expect roughly 50 responses per questionnaire (Baruch & Holtom, 2008).

For each company, our primary focus is on engaging the chief executive officer (CEO), ensuring that the most knowledgeable individual with a comprehensive view of the business and market is providing the response. Using a combination of information on LinkedIn, publicly available registers, and the company's web page, we ensure that we are reaching out to the correct individuals. Utilizing Qualtrics, the links to the questionnaires are sent to the research objects through either text or email. Qualtrics collects both complete and partially complete responses, but we consider only the completed.

The questionnaires for the *vendors* and *clientele* cover many similar topics, and where the subjects overlap, both phrasing of the questions and the available response choices are identical. However, some questions are unique for each questionnaire. For a complete overview of both questionnaires, please see Appendix A.4. Each question is standardized to ensure similar interpretations by all respondents and meticulously designed to capture precise data pertinent to our research question. By neither adopting nor adapting questions used in other questionnaires, we focus on constructing unbiased and cleverly worded questions, to enhance the internal validity of the questionnaires, i.e., the ability to measure what we intend to measure (Saunders et al., 2019, p. 217).

We are predominantly using forced-choice questions, but the use of some open-ended questions provides some qualitative insights. However, as previously mentioned, the qualitative data makes up only a fraction of the primary data, and are meant to supplement the categorical findings, rather than being methodically analyzed. Additionally, (1) utilizing conditional follow-up questions based on specific responses, (2) focusing on a clear and visually appealing presentation for both web and mobile, (3) explaining the purpose of the questionnaires, (4) keeping the questionnaires relatively short, and (5) allowing respondents not to answer, are some of the measures taken to increase response rates and reduce the risk of respondent biases.

3.4.2 Secondary Data

The secondary data consists of longitudinal quantitative *numeric* and *categorical* data, primarily sourced from *Regnskapsdatabasen,* a comprehensive database of core financial and corporate information on Norwegian corporations (Mjøs & Selle, 2022, p. 1). The database, intended for research and educational purposes, has prior to our collecting undergone rigorous quality control to maintain high data quality. Key steps include careful selection of data sources, primarily reputable governmental institutions, and implementing several checks to identify errors in reporting, such as algorithm-based flagging of irregularities in financial data. Additionally, in-depth studies of select firms are conducted for consistency checks, and user feedback is actively sought to identify and correct any errors (Mjøs & Selle, 2022, p. 5). This database is therefore particularly well suited for empirical studies due to its high quality and digitally accessible register data from the generally well-documented economy of Norway.

To supplement the data from Regnskapsdatabasen, which only extends up to 2020, we incorporate additional, up-to-date market information from Proff Forvalt. Proff Forvalt, a branch of Proff AS specializing in credit and accounting information, is widely acknowledged as a trustworthy source for Norwegian businesses and public entities (Proff Forvalt, n.d.). Similar to Regnskapsdatabasen, Proff Forvalt primarily sources its data from reputable government institutions. However, unlike Regnskapsdatabasen, Proff Forvalt does not explicitly state its involvement in rigorous quality control processes. This lack of specified quality assurance could mean that their data is more susceptible to both systematic and non-systematic errors, particularly those arising from manual data entry into official registers by companies (Mjøs & Selle, 2022, p. 5). Therefore, while Regnskapsdatabasen remains our primary source for secondary data, we use Proff Forvalt's data selectively only when appropriate to complement Regnskapsdatabasen.

3.5 Data Analysis Techniques

To analyze the primary and secondary data, we utilize several data analysis techniques for quantitative data. Exploratory data analysis (EDA) is used for presenting and exploring the primary data and the overall market trends and data distributions of the secondary data. Furthermore, statistical methods for examining relationships and trends are done utilizing different types of regression analysis on the secondary data. We divide the secondary data into different subsets by company size for parts of the analysis.

As already established, this paper does not particularly consider qualitative data, but it is worth mentioning that the small amount of such data that is collected, will be analyzed through a *thematic narrative analysis,* i.e. addressing analytical themes in narratives, to complement the categorical findings (Saunders et al., 2019, p. 675). However, we do not extensively elaborate on this due to its insignificance in the overall analysis.

3.5.1 Heterogeneity Analysis

While analyzing the entire population is done for most of the exploratory data analysis, dividing the data into smaller target populations is beneficial for the regression analysis. As there exist vast differences and few common denominators between certain groups, dividing the companies into subsets by size can provide clearer patterns in the regression models. This will give more insightful and thorough answers to Research Question 2. We group the data into *very small, small, medium-sized,* and *large* firms, with a slight adaptation of the size limits for SMBs (Small and Medium-sized Businesses) (Regjeringen, 2019). *Very small* firms are companies with less than 10 employees and with revenue and/or annual balance sheets below 20 MNOK. *Small* firms are companies with less than 50 employees and revenue and/or annual balance sheets below 100 MNOK. *Medium-sized* firms are companies with less than 250 employees and revenue below 500 MNOK and/or annual balance sheets below 430 MNOK, and *large* firms are companies with more than 250 employees. Because we want to measure productivity, defined as $\frac{Sales\ Revenue}{Employees}$, companies with zero employees are not included in the company size segments (Holliday, 2021).

3.5.2 Exploratory Data Analysis

Exploratory Data Analysis consists of techniques used to investigate and summarize the main characteristics of a dataset, often in a visual manner, and is often used for uncovering underlying structures, identifying important variables, detecting outliers and anomalies, and testing assumptions. EDA is a critical first step in the data analysis process as it allows for gaining insights and a deeper understanding of the data's nature and behavior (Saunders et al., 2019, p. 581-597). This approach is particularly useful in guiding the selection of appropriate statistical tools and models for further analysis and in preparing data for more formal and complex analyses.

In Chapter 4**,** we use exploratory data analysis to visualize market trends in the secondary data from the year 2000 to 2022. However, due to the slightly inferior data quality and lack of overlap with Regnskapsdatabasen, as shown in Figure 3 in Section 4.1, data from Proff Forvalt is used only for certain descriptive purposes and not for extensive analysis. Hence, the majority of Chapter 4 revolves around the years 2000-2020, and has a descriptive nature, providing valuable insights for answering Research Question 1.

In Section 6.1, we present the findings of the primary data through proportions and charts, employing Exploratory Data Analysis (EDA) as a first step to gain insights into the market and understand its nature. This analysis, which is both descriptive and exploratory, aims to provide inductive insights that contribute to the overall thesis, specifically addressing Research Question 3. Due to practical limitations in conducting extensive analyses on the secondary data, our approach with the primary data is less statistically intensive. However, this EDA on the primary data lays a solid foundation for future research, offering an initial understanding and paving the way for more in-depth statistical exploration.

3.5.3 Regression Analysis

Regression models are commonly used for estimating relationships between variables, and for identifying trends in data. They are especially valuable in quantifying the strength of the relationship and in making predictions based on the observed data patterns (Hassan, 2023a). Our regression analyses aim to address Research Question 2 with an emphasis on intuitive and straightforward models that can effectively uncover relationships. Therefore, we have opted for simpler models like linear and logistic regressions, prioritizing interpretability over the potentially increased predictive capabilities of more advanced alternatives.

We will consider two types of regression models: (1) linear regression models and (2) logistic regression models. Linear regression models are used to investigate linear relationships with a response variable and one or more explanatory variable(s). A logistic regression model, however, considers the logarithm of the odds (log odds) of a binary outcome based on one or more explanatory variable(s) and is thus able to capture non-linear relationships (Hassan, 2023a).

The linear and the logistic models will be fitted to each of the company size segments mentioned in Section 3.5.1, using the secondary data from Regnskapsdatabasen from 2010 to 2020. Due to the rigorous quality measures done by Regnskapsdatabasen itself, not much cleaning of the data is necessary. However, due to the disturbing effect of outliers on both linear and regression models, extreme values are removed per subset. We adopt the approach of (Soetewey, 2020), and remove those below the 2.5% percentile and above the 97.5%.

Both models feature explanatory variables carefully chosen to best analyze the key characteristics of the software development market, further discussed in Section 5.2.1. Our analysis employs two variants of the linear regression model: one treats the data strictly as cross-sectional for comparative purposes, while the other adjusts for the panel data structure inherent in the dataset through incorporating fixed effects. The results will be discussed and used as a basis for generating new ideas, encompassing the inductive nature of the research.

Baseline Linear Regression

In our linear regression model, we explore the linear relationship between the logarithm of market share (log-transformed market share) and a set of explanatory variables, which will be listed in Appendix A1. The baseline multiple linear regression model is formulated as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n + \varepsilon$$

Here, Y is the response variable (logarithm of market share), and X_i are the explanatory variables. The β coefficients (β_i) represent the marginal effect on Y associated with a unit increase in X_i, while holding other explanatory variables constant. The term β_0 is the baseline value of Y when all X_i are zero. The term ε represents the error term, capturing the variation in the response variable Y that isn't explained by the explanatory variables X_i, as well as the inherent randomness of the data.

For the interpretation of a linear regression model to be valid, it must satisfy several assumptions: (1) a linear relationship between the response and explanatory variables, (2) independent residuals, (3) constant variance of residuals (homoscedasticity), (4) no multicollinearity among explanatory variables, and (5) normally distributed residuals. These assumptions ensure the reliability and validity of the regression results. In Section 5.3.3, we will apply well-known statistical tests to confirm if our data meets these criteria.

In Chapter 5, we start by fitting the linear models in their baseline form and then proceed to adjust for fixed effects.

Linear Regression with Fixed Effects

The fixed effects linear regression models account for firm-specific and/or time-specific influences that might impact the data but are not directly included in the models. For firm-specific factors, we might include industry-specific regulatory frameworks or specific leadership strategies. These factors are intrinsic to individual firms and may significantly influence their performance. For time-specific fixed effects, we might consider yearly changes in economic policies, global market trends, or annual technological advancements. These temporal factors can impact all firms in the dataset but vary over different time periods. When adjusting for fixed effects, some of our models include only firm-specific fixed effects, while others include both firm- and time-specific fixed effects. The formulation of a linear regression model adjusted for only firm-specific fixed effects looks as follows:

$$Y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \cdots + \beta_n X_{nit} + \alpha_i + \varepsilon_{it}$$

In this model, Y_{it} represents the response variable for firm *i* at time *t*, and X_{nit} denotes the explanatory variable *n* for firm *i* at time *t*. The β coefficients indicate the marginal effect on the response variable associated with a unit increase in each explanatory variable, while holding other variables constant. However, this interpretation is specifically within the context of changes within each entity over time, not across different entities.

In many fixed effects models, the overall intercept term β is not explicitly reported or is considered redundant, as the model focuses on the differences within entities over time rather than the absolute levels. Therefore, each firm has its "own intercept" captured by α_i. Lastly, ε_{it} is the error term, accounting for random variation in the response variable that is not explained by the fixed effects or the explanatory variables.

In extending our model to incorporate both firm- and time-specific fixed effects, the formulation is augmented to:

$$Y_{it} = \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \cdots + \beta_n X_{nit} + \alpha_i + g_t + \varepsilon_{it}$$

In this model, we enhance the firm-specific fixed effects approach (α_i) by introducing time-specific fixed effects (g_t). While α_i still captures the unique, unchanging characteristics of each firm, g_t accounts for common factors affecting all entities but varying by time period, such as annual economic shifts or policy changes. This inclusion of g_t allows the model to simultaneously adjust for unobserved variations at both the individual and temporal levels. The β coefficients continue to measure the marginal effects of explanatory variables, now within a more comprehensive framework that controls for firm-specific and time-specific unobserved factors.

Logistic Regression

In the logistic regression model, we want to see if any of the selected explanatory variables discussed in Section 5.2.1 and presented in Appendix A.1 seem to relate to the likelihood of growth in market share. The formulation of a logistic regression model is as follows:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n + \varepsilon$$

In a logistic regression model, the response variable is expressed as the logarithm of the odds (log odds), $log\left(\frac{p}{1-p}\right)$, where p is the probability of a binary outcome occurring. Similar to linear models, X_i denotes the explanatory variables. The β coefficients (β_i) quantify how a unit increase in an explanatory variable affects the log odds of the binary outcome, with other explanatory variables held constant. The term β_0 is the log odds of the outcome when all explanatory variables are zero.

Unlike linear regression, logistic regression has different assumptions for valid interpretation. The key assumptions include: (1) the response variable is binary, representing two possible outcomes (e.g., "growth" or "no growth" in market share), (2) observations are independent, (3) there is an absence of multicollinearity among explanatory variables, (4) the model assumes a linear relationship between continuous explanatory variables and the log odds of the response variable, (5) the sample size must be sufficiently large, and (6) there should be few to none extreme outliers. To ensure these assumptions are met, we apply various statistical tests to our models.

3.6 Ethical Considerations

Regarding the ethical considerations of this study, we will address the sourcing and handling of primary and secondary data. Access to the secondary data was provided from Regnskapsdatabasen, on the premise that it was not to be used outside the scope of the thesis. The data from Proff Forvalt is accessible through a paywall, but with less constrained use cases.

We strictly use the data only for the purpose of our thesis. Additionally, no specific companies are mentioned, avoiding any ethical considerations regarding incorrect descriptions.

For the primary data, although no personal information is gathered, we have taken several measures to ensure that ethical principles are upheld. This commitment is evidenced by both questionnaires being registered at Sikt, a Norwegian governmental administrative agency under the Ministry of Education (Sikt, n.d.). Privacy is a key principle when considering the ethical standards of gathering primary data. The most relevant aspects for us to consider were (1) the voluntary nature of the participants, (2) informed consent, (3) ensuring confidentiality, and (4) responsibility in the analysis of data (Saunders et al., 2019, p. 258).

During the recruitment for participation, all individuals were free to choose their involvement without any coercion or incentives that could influence their decision against their will. Moreover, participants were granted the option to abstain from answering any question and had the freedom to discontinue their participation in the questionnaire at any point. Each questionnaire starts with a consent letter that outlines the implications of participating in the survey. Before starting the survey, participants were required to give their consent to participate in the study. They were informed about the study's objectives and the handling of their data. Those who did not consent were promptly excluded from the survey. Additionally, all participants were assured that their responses would be kept confidential, and their anonymity maintained throughout the thesis.

In the data analysis, all information is presented in a cumulative manner to ensure a responsible evaluation and to maintain the anonymity of the respondents. The collected data is stored securely on the Norwegian School of Economics' server and retained only for the duration necessary to fulfill the research objectives.

3.7 Reliability and Validity

The quality of research is commonly evaluated based on its *reliability* and *validity*. These criteria measure the consistency and reproducibility of the research, as well as the accuracy and appropriateness of the analysis, procedures and measures used (Saunders et al., 2019, p. 213).

Reliability

Reliability in research signifies the extent to which a study's methodologies, experiments and analyses can be replicated by other researchers in similar conditions with identical subjects, yielding consistent outcomes (Hassan, 2023b). This section evaluates the reliability of our study, explaining how the results are not simply coincidental or exclusively influenced by the unique conditions of this study.

The secondary data consists of self-reported financial data from each company. Many limited liability companies and publicly limited companies utilize auditors, which reduces the risk of self-reported errors from the companies. Furthermore, Regnskapsdatabasen has implemented rigorous testing and correction of the data, to ensure reliability of the data set. As a distinguished database renowned for its consistently high-quality data curated specifically for research, it lays a solid groundwork for reproducible outcomes. Moreover, our cautious approach to incorporating data from Proff Forvalt, especially our practice of cross-referencing with Regnskapsdatabasen to ascertain data suitability, further reinforces the reliability of the secondary data. However, it is worth noting that ensuring the reliability of such vast datasets is challenging, and the room for human errors both in the reporting and error-testing should be considered.

In the primary data we consider the *participant error, participant bias*, *researcher error* and *researcher bias.* Regarding participant error, we engaged the participants over multiple weeks at different times. While timing may affect the way in which a participant responds, this approach seeks to average out such effects. Allowing respondents to answer according to their own schedule is a measure to reduce participant bias, which may occur under forced circumstances (Saunders et al., 2019, p. 214). Utilizing questionnaires reduces the researcher error, as it provides clear answers with the possibility for analysis in an appropriate setting. Finally, research bias is reduced by having both researchers involved in the data collection and analysis process, ensuring the analysis is not overly influenced by subjective views.

Validity

Validity refers to the degree to which our methods, measurements, and analyses accurately capture and reflect what they are intended to measure. As Saunders et al. (2019, p. 215) elaborates, it is important to assess the *construct*, *internal* and *external* validity, to demonstrate how the findings reliably represent the dynamics within the software development sector, ensuring they are not merely artifacts of our methodological choices or influenced by external variables.

Construct validity refers to what extent the research measures what it claims to measure (Hassan, 2023c). In this study we wanted to examine unexplored aspects in the software development industry in Norway, regarding historical patterns and relationships within market composition, growth, and innovation. The exploratory nature of the thesis implies an inductive perspective, and utilizing a multi-method quantitative approach we consider both primary and secondary data to get a comprehensive foundation for answering our research questions. Firstly, we use exploratory data analysis to get a historical overview of the market. Next, we utilize statistical techniques to analyze key variable relationships. Finally, incorporating both *vendors* and *clientele* questionnaires uncovers insights that are not available elsewhere. Additionally, by predominantly capturing responses from CEOs, we ensure that the insights and data gathered reflect a high level of expertise and informed perspective. Using standardized questions ensures a consistent response pattern, as all respondents encounter the same set of questions. By providing clear descriptions and ensuring that all respondents interpret the questions similarly, we ensure that we measure what we intend to measure. However, the primary data being collected only from Oslo should be considered, as there might be systematic differences between Oslo as a subset and the rest of the country, making the results less applicable for the Norwegian software development sector.

In further considering the construct validity, it is worth mentioning that the secondary data captures financial data from the relevant NACE-codes (62.010 and 62.020); however, this might include more or less than the population of software development companies in Norway. For instance, 62.020 contains companies with more diverse service offerings than merely software development, implying that the study object might not be entirely representative for the population. Furthermore, companies might have different NACE-codes over time, due to BRCs classification of their core activities (Brønnøysundsregistrene, 2023). Some companies, such as consulting firms, typically include multiple activities in their operations, and over time might be subject to changes in their main NACE-code. This is especially apparent for the largest companies in the market, which are often involved in mergers and acquisitions that might influence the definition of their core activities. The dataset and the market boundaries therefore risk being disturbed by companies leaving and entering the sector due to changes in their NACE code classifications. This could imply that measurements are somewhat affected by aspects not meant to be measured.

Internal validity refers to whether the study's results truly stem from the examined relationships, and not from any design flaws or other external factors (Saunders et al., 2019, p. 215). In this study, the statistical relationships of the secondary data utilize financial data to measure complex phenomena such as innovation and market share growth, which due to its complexity, might be hard to measure. However, utilizing multiple regression models, and incorporating fixed effects, presents a more nuanced perspective on variable relationships. Furthermore, through sensitivity analysis we evaluate the impact of the variables on model fits. Nevertheless, these findings should be viewed as indicating strong correlations within the data, rather than definitive causal relationships. Additionally, while the primary data aims to complement the regression analysis, it is not used for establishing statistically significant relationships.

External validity concerns representability of the findings, i.e. the extent to which the results of a study can be generalized to other settings or groups (Hassan, 2023d). The secondary data consisting of Norwegian financial data may limit the generalization of the findings to other countries, due to specific regulatory and reporting frameworks governed by Brønnøysundregistrene and Regnskapsloven (Norwegian Accounting Act) (Wojtecka, 2023). Additionally, while the research on primary data focuses on Oslo, these localized findings could potentially be extrapolated to other metropolitan areas or regions within Norway or other countries, albeit with some caution. Metropolitan areas often share similar economic and infrastructural characteristics that might make these findings relevant to other urban contexts (Dallhammer, et al., 2019). However, differences in regional market dynamics and local business environments should be taken into account when attempting such generalizations. Accounting for time specific aspects in the statistical analysis while using primary data to consider future trends, is a measure for generalizing the results considering the rapid change and innovation of the sector.

3.8 The Constraints of the Study

In this section, we address the constraints of the study. It's vital to recognize potentially inherent constraints in the methodology to ensure a balanced understanding of our conclusions regarding the past and future trends within the software development market.

Addressing the study object is crucial for considering fundamental limitations. The Norwegian software development sector is characterized by its broad scope and diverse nature, featuring a wide array of companies ranging from traditional IT consulting to niche players (Advania, 2023). These firms therefore might not operate exclusively in the same market, limiting a direct comparison within the sector.

With our data analysis techniques, we focus on capturing overarching market trends in and outside of size specific subsets of the data rather than delving into specific local trends within segments. This approach offers a comprehensive market view but might exclude finer details of certain aspects.

The survey methodology is constructed to accommodate a diverse range of respondents, and we found it necessary to simplify the questions to ensure they were universally applicable. This approach yields a valuable general overview of the market but sacrifices some technical specificity and expertise.

A key limitation of our study is the uncertainty in discussing future market trends. Factors like short term market trends, economic downturns, or unforeseen technological advancements can significantly deviate the outlook.

Chapter 4 - Descriptive Insights of the Sector

This chapter will analyze the recent trends in the software development sector in Norway, leveraging data from Regnskapsdatabasen and Proff Forvalt, spanning from 2000 to 2020 in most cases, and 2022 for some metrics. Our goal is to provide a comprehensive overview of the trends in sales, employment, competition, and overall sector health, as this is crucial for understanding the market's trajectory and for providing valuable insights for answering Research Question 1. It is worth mentioning that a span of 20-22 years in the software development sector represents a substantial portion of the historical trends, as the sector is relatively young.

4.1 Market Growth

When analyzing growth over time, it is common to evaluate the Annual Compound Growth Rate (CAGR). This is calculated, when applicable, using the following formula:

$$CAGR = \left(\frac{Value_{final}}{Value_{initial}}\right)^{\frac{1}{t}} - 1$$

Total CPI Adjusted Sales

First, we consider the total sales in the market, represented by the total amount of revenue generated from sales, accumulated for all companies. To analyze the real increase in sales over the past decades without including inflation, we adjust by the CPI with 2020 as the reference year. Figure 1 illustrates the total sales in the Norwegian software development market, measured in billions of Norwegian kroner (BNOK). This market has experienced remarkable growth over the past 22 years. With 20.1 BNOK in 2000 adjusted for CPI, sales revenue increased to 87 BNOK by 2020 and further to 104 BNOK in 2022. This fivefold increase highlights the market's significant expansion. It's important to note that the data we have from Regnskapsdatabasen covers the period up to 2020. To extend our insights to 2022, we include data from Proff Forvalt. While these two sources don't align perfectly, as discussed in Section 3.4.2, the overall trends show a consistency regarding total sales.

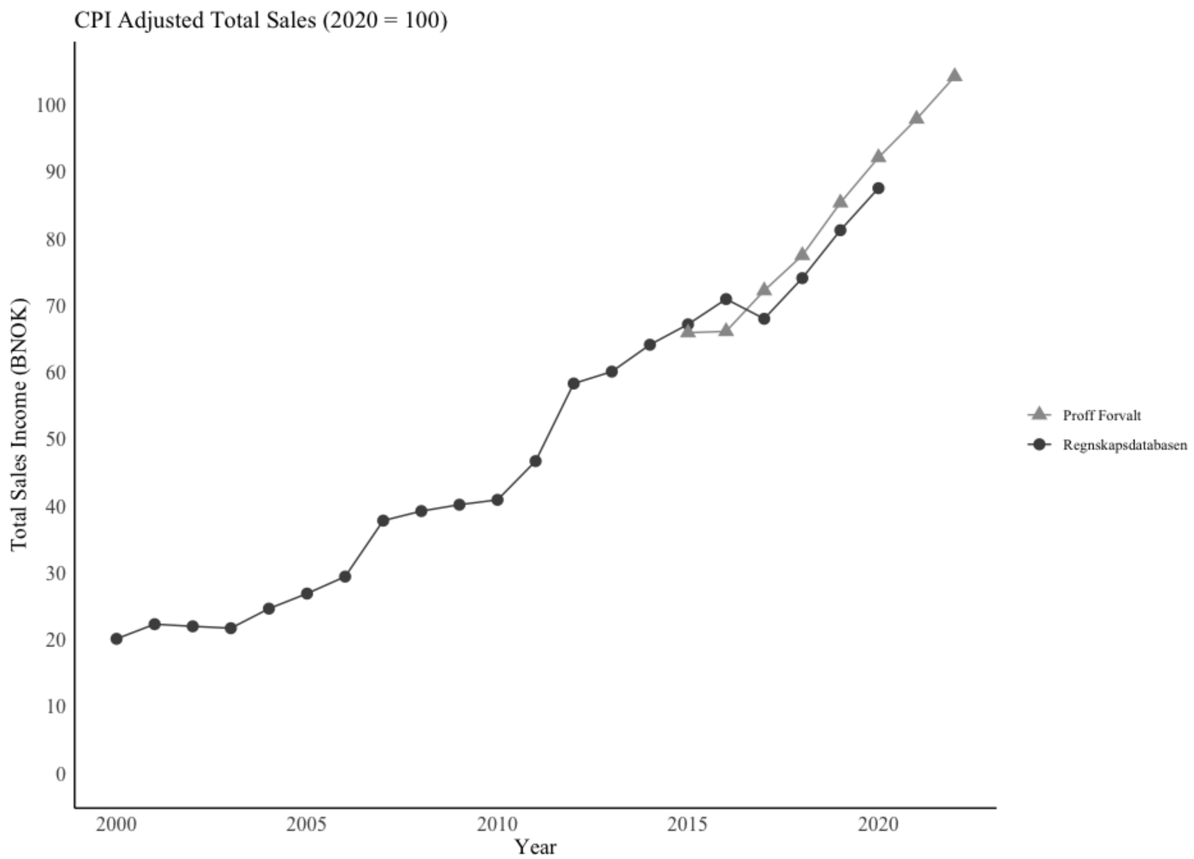


Figure 1. CPI Adjusted Total Sales (2020 = 100) *Data source: Regnskapsdatabasen & Proff Forvalt*

We notice a robust growth in total sales with a CAGR of 7.76%. A significant inflection point occurred around 2010, when the CPI adjusted total sales was around 40 BNOK. Since then, in the last 12 years, there has been a higher linear increase in total sales than between 2000-2010.

Interestingly, the rapid growth in the years after 2010 suggests that the great recession started by the financial crisis in 2008 did not seem to adversely affect the sector. This pattern appears to have been the case for IT sectors not only in Norway, but several other economies (Rosenberg, 2018). Furthermore, as stated by LaBerge et al. (2020), the Covid-19 pandemic in 2020 served as a digital catalyst for many industries. This could potentially explain why the software development sector does not appear to have been significantly negatively affected during this recent crisis either.

Number of Companies

While the total sales reveal a substantial increase in the past decades, it’s also important to consider the change in the number of unique companies within the sector when assessing market growth. As evidenced in Figure 2, there has been a substantial increase in the number of unique active companies. From 1,341 companies in 2000, the market consists of 11,557 in

2022. We notice more than an eightfold increase, and a CAGR of 10.3%. The year 2010 also marks a key point of acceleration regarding the total number of companies. With a market composition of 3,433 companies in 2010, an increase of 8124 companies up until 2022 reveals that the subsequent growth has been especially high, underpinning the discussion regarding the years post the global crises in the previous paragraph.

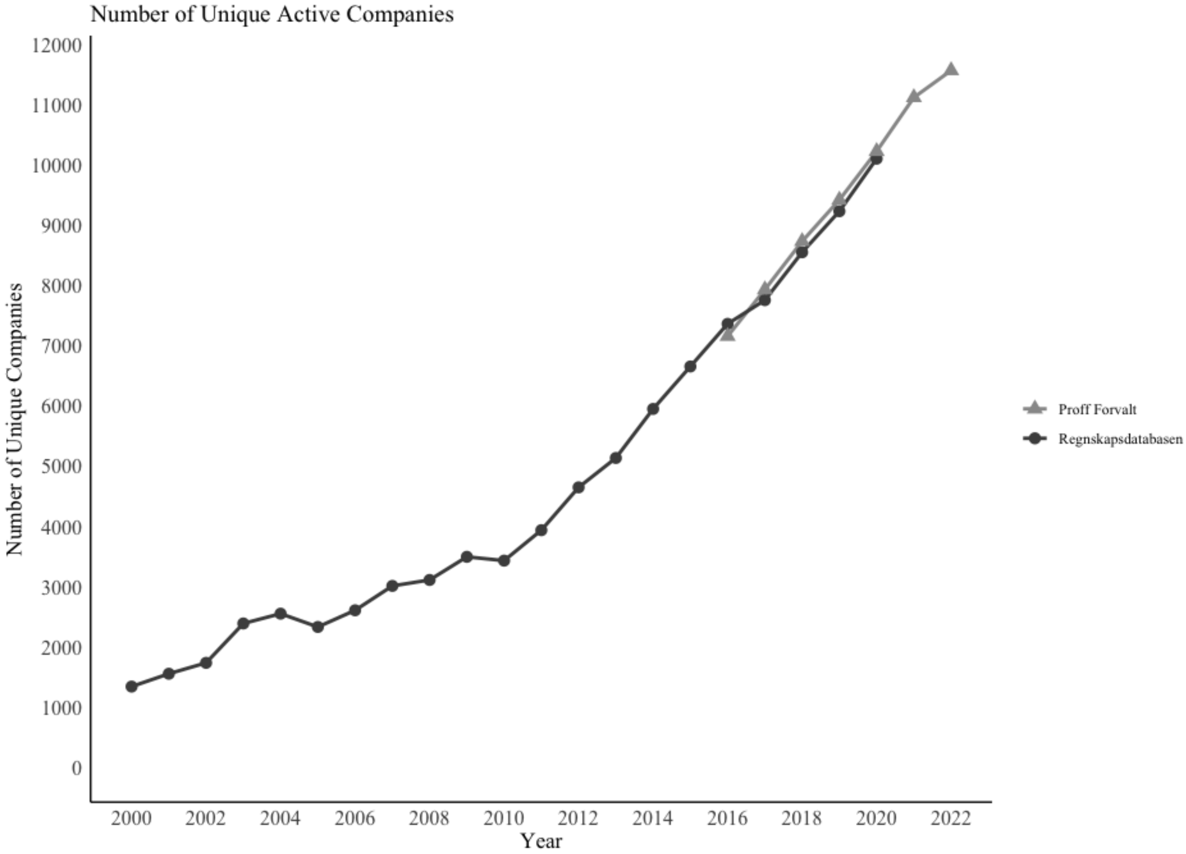


Figure 2. Number of Unique Active Companies *Data source: Regnskapsdatabasen & Proff Forvalt*

The increased number of companies indicates that the market is expanding and branching out. Contrary to a scenario of market consolidation, where growth is driven predominantly by existing actors, the Norwegian software development sector appears to be moving towards a more fragmented landscape. We notice that the CAGR of the total number of companies is higher than the CAGR for CPI adjusted total sales. This trend implies that the sector's expansion is not solely concentrated among a few dominant companies but is instead characterized by a broadening of the competitive landscape.

Employment patterns

It is natural to expect that as the number of companies increases, so does the total employment. Upon examining the employment patterns over the past two decades, it becomes evident from Figure 3 that the data from Proff Forvalt is not suitable for analysis in this context. Unlike for total sales and number of unique companies, the employment data from Proff Forvalt contain numerous missing values and inconsistencies. This results in an unrealistic and dramatic decline in the number of employees. These inconsistencies are evident in several other metrics, and given these anomalies, we will exclude Proff Forvalt data from any further analysis. This also applies to the regression analyses in Chapter 5, as there are too many missing values to derive meaningful and consistent insights from the data. However, the discrepancy is an interesting observation, highlighting the impact of systematic or non-systematic errors of the self-reported financial data.

Interpreting only the data from Regnskapsdatabasen, we are limited to the year 2020. In 2000, the sector employed 11,549 people. Since then, there has been a steady increase over the years, reaching 44,118 by 2020. This more than fourfold increase in employment is notable, giving a CAGR of 6.93%.

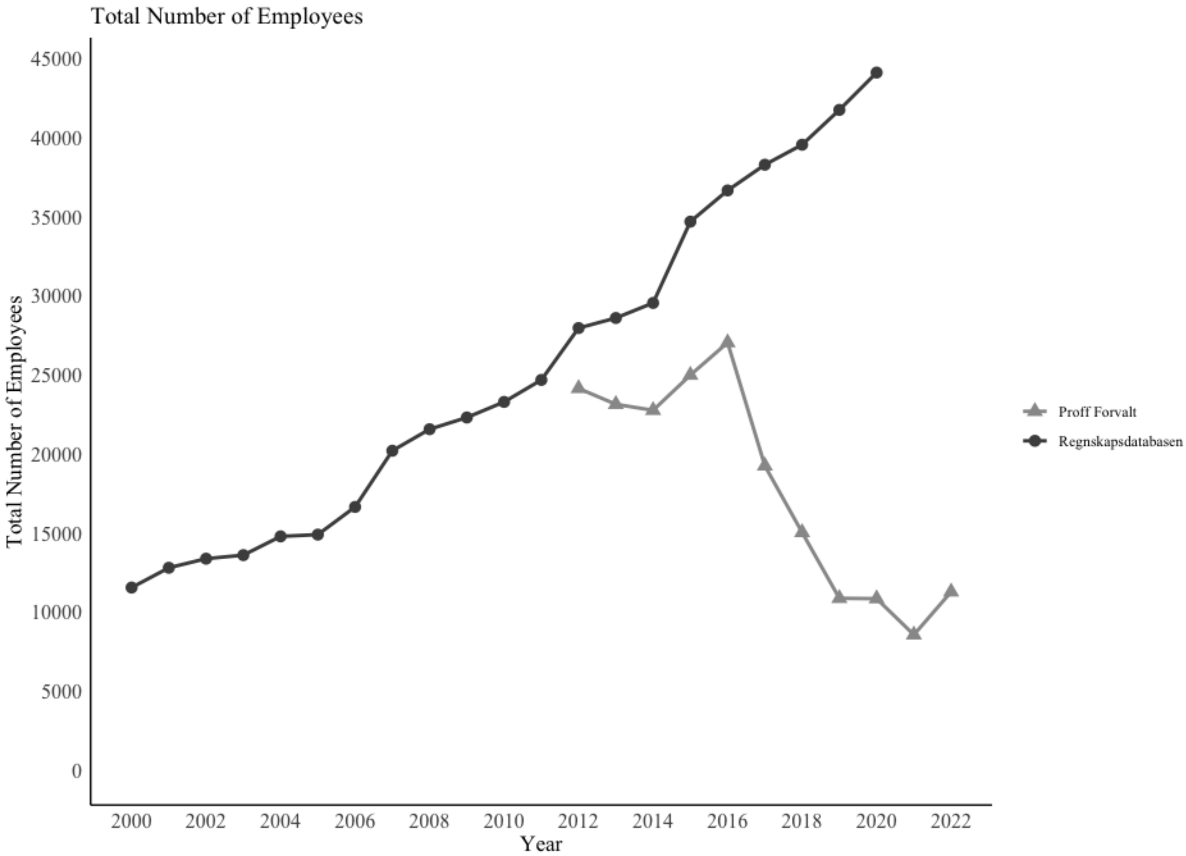


Figure 3. Total Number of Employees *Data source: Regnskapsdatabasen & Proff Forvalt*

The disparity between the growth rates in total sales and the number of employees is also particularly revealing. While both metrics have shown an upward trend, the sales income has grown disproportionately higher than the total number of employees, with a CAGR of 0.83

percentage point more. This could imply an increase in productivity, which will be further examined in the next section. As for the other metrics, the sector does not seem particularly affected by the great recession regarding unemployment, with a steadily increasing employment pattern throughout the years.

4.2 Average Metrics

Average Company Size Measured by Sales

Moving on from total metrics considering the market growth, we will now consider average metrics to gain insights into trends of the characteristics for the average company over the past decades. First, we examine the yearly CPI adjusted average sales. Figure 4 reveals that, in year 2000, the CPI adjusted average sales was at approximately 15 MNOK. However, during the dotcom burst between 2000-2002, the sector experienced a substantial decline, before stabilizing (Lebo, 2019). By 2003, the average sales had dropped to about 11.3 MNOK. Between 2003 and 2012, the average sales fluctuated, reaching an average sales income adjusted for CPI at around 12.5 MNOK in 2012. Post 2012, there was a five-year period of a sharp decline in the average sales per company. This could be attributed to the increase in number of companies, as discover in Figure 2, and supports the narrative of a global increase in tech-startups, thereby driving down average sales, as startups typically have lower sales than established firms (Moayed, 2021). In 2017, the CPI adjusted average sales was about 8.8 MNOK. From 2017 to 2020, the CPI adjusted average sales income has remained relatively stable, with 2020 having an average of roughly 8.7 MNOK.

$$\frac{Sum\ of\ CPI\ adjusted\ sales_t}{Number\ of\ unique\ active\ companies_t} \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

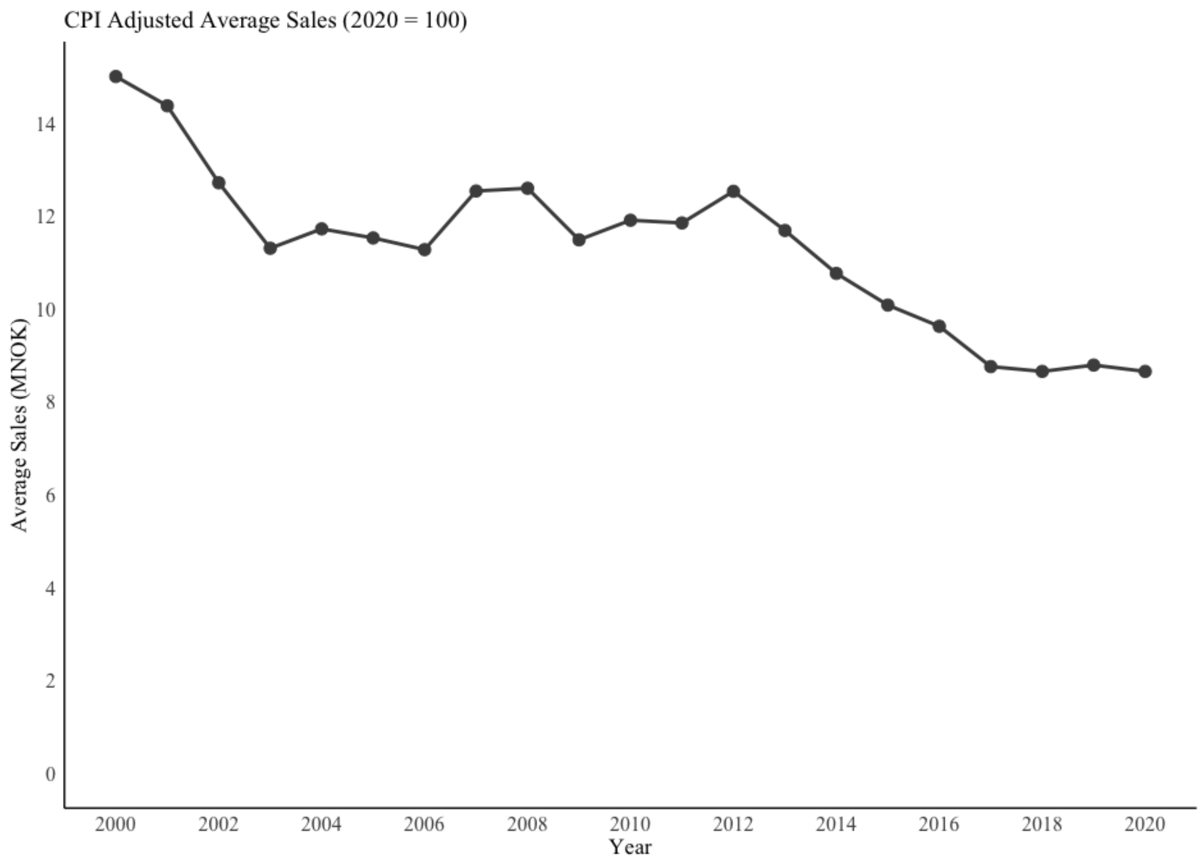


Figure 4. CPI Adjusted Average Sales (2020 = 100) *Data source: Regnskapsdatabasen*

Average Number of Employees per Company

Next, we consider the average number of employees per company, where Figure 5 reveals an interesting pattern over time. Despite yearly fluctuations, a distinct downward trend is apparent, indicating a decrease in the average number of employees per company. In 2000, the average company had 8.69 employees. By 2020, the average number of employees per company was 4.36, marking a 50% reduction over two decades. Notably, the 2020 number represents the record low in the last 20 years. The highest number of employees per company was reached in 2007, with an average of 9.16 employees per company. 2008 marked the year of the financial crisis and experienced a substantial decline. Subsequent years are characterized by a general decline with yearly fluctuations. The overall increase of total employment, revealed in Figure 3**,** combined with the notable decline in the average number of employees per company evident in Figure 5, suggests a growing presence of smaller companies in the sector. Average number of employees per company is calculated as:

$$\frac{Sum\ of\ employees_t}{Number\ of\ unique\ active\ companies_t} \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

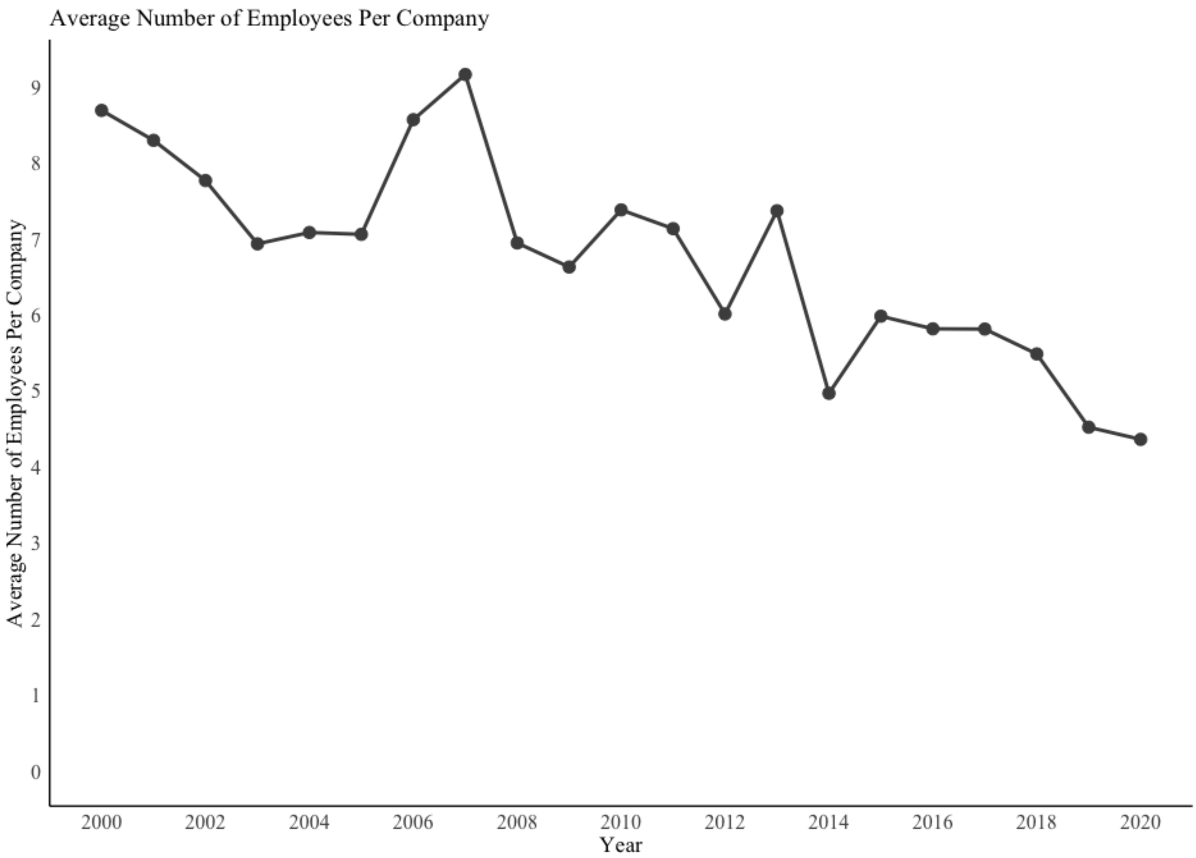


Figure 5. Average Number of Employees Per Company *Data source: Regnskapsdatabasen*

Average Profitability as Measured by Weighted Average EBITDA Margin

Regarding the average profitability, we are using the average EBITDA margins (Earnings Before Interest, Taxes, Depreciation, and Amortization) weighted by sales, per company. EBITDA margin is widely recognized as a reliable measure of operational profitability because it is similar to the companies' cash flows (Westberg, 2023). The weighted average EBITDA margin is calculated using the following formula, were N_t is the number of unique companies in a given year *t*, and *i* represents each unique company:

$$\left(\frac{\sum_{i=1}^{N_t}(EBITDA\ margin_{it} * Sales_{it})}{\sum_{i=1}^{N_t}(Sales_{it})}\right) \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

Figure 6 indicates that the weighted average EBITDA margin has experienced significant fluctuations between 2000 and 2020. In 2001, the year after the dotcom crack, there was a sharp decline in average profitability. Profitability dropped to a negative 0.9%, indicating a substantial average loss across the sector.

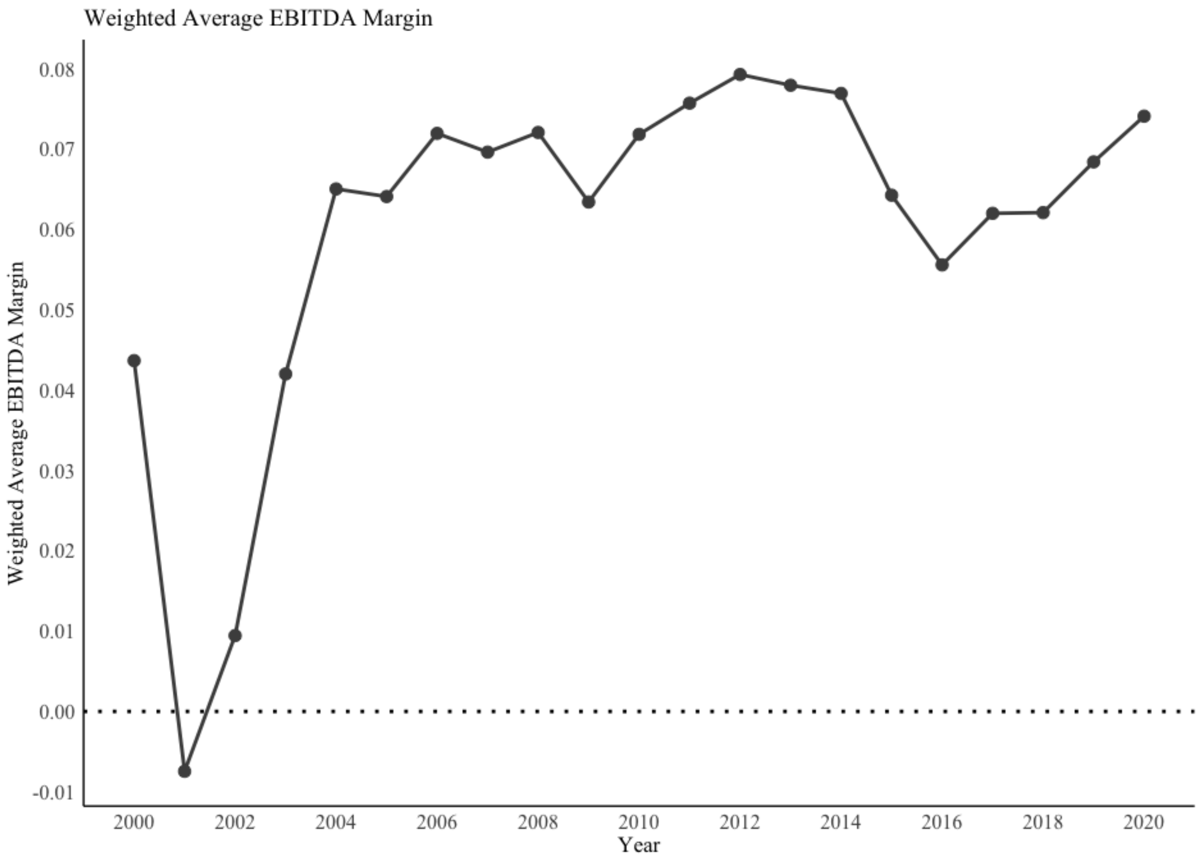


Figure 6. Weighted Average EBITDA Margin *Data source: Regnskapsdatabasen*

However, post 2003, the sector demonstrated a more stable profitability range. From 2004 to 2020, the average EBITDA margin consistently hovered between 6% and 8%. Specific years like 2010, 2012, and 2020 saw margins of 7.2%, 7.8%, and 7.4%, respectively. These figures reflect a reasonably healthy profitability level in relation to sales, suggesting a robust margin for the market. The great recession did not seem to affect the profitability in the market as strongly as the dot com crash did.

Average Productivity

As mentioned in Section 4.1**,** the growth in CPI adjusted sales has been higher than that of the number of employees during the past two decades. As previously mentioned, a common metric for productivity is sales in relation to number of employees. As we are interested in the yearly productivity in the sector, we reduce the impact of potential outliers by calculating it as follows:

$$\frac{Sum\ of\ CPI\ adjusted\ sales_t}{Sum\ of\ employees_t} \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

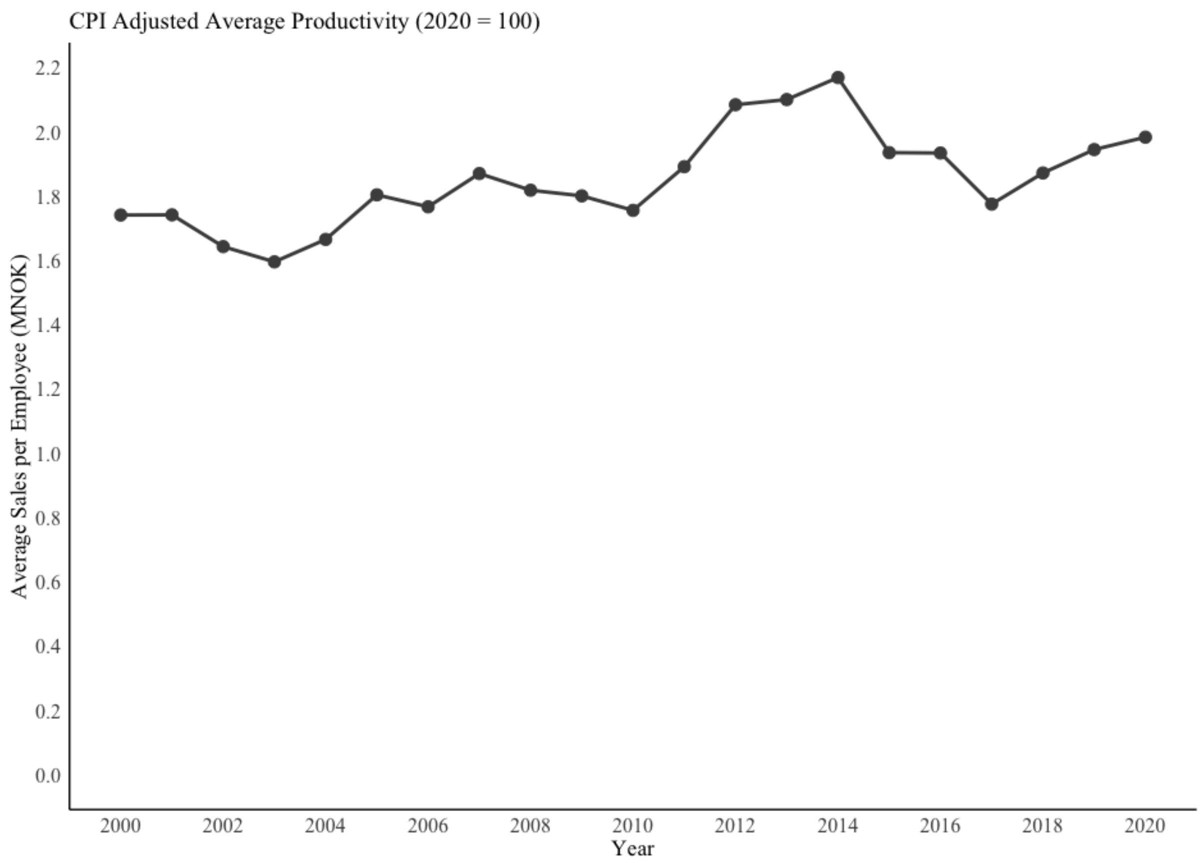


Figure 7. CPI Adjusted Average Productivity (2020 = 100) *Data source: Regnskapsdatabasen*

Examining Figure 7, we notice some fluctuations in productivity throughout the years, but with an underlying positive trend. In 2000 there was an average of 1.74 MNOK in CPI adjusted sales per employee, increasing to 1.98 MNOK in 2020. The least productive year appears to be 2003, with an average of 1.6 MNOK in CPI adjusted sales per employee. 2014 seems to be a peak year for productivity, with an average of 2.17 MNOK in CPI adjusted sales per employee. As will be elaborated in Section 4.3, however, some anomalies may have had an impact between the years 2011 and 2016.

Average R&D Expenses

Next, we examine research and development expenses, as these can be used as a benchmark for measuring the innovative activities in a company (Zebrabi, 2023). Examining the average of this parameter can therefore reveal some interesting trends regarding innovation. The metric is calculated as follows:

$$\frac{Sum\ of\ CPI\ adjusted\ R\&D\ expenses_t}{Number\ of\ unique\ active\ companies_t} \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

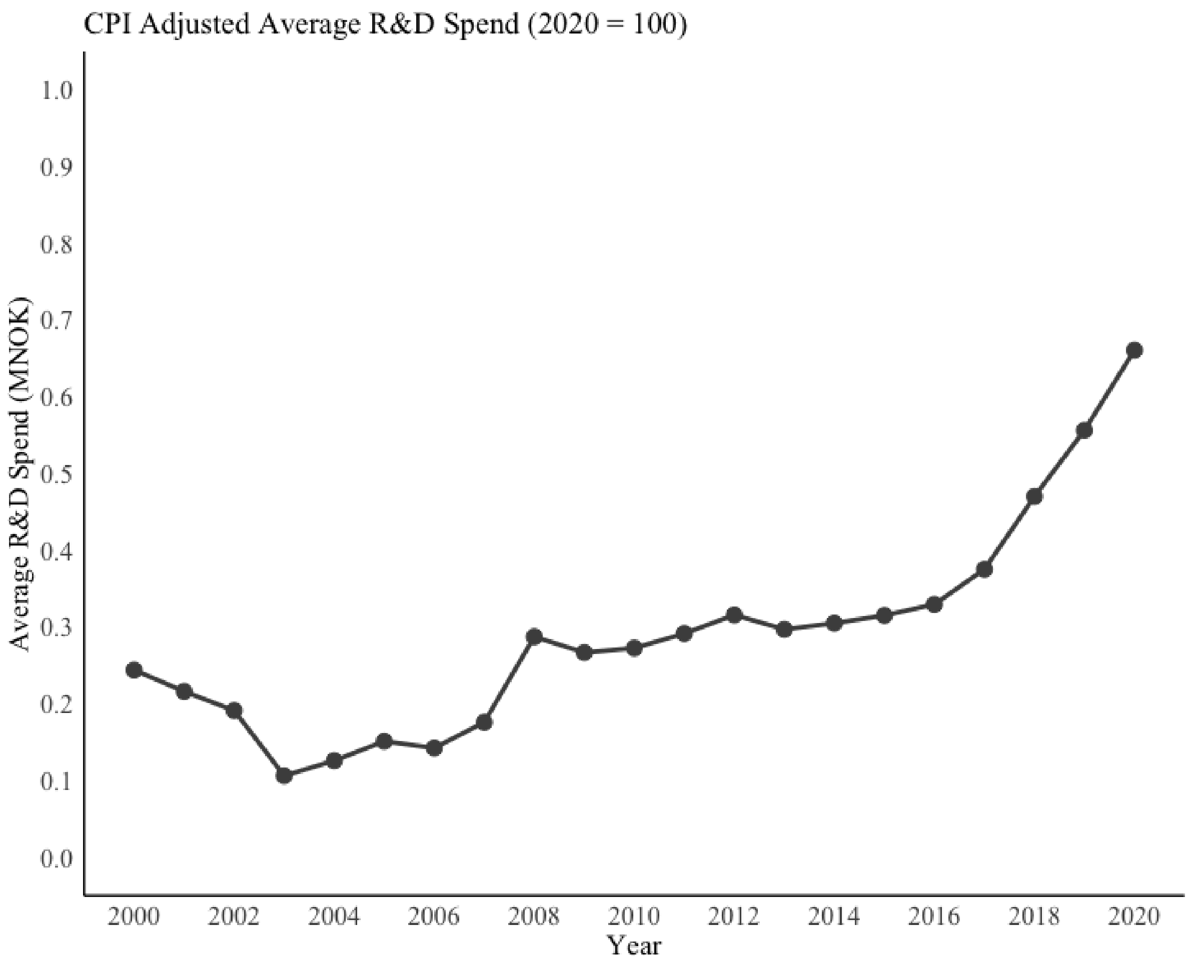


Figure 8. CPI Adjusted Avereage R&D Spending (2020 = 100) *Data source: Regnskapsdatabasen*

In Figure 8**,** we observe a positive trend in average R&D expenses over the past two decades. The average spending has grown from 0.24 MNOK in 2000 to 0.661 MNOK in 2020, nearly tripling in size. Particularly notable is the increase in R&D expenses post 2017. This period aligns with key developments in artificial intelligence, most significantly marked by the publication of the *Attention Is All You Need* paper (Vaswani et al., 2017). This paper introduced the Transformer model, a novel approach to training Large Language Models, which has since had a profound impact on the evolution of AI and its applications across various industries (Murgia, 2023). The correlation between this landmark publication and the surge in R&D

spending could potentially underscore the software development sector's commitment to advancing AI technologies. This observation is noteworthy: despite a decrease in the average size of companies, as indicated by both sales and employee numbers, there has been a notable increase in average R&D expenditure. This trend underscores a significant emphasis on innovation in the sector.

4.3 Market Concentration

Market Shares of the Largest Companies

We will now consider market shares, as measured by proportions of total sales, offering insights into market dominance and evolving trends. In Section 4.1, we noticed that the CAGR for total unique companies has been greater than for CPI adjusted sales in the last two decades. To investigate this further, we'll examine the market share of the top 5 highest grossing companies for each year. However, it should be noted that some of these companies may have undergone mergers and ceased operations. The top 5 highest grossing companies, therefore, do not necessarily consist of the same companies over time. The market share of these companies per year, is calculated as the following:

$$\frac{Sum\ of\ CPI\ adjusted\ sales\ in\ 5\ highest\ grossing\ firms_t}{Total\ CPI\ adjusted\ sales_t} \quad \forall t \in \{2000, 2001, \ldots, 2020\}$$

Examining Figure 9, it is evident that the proportion of the total sales attributed to the top 5 highest grossing companies has steadily decreased. In 2000, these firms accounted for 40% of the market's total sales. By 2020, their share of the market's total sales had dropped to 14%. The decline was relatively stable between 2000 and 2011, reaching 18% in 2011. A temporary increase to around 25% occurred in 2012-2013, a spike caused solely by a tech giant, as they acquired another large company, causing them to be classified within NACE code 62.020 (Jørgenrud, 2010). However, post 2016 the merged organization was no longer categorized as an IT-consulting firm. After the exit, the decline continued towards 14% in 2020.

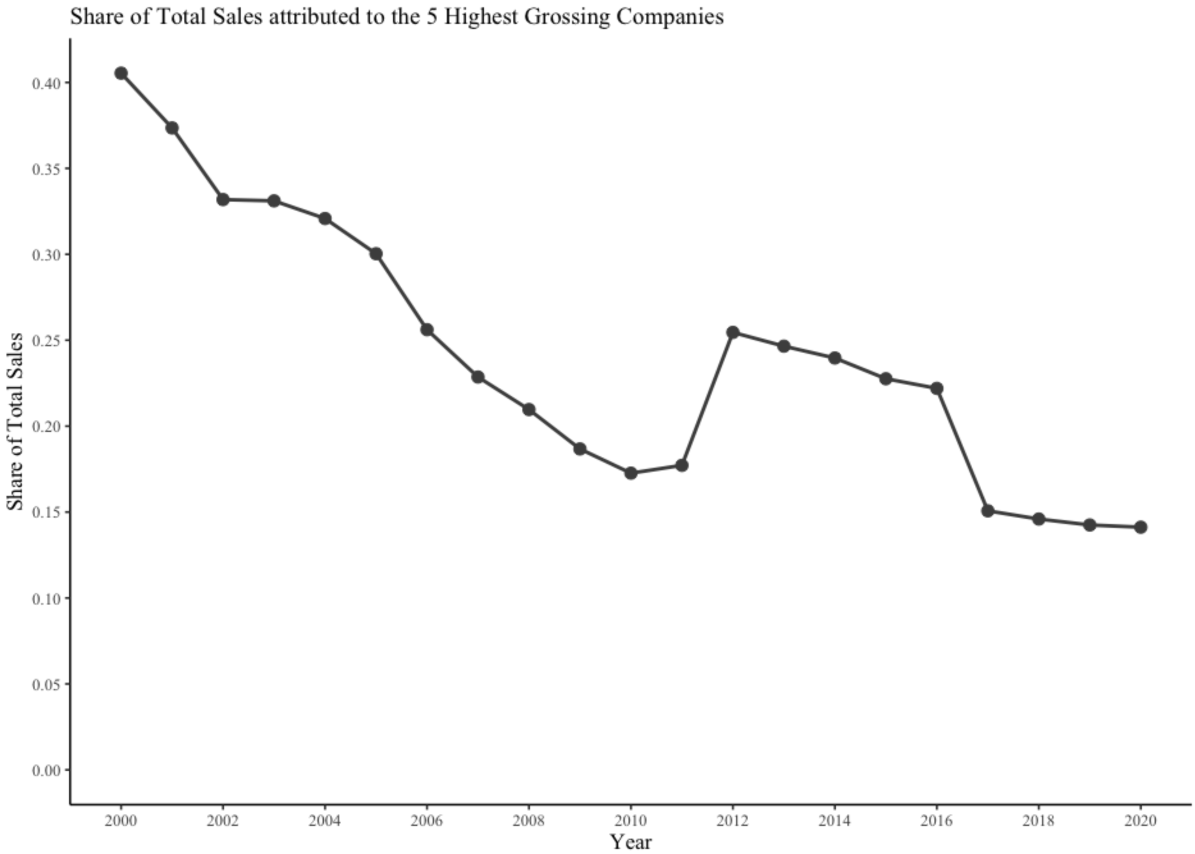


Figure 9. Share of Total Sales attributed to the 5 Highest Grossing Companies *Data source: Regnskapsdatabasen*

Figure 9 suggests that the highest grossing companies are losing their previous dominance in the market. While the largest firms haven't necessarily ceased growing, the market itself has expanded significantly, with an influx of new companies, thus diluting their relative share amidst this broader landscape, as discovered in the previous sections.

Market Shares of Company Size Segments

Further addressing the changes in market composition, we analyze the evolving trends in the market shares of the different company size segments; *very small, small, medium-sized* and *large* firms (Section 3.5.1). Market share is measured as a company's sales in relation to the total sales of the sector, and thereby calculated as follows:

$$\frac{(Sum\ of\ CPI\ adjusted\ sales\ in\ the\ size\ segment)_t}{(Total\ CPI\ sales)_t} \qquad \forall t \in \{2000, 2001, \dots, 2020\}$$

Revealed by Figure 10, in 2020, the market share for *very small* firms was 14.8%, while *small* firms had 29.2%, *medium-sized* firms 31%, and *large firms* 24.9%. Compared to 2010, we notice a notable change in the market dynamics. In 2010 *very small firms* had a total market share of 19.3%, *small* firms had 26.8%, *medium-sized* firms 27.2%, and *large* firms 26.8%. *Very small* and *large* firms constituted more of the market in 2010 than in 2020, while *small* and *medium-sized* firms have expanded their shares of the market.

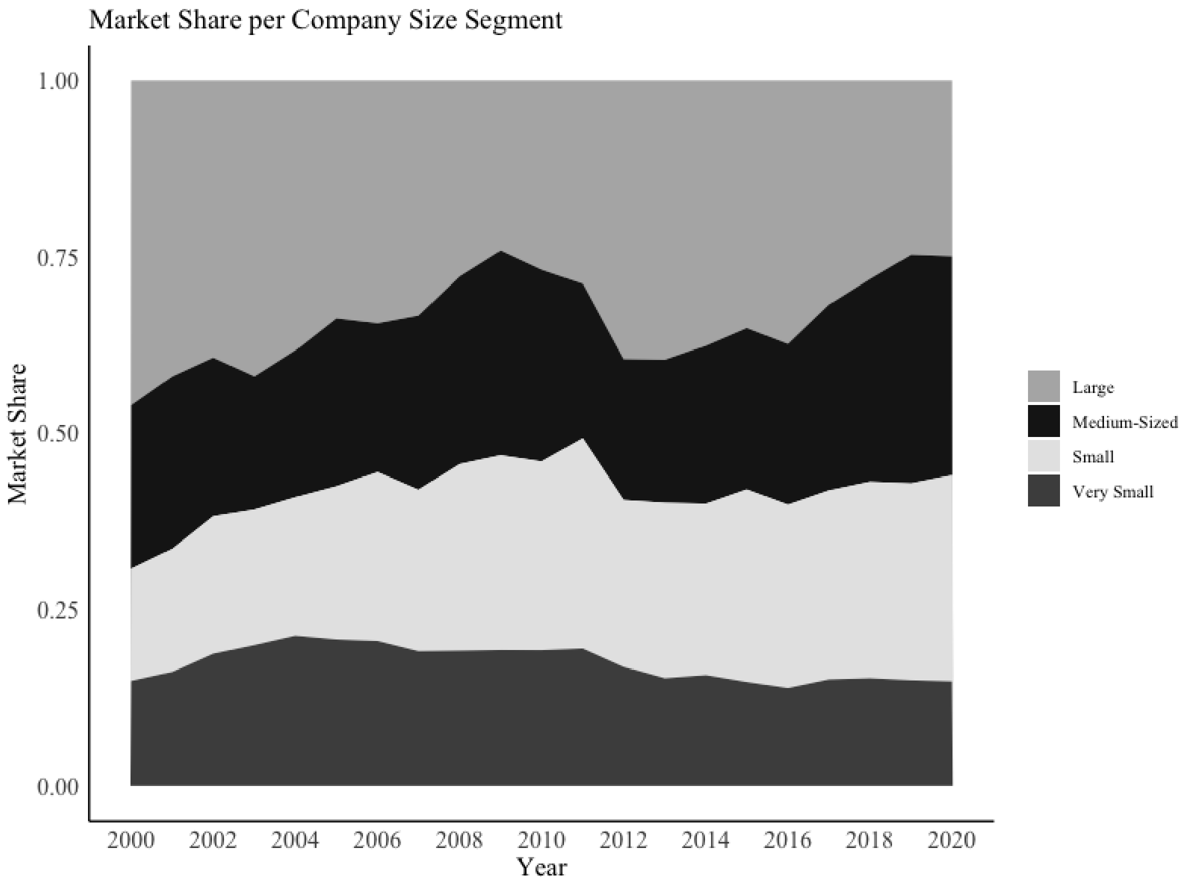


Figure 10. Market Share per Company Size Segment *Data source: Regnskapsdatabasen*

Going back to 2000, *very small* firms had a 14.9% market share, *small* firms 15.9%, *medium-sized* firms 23.8%, and *large* firms a substantial 46.0%. The company size segments have shown relatively consistent trends over time, with notable shifts such as *large* firms losing market share while *small* and *medium-sized* firms gaining.

4.4 Brief Summary

In this chapter, we explored the recent trends in the software development sector and market, both characterized by high growth, in terms of increased total sales, increased number of companies, and a rise in overall employment. Despite the sector's growth, there's a trend towards smaller units, with new and smaller companies contributing to lower average sales and fewer employees per company.

Profitability, measured by the weighted average EBITDA margin, has maintained a steady range of 6-8% from 2003 to 2020. The market also exhibits higher productivity over time, measured by average sales per employee, alongside an increase in average R&D spending per company.

We observe an increased market fragmentation, with a surge in new companies diluting the market shares of the largest players. This is not a result of larger companies declining but rather the overall market expanding around them.

Building on the insights from Chapter 4, which detailed the evolving trends in Norway's software development sector, Chapter 5 aims to extend this understanding through a statistical lens. Specifically, we will explore how existing market position, EBITDA margins, R&D expenditures and productivity correlate with market share dynamics and the likelihood of market share growth. While Chapter 4 provided a comprehensive descriptive overview, Chapter 5 delves into the underlying mechanisms that potentially could drive these trends, employing regression analysis to discern patterns and relationships within the data.

Chapter 5 - Regression Results

5.1 Brief Overview

In this chapter we are moving on from the descriptive nature of Chapter 4, to explore the characteristics of market share dynamics and likelihood of market share growth through statistical analyses. We employ two primary regression models for each company size segment: (1) a linear model that examines current market share, and (2) a logistic model that investigates growth or no growth in market share. First, we consider baseline linear regression models, meaning we do not fully account for the panel data structure of the datasets. Subsequently, we adjust for fixed effects, and lastly, we run the logistic regression models. To account for violations of model assumptions, the coefficients in the fixed effects linear regression and the logistic regression models, are adjusted for robust standard errors. All relevant metrics in the data are CPI adjusted based on numbers from SSB (SSB, 2023).

The chapter is structured in a way that first provides an overview of the variable selection and descriptive statistics of the data used in the regression models. Following is a presentation of the results obtained from the linear models, followed by the results obtained from the logistic models. Considerable importance is attributed to addressing issues related to autocorrelation and heteroskedasticity for the transparency and soundness of the research findings. Sensitivity analyses are conducted to further strengthen the reliability and validity of the results.

5.2 Data Summary

5.2.1 Variable Selection

When selecting variables for the regression models, we carefully evaluated how they might relate to each other. By addressing possible correlations between explanatory variables prior to model creation, we avoid any issues regarding multicollinearity, which is a necessity for the validity of both the linear and logistic regression models.

In the linear regression, we use the natural logarithm of market share as the dependent variable, to capture potential non-linear relationships and offer a deeper insight into market dynamics. The market share per company is calculated by dividing company sales by the total yearly market sales. Our logistic regression models employ a binary variable as the response to distinguish firms based on whether they experience growth in market share or not. This method highlights key differences between growing and non-growing firms in terms of market share. Focusing on market share, rather than sales, allows us to analyze relative market dominance more efficiently. The analysis encompasses both position in the market and likelihood of growth, aiming to identify factors linked to a firm's market share in different segments and characteristics associated with market share growth. This will help us answer Research Question 2.

We focus on the following key explanatory variables to analyze market dynamics: research and development (R&D) expenditure relative to sales, the natural logarithm of sales per employee (for the linear regression only), EBITDA margin, and the natural logarithm for current market share (for the logistic regression only). R&D spend ratio provides insights into innovation levels across segments, reflecting how firms allocate resources to research and development relative to their sales. The logarithm of sales per employee addresses productivity's non-linear link with the response variable, helping us analyze variations in productivity across firm sizes (Appendix A.3.1). EBITDA margin is crucial to assess any correlation between profitability and firm size or growth likelihood. Lastly, the logarithm of current market share is vital in the logistic regression models for understanding the relationship between a firm's size and its likelihood for market share growth. These variables collectively aim to offer a comprehensive view of the characteristics within the Norwegian software development sector.

To enhance the robustness and accuracy of our models, we also include several explanatory variables that, while not central to our primary analysis, are valuable for controlling for certain circumstances. These variables encompass "Year" as a categorical variable to account for time-based market trends and fluctuations, and a binary variable accounting for whether a firm is in Oslo or not, to control for metropolitan effects on performance and behavior. As mentioned by (Golding, 2023), IT companies tend to cluster in metropolitan areas. We also consider the ratio of cash holdings to sales, a potential indicator of a firm's financial health and growth potential. The age of the company is accounted for to assess how its market position and growth prospects are influenced by its maturity and industry experience. Lastly, ownership concentration as measured by the Herfindahl Index is included to account for its possible impact on strategic decisions and market performance (Mjøs & Selle, 2022, p. 10). By integrating these variables, we construct a more comprehensive and nuanced framework.

For a detailed elaboration of each variable, how it is calculated, and its abbreviation, please see Appendix A.1.

5.2.2 Descriptive Statistics

As outlined in Section 3.5.1, the secondary data from Regnskapsdatabasen, encompassing extensive financial data for Norwegian companies spanning from 2010 to 2020, has been methodically divided into four subsets, segmenting companies based on size: *very small*, *small*, *medium-sized*, and *large* firms. This stratification is useful for capturing the distinctive characteristics and dynamics prevalent within each company size segment of the sector. Due to regression models' sensitivity to extreme values, outlier removal for all subsets is done as described in Section 3.5.3**,** before further analysis. Note that the slight discrepancies in the explanatory variables for the linear and logistic regression models result in minor deviations in the company size segments after outlier removal.

In this section, we delve into the descriptive statistics of the key variables described in Section 5.2.1. This analysis will identify the distribution patterns and central tendencies for every variable within each company size segment. Additionally, notable correlations among variables will be discussed, with a focus on their predictive power. For comprehensive correlation matrices please see Appendix A.2.

5.2.2.1 Numeric Variables

Examining the distributions of each numeric variable gives a solid understanding for how the data behaves, and reveals interesting differences between the company size segments.

Very Small Firms

Examining Table 1, we notice a standard deviation of 1.130 for the logarithm of market share, along with a large interval between minimum and maximum values, which indicates that for *very small* firms, there is a significant dispersion in the distribution for the response variable. A moderate deviation between the mean and the median suggests a distribution that is highly symmetrical around the central tendency.

Statistic	Max	Mean	Median	Min	N	Pctl(25)	Pctl(75)	St. Dev.
Logarithm of Market Share	-8.371	-10.510	-10.458	-14.129	17,080	-11.140	-9.724	1.130
EBITDA Margin	0.644	0.086	0.097	-2.829	17,080	0.004	0.256	0.335
Cash per Sale	5.404	0.428	0.251	0.004	17,080	0.103	0.501	0.591
R&D Per Sale	1.087	0.016	0.000	0.000	17,080	0.000	0.000	0.093
Log of Sales per Employee	8.345	6.824	6.984	3.447	17,080	6.420	7.417	0.849
Age of the Company	27	8.686	7	2	17,080	4	12	6.032
Ownership Concentration	1.000	0.766	1.000	0.188	17,080	0.500	1.000	0.281

Table 1. Desriptive Statistics of Numeric Variables for Very Small Firms *Data source: Regnskapsdatabasen*

The median for the EBITDA margin is slightly higher than the mean value of 0.086, indicating a slight positive skewness in profitability. Cash per sale shows considerable variation, with a median value lower than the mean. The maximum value of 5.404 suggests a distribution characterized by a larger right tail. A median value of 0 indicates that only a few companies invest in R&D. However, there is a skewed distribution towards larger values, with a maximum expenditure of 1.087. The productivity parameter, logarithm of sales per employee, shows quite a bit of variation with a standard deviation of 0.849, and a slightly left tailed distribution. The standard deviation of 6.032 for the age of the company indicates that there is significant variation in the age distribution of the enterprises. However, the median age reflects that most firms are rather young. A median value of 1 indicates that most organizations exhibit a complete ownership concentration, indicating only a few owners as measured by the Herfindahl Index (Mjøs & Selle, 2022, p. 10).

Small Firms

In *small* enterprises, Table 2 shows that the distribution of the logarithm of market share is narrower than that of *very small* firms, with a standard deviation of 0.684 compared to 1.130. The similarities between the mean and median values suggest a distribution that exhibits a rather balanced nature.

Statistic	Max	Mean	Median	Min	N	Pctl(25)	Pctl(75)	St. Dev.
Logarithm of Market Share	-6.626	-7.929	-7.888	-10.388	3,695	-8.381	-7.410	0.684
EBITDA Margin	0.399	0.060	0.085	-2.312	3,695	0.025	0.158	0.229
Cash per Sale	2.132	0.215	0.161	0.009	3,695	0.068	0.283	0.228
R&D Per Sale	1.857	0.068	0.000	0.000	3,695	0.000	0.000	0.211
Log of Sales per Employee	8.186	7.223	7.268	4.921	3,695	6.983	7.541	0.482
Age of the Company	31	11.792	11	2	3,695	6	16	7.052
Ownership Concentration	1.000	0.550	0.460	0.080	3,695	0.241	1.000	0.340

Table 2. Descriptive Statistics of Numeric Variables for Small Firms *Data source: Regnskapsdatabasen*

There is a similar, though smaller, spread in the EBITDA margin compared to *very small* firms. With a mean of 0.060 and a median at 0.085, we see a reasonable level of profitability. Cash per sale shows a more moderate right tail than for *very small* firms, albeit a rather unbalanced distribution. R&D per sale appears to be rare also for *small* firms, with a median of zero. The logarithm of sales per employee appears to have a left tail and a moderate amount of variation. A median age of 11 years suggests the presence of relatively established businesses, but the standard deviations suggest a wide range of ages. The ownership concentrations tend to be lower in comparison to that of *very small* enterprises, with a median value of 0.467.

Medium-Sized Firms

Table 3 reveals that *medium-sized* firms have a mean logarithm of market share of -6.118, with a median that closely approximates the mean. With a relatively small standard deviation, this implies a narrower dispersion of market shares around the central market position.

Statistic	Max	Mean	Median	Min	N	Pctl(25)	Pctl(75)	St. Dev.
Logarithm of Market Share	-5.009	-6.118	-6.125	-7.441	541	-6.501	-5.743	0.534
EBITDA Margin	0.246	0.085	0.085	-0.188	541	0.044	0.128	0.068
Cash per Sale	0.749	0.141	0.105	0.0005	541	0.033	0.199	0.130
R&D Per Sale	0.393	0.018	0.000	0.000	541	0.000	0.005	0.049
Log of Sales per Employee	8.266	7.438	7.411	6.588	541	7.238	7.611	0.298
Age of the Company	38	15.168	14	2	541	9	20	8.126
Ownership Concentration	1.000	0.763	1.000	0.077	541	0.479	1.000	0.345

Table 3. Descriptive Statistics of Numeric Variables for Medium-Sized Firms *Data source: Regnskapsdatabasen*

The EBITDA margin's median and mean values are both 0.085, suggesting an even level of profitability within the subset. The cash per sale and R&D per sale demonstrate minimal variability and asymmetry. The logarithm of sales per employee has a similar mean and median value and appears to have a rather balanced distribution. The distribution for the company age has a wider range in comparison to smaller firms, with the median age being 14 years and a standard deviation of 8.126. A median value of 1 for ownership concentration indicates most *medium-sized* firms have few owners.

Large Firms

Large firms have the largest gap between the mean logarithm of market share and the median, as evident from Table 4. However, the distribution appears generally symmetrical around the central tendency.

Statistic	Max	Mean	Median	Min	N	Pctl(25)	Pctl(75)	St. Dev.
Logarithm of Market Share	-2.942	-4.023	-3.765	-5.712	98	-4.537	-3.627	0.602
EBITDA Margin	0.188	0.073	0.064	-0.027	98	0.043	0.109	0.048
Cash per Sale	0.332	0.107	0.086	0.000	98	0.021	0.191	0.095
R&D Per Sale	0.018	0.001	0.000	0.000	98	0.000	0.000	0.003
Log of Sales per Employee	8.325	7.582	7.563	5.974	98	7.357	7.775	0.324
Age of the Company	80	26.786	20.5	3	98	13	33	19.072
Ownership Concentration	1.000	0.952	1.000	0.112	98	1.000	1.000	0.194

Table 4. Descriptive Statistics of Numeric Variables for Large Firms *Data source: Regnskapsdatabasen*

As for the other subsets, most *large* firms are profitable, with a mean EBITDA margin of 0.073 and a median of 0.064. Notably, there appears to be less variation in profitability compared to the other subsets. As for the other subsets, there is a right tail in the distribution of cash per sale and a somewhat lower ratio with a mean of only 0.107. The right tail is also present in the R&D per sale distribution. The distribution of the logarithm of sales per employee appears to be balanced. This sector includes both older and younger organizations, with a median age of 20.5 years and a standard deviation of about 19. Near-complete ownership concentration, with the median at 1 and mean of 0.952, indicates significant control by owners or primary shareholders in most *large* firms.

5.2.2.2 Categorical Variables

There exists a disparity in the distribution of enterprises in Oslo and their corresponding levels of market growth, which is contingent upon the size of the firms. Examining Table 5, the city of Oslo exhibits a notable concentration of *large* firms, suggesting a tendency of large companies being located in the capital.

	Very Small	Small	Medium Sized	Large
Percentage per subset with HQ in Oslo	27.9%	37.7%	55.4%	64.3%
Percentage per subset with Market Share Growth	41.8%	52.0%	45.8%	49.0%

Table 5. Descriptive Statistics of Categorical Variables *Data source: Regnskapsdatabasen*

Examination of the number of unique companies per year per company size segment indicates a consistent upward trajectory, as seen in Table 6. In recent years, notable increases have been observed in *very small, small,* and *medium-sized* enterprises, suggesting a potential alteration in market dynamics also evident by Figure 10 in Section 4.3.

Size	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Large	7	8	8	7	8	9	10	10	10	11	10
Medium	33	34	41	40	44	50	49	56	60	67	67
Small	225	228	245	282	298	380	403	366	407	422	439
Very Small	1,067	1,121	1,187	1,213	1,405	1,658	1,684	1,773	1,846	1,984	2,142

Table 6. Number of Companies per Company Size Segment *Data source: Regnskapsdatabasen*

5.2.3 Pearson's Correlation Coefficients

To get a comprehensive understanding of the interrelationships in the data we analyze the Pearson correlation coefficients. Pearson's correlation coefficients reveal the degree of linear relationship between two variables, where values closer to -1 or +1 indicate a strong negative or positive linear relationship, respectively, and values around 0 suggest no linear correlation (Hahs-Vaughn, 2023). Interpreting these coefficients, we uncover a nuanced landscape of predictors for current market position (measured by the logarithm of market share) and market share growth within each company size segment. For complete correlation matrices, please see Appendix A.2.

Current market position

The productivity parameter, logarithm of sales per employee, has a consistently strong positive correlation with current market position across all firm size segments, with its influence being most pronounced in *very small* firms (0.77) and gradually lessening in larger firms. This could indicate that productivity is more profoundly related to higher market share in smaller firms.

In *very small* firms, cash holdings per sale and ownership concentration emerge as additional predictors for the current market position, but with negative correlations (-0.253 and -0.279, respectively). This suggests that in this size segment, higher cash reserves or concentrated ownership may not necessarily drive market share. In contrast, for *small* firms, EBITDA margin and R&D per sale show mild positive and negative correlations (0.196 and -0.178), indicating a complex interplay between profitability, R&D investment, and market share.

Medium-sized firms present a different dynamic, with being located in Oslo and ownership concentration showing positive correlations (0.164 and 0.3) with the current market position, hinting at geographical and ownership influences. *Large* firms display a distinct pattern, with EBITDA margin and company age as notable predictors, showcasing a negative correlation (-0.347) and a positive correlation (0.361) with the logarithm of market share, respectively, suggesting that for *large* firms, factors like long-term establishment and operational efficiency are more influential.

A consistent negative correlation between year and the current market position across all company size segments indicates a broader trend, possibly due to evolving market dynamics or increased competition as revealed and discussed in Chapter 4.

Growth in market share

Contrary to the current market position, the Pearson correlation coefficients reveals no uniform best predictor for market share growth across company size segments. For *very small* firms, the logarithm of market share and age of the company are most important, with mixed correlations (0.128 and -0.133). In *small* firms, company age emerges as a significant negative predictor (-0.206). *Medium-sized* firms show EBITDA margin and company age as mixed predictors (0.079 and -0.105), while for *large* firms, cash per sale and location in Oslo are positively correlated (0.163 and 0.267).

5.3 Baseline Linear Regression

Moving on from the variable selection and correlation analysis, we will now analyze the company size segments utilizing linear regression. The subsequent sections provide an overview of the outcomes obtained from the baseline linear regression models. The findings presented below offer an initial comprehension of the variables that could potentially impact market share, establishing a foundation for subsequent examination incorporating fixed effects. We will present the coefficients of each explanatory variable but computing their marginal effects on the response variable is delayed until Section 5.4.2, *after* adjusting for fixed effects and robust standard errors. It should be noted that the findings for each variable assume all other variables are kept constant. For the detailed regression table, please see Table 7.

5.3.1 Interpreting the Baseline Linear Models' Coefficients

In the context of the baseline linear regression model, as discussed in Section 5.2.1, we examine which explanatory variables significantly correlated with market share (Section 5.2.3). Our primary focus is on the relationship between market share and the following variables: EBITDA margin, R&D expenditure relative to sales, and the logarithm of sales per employee. Although we will also assess the coefficients of other explanatory variables, these are primarily treated as control variables. In our analysis, we have included year dummies, but we omit their coefficients from Table 7, as they do not provide additional insight beyond the negative time trend already established in Section 5.2.3. This decision is supported by a desire to concentrate on the main variables, while acknowledging the importance of controlling for time-related effects in the model.

EBITDA Margin

Controlling for the other variables, we find in Table 7 a negative relationship between EBITDA margin and logarithm of market share ($\beta_{\text{very small}}$ = -0.079, $p < .01$) among *very small* firms, suggesting that there might be a compromise between profitability and market share in this company size segment. In contrast, there exists no significant positive relationship between the logarithm of market share and EBITDA margin among *small* firms (β_{small} = -0.036, $p > .1$), but the sign of the coefficient is similar to that of *very small* firms. For *medium-sized* firms, there is a barely statistically significant positive relationship between EBITDA margin and the response variable ($\beta_{\text{medium-sized}}$ = 0.438, $p < .1$). There is, however, a highly significant negative correlation between EBITA margin and logarithm of market share for *large* firms (β_{large} = -

3.556, $p < .01$), suggesting a reduced profitability for companies with larger market shares within this company size segment.

R&D per Sale

For the same regression models, we derive the relationship of the market share with R&D expenditures per sale from Table 7. For *very small* firms, increased R&D per Sale ($\beta_{\text{very small}}$ = 0.565, $p < .01$) is significantly associated with a higher logarithm of market share. This might suggest that companies with higher market shares typically have higher R&D expenses relative to sales. It should be noted that R&D per sale is not statistically significant for *small* firms (β_{small} = 0.036, $p > .1$), *medium-sized* firms ($\beta_{\text{medium-sized}}$ = 0.431, $p > .1$), or *large* firms (β_{large} = -15.769, $p > .1$). Although only *very small* firms have a statistically significant relationship with R&D per sale and the logarithm of market share, the sign of the coefficients are similar for all firm-size categories except for *large* firms, which has a large and negative coefficient.

Log of Sales per Employee

We also investigate, using the results in Table 7, the correlation between market share and productivity, as measured by sales per employee. The relationship between the logarithm of market share and the logarithm of sales per employee is positive and statistically significant ($p < .01$) across all company size segments ($\beta_{\text{very small}}$ = 1.005; β_{small} = 1.040; $\beta_{\text{medium-sized}}$ = 0.971; β_{large} = 0.548). The sign of the coefficients indicates that increased productivity is a common trait for companies with high market shares.

Cash per Sale

Next, we use Table 7 to discuss how the market share is related to cash holdings per sale. Cash per sale has an adverse relationship with the logarithm of market share for *very small* ($\beta_{\text{very small}}$ = -0.178, $p < .01$), *small* (β_{small} = -0.107, $p < .01$) and *medium-sized* firms ($\beta_{\text{medium-sized}}$ = -0.594, $p < .01$). For *large* firms, there is no statistically significant relationship (β_{large} = 0.570, $p > .1$). However, we note that while the coefficients for *very small, small,* and *medium-sized* firms have negative signs, the sign for *large* firms is positive, suggesting a potential but uncertain difference between this particular company size segment and the rest.

Company is in Oslo

As revealed in Section 5.2.2, a substantial amount of the companies within each company size segment is clustered in Oslo, supporting the discussion in Section 5.2.1 regarding how companies tend to cluster around a country's main city. Table 7 reveals that for *medium-sized* firms, there is a positive and significant link between a company being in Oslo and the logarithm of market share ($\beta_{\text{medium-sized}} = 0.194$, $p < .01$). There is, however, no significant link ($p > .1$) between being located in the Norwegian capital and market shares for any of the other company size segments ($\beta_{\text{very small}} = -0.002$; $\beta_{\text{small}} = -0.001$; $\beta_{\text{large}} = 0.076$).

Age of the Company

We also discuss the relationship between the age of the company and market share based on the regression output in Table 7. For all company size segments, the age of the company has a statistically significant, positive correlation with the response variable ($\beta_{\text{very small}} = 0.007$, $p < .01$; $\beta_{\text{small}} = 0.004$, $p < .01$; $\beta_{\text{medium-sized}} = 0.004$, $p < .05$; $\beta_{\text{large}} = 0.008$, $p < .01$). This implies that older firms in each company size segment typically have higher market shares, which is not surprising as they will have had more time to consolidate. The significance of the coefficients underscores the importance of controlling for company age.

Ownership Concentration

Lastly, we discuss how market share is correlated with the ownership structure. The relationship between ownership concentration and the response variable varies between company size segments, as evident in Table 7. For *very small* firms, there is a significant inverse relationship ($\beta_{\text{very small}} = -0.946$, $p < .01$), but for *medium-sized* firms the relationship is positive ($\beta_{\text{medium-sized}} = 0.364$, $p < .01$). For *small* ($\beta_{\text{small}} = 0.020$, $p > .1$) and *large* ($\beta_{\text{large}} = -0.011$, $p > .1$) firms, there are no significant relationships. This suggests that even though there are some relationships between ownership concentration and market share in certain segments, there is no holistic interpretable correlation. The sign of the coefficients also varies between the company size segments, with negative signs for *very small* and *large* firms, and positive for *small* and *medium-sized* firms.

Constant Terms

The constant terms in our linear regression models reveal significant baseline differences in market share across various company size segments. These constants, which represent the baseline logarithm of market share when all explanatory variables are at zero, vary notably between *very small, small, medium-sized,* and *large* firms. The less negative constants for larger firms indicate a higher baseline level of market share, while the more negative constants in smaller firms highlight a lower starting point in market share. These differences in constants are crucial as they point to the inherent market dynamics and baseline characteristics unique to each company size segment.

Baseline Linear Regression Models

	DV: Logarithm of Market Share			
	Very Small	Small	Medium-Sized	Large
EBITDA Margin	-0.079***	-0.036	0.438*	-3.556***
	(0.015)	(0.031)	(0.246)	(1.313)
R&D per Sale	0.565***	0.036	0.431	-15.769
	(0.051)	(0.034)	(0.336)	(18.123)
Log of Sales per Employee	1.005***	1.040***	0.971***	0.548***
	(0.006)	(0.015)	(0.055)	(0.195)
Cash per Sale	-0.178***	-0.107***	-0.594***	0.570
	(0.008)	(0.031)	(0.130)	(0.632)
Company is in Oslo	-0.002	-0.001	0.194***	0.076
	(0.011)	(0.014)	(0.033)	(0.127)
Age of the Company	0.007***	0.004***	0.004**	0.008***
	(0.001)	(0.001)	(0.002)	(0.003)
Ownership Consentration	-0.946***	0.020	0.364***	-0.011
	(0.017)	(0.021)	(0.047)	(0.291)
Constant	-16.153***	-15.081***	-13.298***	-8.069***
	(0.049)	(0.116)	(0.404)	(1.447)
Observations	17,080	3,695	541	98
Adjusted R^2	0.707	0.633	0.530	0.306
Residual Std. Error	0.611 (df = 17062)	0.414 (df = 3677)	0.366 (df = 523)	0.502 (df = 80)
F Statistic	2,426.654*** (df = 17; 17062)	375.234*** (df = 17; 3677)	36.758*** (df = 17; 523)	3.514*** (df = 17; 80)
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Yearly effects accounted for all company size segments

Table 7. Baseline Linear Regression Output *Data source: Regnskapsdatabasen*

5.3.2 Baseline Linear Models' Performance

In our examination of the models' performances as presented in Table 7, we note a range in the explanatory power, adjusted for the number of predictors (reflected by the adjusted R^2), from 30.6% for *large* firms to 70.7% for *very small* firms. This variation highlights that our models can explain a significant portion of the variance in the logarithm of market share, with the adjusted R^2 providing a more nuanced understanding of the models' effectiveness considering the number of predictors used.

In Table 7 we also observe the significance of the F-statistic across all models, indicating that our models collectively possess a meaningful predictive power for market share, surpassing that of a simplistic constant model. As revealed in Section 5.2.2**,** there is a relatively narrow span in the response variable across each company size category, which in turn makes the size of the residual standard errors more pronounced. While these errors suggest that the models' predictions are not extremely precise, it is important to contextualize this within the primary objective of our analysis: to explore the relationship between specific independent variables and market share. In this light, the lesser precision of predictions does not critically undermine our analysis.

However, it's essential to approach the output of Table 7 with a degree of caution. The results should be viewed as an initial indication of the models' explanatory capacity rather than a definitive demonstration. This caution is particularly warranted given the unaccounted nuances in the panel data structure of our dataset. Such considerations are crucial for accurately interpreting and applying the findings of our analysis and will be assessed in Section 5.4.

5.3.3 Evaluating the Baseline Linear Models

We will now evaluate the underlying assumptions for linear regression models to assess their validity. Our examination includes: (1) checking for linearity between the predictors and the response variable, (2) verifying that the residuals are normally distributed, and (3) ensuring that the residuals are independent of each other. As discussed in Section 5.2.1, through careful selection of independent variables, we have mitigated the risk of multicollinearity in the models. In Section 5.4, we account for heteroskedasticity by adjusting for robust standard errors, and therefore, we deem it unnecessary to further evaluate the assumption of homoskedasticity.

Linearity

Continuing our evaluation of the linear regression models' assumptions as outlined in the introduction, the first aspect we examine is the linearity between predictors and the response variable. Our analysis revealed no substantial non-linear relationships. Particularly, the logarithm of sales per employee displays a clear linear relationship with the response variable, which underscores its significant explanatory power. For more detailed visualizations of this linearity, please see the linearity plots in Appendix A.3.

Normality

Moving on from the linearity assessment, the next critical assumption we address is the distribution of residuals. To determine whether the residuals are normally distributed, we utilize the Shapiro-Wilk (SW) test. This test is regarded as one of the best for assessing normality, even in small samples (Malato, 2023). For methodological consistency, we apply this test for all company size segments. The W-statistic from the SW test ranges between 0 and 1, where values closer to 1 indicate a higher likelihood of the data being normally distributed. The null hypothesis of the SW test posits that the data is normally distributed, and a low p-value leads to rejection of this hypothesis, implying non-normal distribution. For *very small* ($W = 0.967$, $p < .01$), *small* ($W = 0.956$, $p < .01$), and *medium-sized* ($W = 0.983$, $p < .01$) firms, we reject the null hypothesis, suggesting that the residuals are not normally distributed. However, the W-statistics are relatively close to 1, indicating that the deviations are moderate. For *large* firms, the null hypothesis is not rejected ($W = 0.984$, $p > .1$), but it is noteworthy that this company size segment contains substantially fewer observations than the others.

Given the Shapiro-Wilk test's sensitivity to small deviations from normality, we combine this statistical test with graphical interpretations, for a more nuanced evaluation. In the QQ-plots in Table 11, the straight line represents perfect normality. Deviations from this line in the models fitted to *very small* and *small* company size segments suggest that the residuals are not perfectly normally distributed, with the most notable deviations occurring in the "tails". This indicates that, while there are some deviations from normality, the assumption is not entirely violated. For models fitted to *medium-sized* and *large* firms, deviations from normality are less pronounced.

Our assessment of the normality assumption, using both Shapiro-Wilk tests and QQ-plots, reveals moderate deviations from normal distribution for *very small*, *small*, and *medium-sized* firms. Although these deviations are present, they are not so substantial as to significantly undermine the statistical inference capabilities of our models. These findings suggest that while some caution is warranted in interpreting results, the deviations from normality do not pose dramatic concerns for our analysis. This balanced view, considering both statistical tests and graphical interpretations, allows us to proceed with a reasonable level of confidence in our models' overall validity.

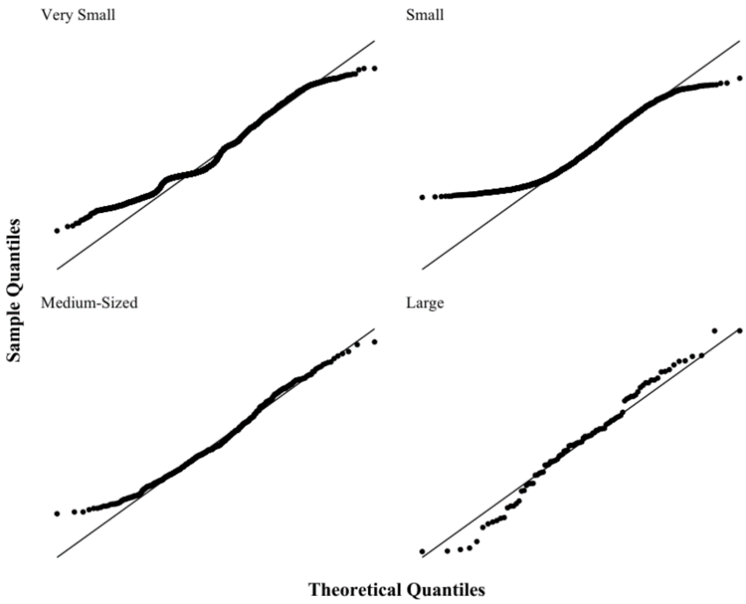


Figure 11. QQ-plot of Residuals *Data source: Regnskapsdatabasen*

Independent Residuals

Building on our balanced evaluation of the normality assumption, we next turn to the independence of residuals, a critical aspect of model validity. To assess this, we employ the Durbin-Watson (DW) test, a widely used method for detecting autocorrelation in the residuals. The DW-statistic ranges from 0-4, were values close to 2 indicates no autocorrelation in the residuals, and values close to 0 or 4 indicates positive or negative autocorrelation, respectively (Krämer, 2014, pp. 408-409). Under the null-hypothesis, there is no autocorrelation, i.e. independent residuals. As evident in Table 8, we reject the null-hypothesis for all company size segments. The significant positive autocorrelation indicated by the Durbin-Watson test reveals that none of the models fully satisfy the assumption of independent residuals. It is important to note, however, that the Durbin-Watson test does not accommodate the panel data structure of the datasets, which places some

Model	DW	p-value
Very Small Firms	0.741	<.01
Small Firms	0.794	<.01
Medium-Sized Firms	0.738	<.01
Large Firms	0.498	<.01

Table 8. Durbin-Watson Statistic for Linear Regression
Data source: Regnskapsdatabasen

constraints on the direct interpretation of the results. Bearing these limitations in mind, the test still serves as a useful indicator of potential autocorrelation issues in our models.

5.4 Linear Regression with Fixed Effects and Robust Standard Errors

In the baseline linear regression models, we found some violations with the models' assumptions, specifically regarding the assumptions of independent residuals. To address these findings, we adjust for both fixed effects and robust standard errors.

Considering the panel data structure of the company size segments, incorporating fixed effects can help mitigate the effect of some of the complexities not accounted for in the baseline models. Fixed effects adjust for unobserved variables that vary across firms and are constant, thereby controlling for potential omitted variable bias (Huntington-Klein, 2023). This approach is particularly beneficial in longitudinal data where inherent characteristics of entities may influence the observed outcomes.

Following our consideration of fixed effects to control for unobserved, firm-specific variables in the panel data, we turned our attention to the temporal structure of the data. Utilizing the Durbin-Watson test, a crucial step in the analysis of longitudinal data, we uncovered significant autocorrelation in the linear regression model. Autocorrelation, or the correlation of a variable with itself across time, is a common challenge in longitudinal studies. Its presence can lead to overestimated standard errors, which in turn may inflate the significance levels of the model's coefficients (Krämer, 2014, pp. 408-409). This situation can result in erroneously interpreting non-significant findings as significant.

To address this issue and enhance the validity of our findings, we have decided to adjust for robust standard errors. This adjustment is vital for obtaining the most valid interpretation of the relationships between variables. Robust standard errors are less sensitive to violations of assumptions such as homoskedasticity and, thereby providing a firmer ground for statistical inference in the presence of such complexities (Taboga, 2021). By implementing these adjustments, we aim to ensure that our conclusions not only are statistically sound but also meaningful in the context of the real-world phenomena they seek to explain.

5.4.1 Adjusting the Linear Models

To incorporate fixed effects in our linear regression models for panel data, we utilized the *plm()* function in R, a widely-used approach for such data (Croissant & Millo, n.d.). For *very small* and *small* firms, we applied a *"within, two-ways"* transformation, addressing both time-specific and firm-specific effects. Conversely, for *medium-sized* and *large* firms, we used a *"within, individual"* transformation, as these segments contain significantly fewer unique companies. This approach was chosen to control for unobserved factors that are constant over time for each firm but may vary across firms. Yearly dummies are included in the firm-specific fixed effects models, but excluded from the regression output following the same logic discussed in Section 5.2.1. To account for potential heteroskedasticity, we adjusted for robust standard errors using White's method. This adjustment provides more reliable standard error estimates in the presence of heteroskedasticity, which can affect the confidence intervals and statistical significance of certain coefficients (Zach, 2022). By adjusting for White's standard errors, we can interpret the coefficients without relying on the assumption of homoskedasticity.

Considering the implementation of robust standard errors, the analysis for *medium-sized* and *large* firms employs the heteroskedasticity-consistent (HC) estimator HC3. This choice is motivated by HC3's enhanced adjustment for high-leverage observations, which becomes particularly relevant in datasets with a smaller number of observations, where individual data points can have a more pronounced impact (Pinzon, 2022). In contrast, for the analysis involving *very small* and *small* firms, which comprise a larger number of observations, the HC1 estimator is used. HC1 offers a correction for heteroskedasticity with a degrees of freedom adjustment, making it suitable for relatively larger samples where the influence of individual observations is less significant compared to smaller datasets.

5.4.2 Interpreting the Adjusted Linear Models' Coefficients

Naturally, the regression outputs undergo notable changes when we adjust for fixed effects. For *very small* and *small* firms, our analysis now focuses on variations both firm- and time-specific variation, utilizing a *"twoways"* fixed effects approach. This method allows us to capture the nuances of firm-specific characteristics and temporal changes. Conversely, for *medium-sized* and *large* firms, we concentrate on differences among individual firms, implementing an *"individual"* fixed effects model. This approach is particularly suitable given the smaller number of observations in these categories, where each firm's unique characteristics play a more significant role. The incorporation of fixed effects, along with robust standard errors, substantially influences the linear regression models' outcomes, offering a more nuanced understanding of the dynamics at play in the different company size segments.

Table 9 displays the results for the fixed effects linear regression model with robust standard errors relating the market share to the independent variables such as profitability, productivity, and innovative activities. In the analysis of regression outputs, it's crucial to contextualize the coefficients differently between the company size segments. For *very small* and *small* firms, coefficients should be interpreted considering both time- and firm-specific variations. This is due to our use of a *"twoways"* fixed effects approach, which is apt for these company size segments as it allows us to account for both temporal changes, such as economic cycles or industry trends, and unique characteristics of individual firms. This comprehensive method aids in understanding the diverse factors relating to market share in these company size segments.

On the other hand, for *medium-sized* and *large* firms, our analysis employs an *"individual"* fixed effects approach. This method is chosen due to the smaller number of unique companies in these categories, where firm-specific attributes are more pronounced and pivotal. In these segments, each company's distinct characteristics, rather than time-specific factors, play a more critical role in affecting market share. This approach reflects the need for focused analysis on firm-specific dynamics, especially in contexts with a lower degree of data aggregation, as is the case with *medium-sized* and *large* firms.

Examining Table 9, we observe that after implementing fixed effects and adjusting for robust standard errors, several coefficients that were significant in the baseline models are no longer deemed so, and vice versa. For example, in the baseline model for *very small* firms presented in Table 7, R&D expenses in relation to sales had a significant positive correlation with the response variable, a relationship that no longer holds statistical significance. Additionally, the control variable for Oslo and company age are left out of the analysis as these are assumed to be constant over time, and thereby not attributing any variation on firm-specific level.

We now turn our attention to investigate the coefficients of each explanatory variable in our model, focusing particularly on their marginal effects on the response variable for the coefficients that are significantly different from zero. In our linear models, where the response variable is log-transformed, the coefficients indicate the expected percentage change in the response variable for a one-unit increase in the explanatory variable, assuming all other variables remain constant. Though clearly illustrated in Table 9, it's important to note that the constant term is absorbed in each model, a direct result of implementing fixed effects. Our analysis of the marginal effects is specifically conducted for variables that show statistical significance. However, this interpretation conceptually extends to the coefficients of the insignificant variables as well, though without the level of certainty attributed to the significant ones.

EBITDA Margin

We start by analyzing the profitability parameter in Table 9, indicated by EBITDA margin, and find a significant relationship with the logarithm of market share for *very small* ($\beta_{\text{very small}}$ = 0.162, $p < .01$) and *small* (β_{small} = 0.090, $p < .1$) firms. For these segments, a one percentage point (one hundredth of a unit) increase in EBITDA margin is associated with an approximate increase in market share of 0.2% for *very small* firms and about 0.1% for *small* firms. However, for *medium-sized* ($\beta_{\text{medium-sized}}$ = -0.061, $p > .1$) and *large* (β_{large} = 1.160, $p > .1$) firms, the relationship is not statistically significant, with a noteworthy small, negative sign for *medium-sized* firms. This suggests that the patterns observed in the models for smaller firms do not universally apply across the entire market. Consequently, the relationship between profitability and market share in larger segments appears to be more variable and less straightforward, rendering interpretation more complex.

R&D per Sale

Next, we examine the relationship between market share and innovative activities, measured by R&D expenses in relation to sales. Upon incorporating fixed effects and adjusting for robust standard errors, we find in Table 9 that R&D per sale does not show a statistically significant relationship ($p > .1$) with the logarithm of market share for any of the company size segments ($\beta_{\text{very small}}$ = 0.015; β_{small} = -0.129; $\beta_{\text{medium-sized}}$ = -0.399; β_{large} = 3.667). Consequently, while there may be an underlying relationship between market share and R&D expenses relative to sales, the association is uncertain, inconsistent, and ambiguous, especially after adjustments for fixed effects and robust standard errors have been made.

Log of Sales per Employee

When analyzing the productivity parameter, measured by the logarithm of sales per employee as shown in Table 9, it's important to note the logarithmic nature of the explanatory variable. We observe a positive and statistically significant association with the logarithm of market share for *very small* ($\beta_{\text{very small}} = 0.725$, $p < .01$), *small* ($\beta_{\text{small}} = 0.699$, $p < .01$), and *medium-sized* ($\beta_{\text{medium-sized}} = 0.725$, $p < .01$) firms, suggesting, as the explanatory variable is in logarithmic form, that a 1% increase in sales per employee is correlated with an approximate 0.7% increase in market share. For *large* firms ($\beta_{\text{large}} = 0.507$, $p > .1$), there is no significant relationship, but we notice that the sign of the coefficient is similar to that of the other company size segments.

These findings suggest that improved productivity, indicated by higher sales per employee, appears to consistently correlate with higher market shares. The positive coefficients suggest that productivity is a common trait for companies with greater market share in all segments.

Cash per Sale

In addition to the primary variables analyzed, we also examine control variables such as cash holdings per sale and ownership concentration, as shown in Table 9. Considering cash per sale, the data reveals that its relationship with market share is consistently negative and statistically significant across *very small* ($\beta_{\text{very small}} = -0.131$, $p < .01$), *small* ($\beta_{\text{small}} = -0.163$, $p < .01$), and *medium-sized* firms ($\beta_{\text{medium-sized}} = -0.272$, $p < .05$). This indicates that a one percentage point (one hundredth of a unit) increase in cash holdings per sale corresponds to a decrease in market share of approximately 0.1% for *very small* firms, 0.2% for *small* firms, and 0.3% for *medium-sized* firms. For *large* firms, the relationship is not statistically significant ($\beta_{\text{large}} = -0.432$, $p > .1$), but the negative coefficient suggests a similar trend. This may indicate a tendency for firms with larger market shares to maintain relatively smaller cash reserves.

Ownership Concentration

Lastly, we examine the relationship between concentration of ownership and market share. Table 9 reveals a significant negative link for *very small* ($\beta_{\text{very small}} = -0.453$, $p < .01$) and *small* ($\beta_{\text{small}} = -0.124$, $p < .05$) firms. The negative relationship for *very small* and *small* firms implies that an increase of a percentage point (one hundredth of a unit) in ownership concentration is associated with roughly a 0.5% and 0.1% decline in market share, respectively. While the coefficient for *medium-sized* and *large* firms is not significant ($\beta_{\text{medium-sized}} = 0.204$, $p > .1$; β_{large}

= 0.254, p > .1), the positive signs suggest a different pattern than for *very small* and *small* firms. These findings imply that there might be some systematic differences in the relationship between ownership concentration for the different company size segments.

Fixed Effects Linear Regression Models with Robust Standard Errors

	Response Variable: Logarithm of Market Share			
	Very Small *(time- & firm-specific)*	Small *(time- & firm-specific)*	Medium-Sized *(firm-specific)*	Large *(firm-specific)*
EBITDA Margin	0.162***	0.090*	-0.061	1.160
	(0.016)	(0.051)	(0.352)	(1.665)
R&D per Sale	0.015	-0.129	-0.399	3.667
	(0.090)	(0.083)	(0.499)	(3.361)
Log of Sales per Employee	0.725***	0.699***	0.725***	0.507
	(0.012)	(0.040)	(0.116)	(0.541)
Cash per Sale	-0.131***	-0.163***	-0.272**	-0.432
	(0.010)	(0.058)	(0.132)	(0.454)
Ownership Concentration	-0.453***	-0.124**	0.204	0.254
	(0.057)	(0.056)	(0.126)	(2.642)
Observations	17,080	3,695	541	98
Adjusted R^2	0.508	0.125	0.147	0.183
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Yearly effects accounted for in the medium-sized and large models

Table 9. Fixed Effects Linear Regression Output with Robust Standard Errors *Data source: Regnskapsdatabasen*

5.4.3 Adjusted Linear Models' Performance

The incorporation of fixed effects into our models has led to a noticeable decline in explanatory power (Adjusted R^2), which is a typical consequence when adjusting for unobserved heterogeneity. From Table 9, for *very small* and *small* firms, the models now account for both time-varying and firm-specific effects, leading to adjusted R^2 values of 50.9% and 12.5% respectively. This approach allows us to capture changes within each firm over time, as well as variations that occur across different firms in each time period.

On the other hand, in Table 9 we observe that for *medium-sized* and *large* firms, the models are focused on isolating and understanding the changes within each firm over time, while not specifically accounting for variations across different time periods. This results in adjusted R^2 of 14.8% and 17.0% for these categories. Compared to the baseline models presented in Table

9, which had adjusted R^2 values of 70.7%, 63.5%, 54.2%, and 30.6% respectively, the reduction in explanatory power is significant.

This decreased explanatory power is primarily due to the models' emphasis on capturing the nuances of changes within each firm, rather than between firms, for *medium-sized* and *large* firms. Similarly, for *very small* and *small* firms, the models' focus on both within-firm variations over time and across different firms at each time point removes variance that was previously attributed to individual firm differences. This refinement, although resulting in a lower adjusted R^2, provides a more accurate and insightful understanding of firm-specific dynamics and temporal variations, offering a more nuanced perspective on the characteristics of current market share.

5.4.4 Evaluating the Adjusted Linear Models

The introduction of fixed effects into our linear model has led to a notable alteration in the findings, reflecting the methodological rigor of our approach. We've applied *"two-ways"* fixed effects for *very small* and *small* firms to capture both time- and firm-specific variations. This approach leverages the rich dataset for these segments, allowing a comprehensive analysis that encompasses variations within and across firms. Conversely, for *medium-sized* and *large* firms, we've adopted *"individual"* fixed effects that focus on unique characteristics and changes within each firm over time, suitable due to the relatively limited number of observations in these categories. While this enhances model precision, it may reduce the overall explanatory power and generalizability of the results.

Incorporating robust standard errors further refines our analysis by mitigating potential heteroskedasticity and serial correlation, ensuring the statistical inferences drawn are both accurate and reliable across all models. These methodological enhancements collectively strike a balance between precision and breadth of analysis, allowing us to gain more precise insights regarding the characteristics associated with higher market share and likelihood of market share growth.

To further evaluate our refined models, our primary focus is on examining any residual autocorrelation. Despite accounting for the panel data structure with fixed effects, dependencies in the residuals may persist. In Section 5.3.3, we identified autocorrelation as the most prominent violation of our model assumptions, which is why we concentrate on this aspect here. For other underlying assumptions, such as normality, multicollinearity, and linearity, we refer to the detailed analysis in Section 5.3.3 and will not reexamine them in this section. We have addressed any potential heteroskedasticity by implementing robust standard errors using White's method.

To assess the extent of autocorrelation that remains in our models, we employ the Panel-Durbin-Watson test along with the Breusch-Godfrey/Wooldridge test for a more comprehensive analysis. The latter is particularly effective at identifying higher-order autocorrelation in panel data due to its consideration of the panel's structure and the inclusion of multiple time lags (Riveros, n.d.). Our Panel-Durbin-Watson test results indicate a rejection of the null hypothesis ($p < .01$), which posits no autocorrelation, across all company size segments ($DW_{very\ small} = 1.49$; $DW_{small} = 1.36$; $DW_{medium\text{-}sized} = 1.21$; $DW_{large} = 1.25$). Similarly, the Breusch-Godfrey/Wooldridge tests yield p-values below .01 for all models, confirming the presence of serial correlation in the idiosyncratic errors.

Despite the detected autocorrelation, it is noteworthy that the Durbin-Watson statistics for all models show an increase towards the value of 2 when compared with the baseline models discussed in Table 8 in Section 5.3.3. This shift closer to 2, which represents the absence of first-order autocorrelation, is indicative of an improvement in our advanced model specifications over the initial models.

Adjusting our models for fixed effects and robust standard errors has enhanced their comprehensibility and applicability. However, these adjustments also introduce certain limitations that need to be acknowledged. While fixed effects *(*firm and time-specific for *very small* and *small* firms, and firm-specific for *medium-sized* and *large* firms) enable us to control for unobserved heterogeneity, they also focus our analysis more narrowly on within-firm and/or within-time variations. This focus might limit our ability to make broader generalizations. Additionally, while robust standard errors address potential issues of heteroskedasticity and serial correlation, they might influence the perceived significance of certain variables.

5.5 Logistic Regression Models

In transitioning from our analysis using a linear regression model, we now turn our attention to a logistic regression model. This model, which we first introduced in Section 3.5.3, is designed to explore the relationship between the likelihood of market share growth and several key variables. Our primary focus is on three significant explanatory variables: the logarithm of market share, the EBITDA margin, and R&D spending as a proportion of sales. While we include additional variables for control purposes, these three are at the core of our investigation.

It's important to note that our model also integrates year dummies to adjust for time-based variations. However, we have omitted their coefficients from Table 10. This decision stems from our goal to zero in on the essential variables and their marginal effects on the response variable, ensuring that the model accounts for significant temporal factors without overcomplicating the interpretation.

Please be aware that the analysis and conclusions drawn for each variable assume that all other variables remain constant. For a comprehensive view of the regression results, refer to Table 10.

5.5.1 Addressing the Panel Data Structure

Before delving into the interpretation of our logistic regression model's coefficients, it is important to consider the impact of the panel data structure, an aspect we addressed through fixed effects and robust standard errors in the linear regression models. However, in our logistic regression analysis, we do not implement fixed effects due to several aspects. Firstly, incorporating fixed effects into logistic regression models is complex and not as straightforward as in linear models (Stammann et al., 2016, p. 2). One potential method, which involves including dummy variables for each company, poses a significant risk of multicollinearity. This could lead to distorted and unreliable results in our model. Additionally, the nature of our response variable, a binary indicator of market share growth, is expected to exhibit less autocorrelation over time compared to a continuous variable like market share. Growth in market share does not necessarily correlate strongly over time, reducing concerns about autocorrelation. While some level of autocorrelation may still exist, as we will be discussed in Section 5.5.4, it is assumed to be less pronounced than in the linear model context. Therefore, the need to account for fixed effects is not as imperative in our logistic regression model. To ensure accurate results in our study, we applied White's method for robust standard errors. This technique addresses issues like intraclustered autocorrelation in the data, the presence of outliers, and small sample sizes (Mansournia et al., 2020).

5.5.2 Interpreting the Logistic Models' Coefficients

In interpreting the coefficients of a logistic regression model, a unit increase in a given explanatory variable, while holding all others constant, is understood as the change in the logarithm of the odds (log odds) of a positive outcome occurring in the response variable.

Logarithm of Market Share

In the linear regression models, the logarithm of market share served as the response variable. However, in our logistic regression analysis, it is included as one of the primary explanatory variables. In Table 10, we explore the correlations between the logarithm of market share and the likelihood of market share growth across each company size segment. Generally, the results suggest that larger firms are more likely to experience market share growth, irrespective of the company size segment. Specifically, the coefficients for *very small* and *small* firms reveal a significant positive correlation with the likelihood of market share growth, with coefficients of $\beta_{\text{very small}} = 0.218$ ($p < .01$) and $\beta_{\text{small}} = 0.228$ ($p < .01$). The trend is even more pronounced for *medium-sized* firms, with a coefficient of $\beta_{\text{medium-sized}} = 0.478$ ($p < .05$).

Given that the explanatory variable is in logarithmic form, these results imply that a 1% increase in market share is associated with an increase in the odds of market share growth of approximately 25% for *very small* and *very small* firms, and 60% for *medium-sized* firms. It's important to note that these figures represent percentage increases in the odds of market share growth, which are always relative to the existing odds. Although a 60% increase might appear substantial, it is crucial to interpret this in the context of the logistic regression model, where such changes are in terms of odds rather than direct probabilities.

Additionally, it is worth mentioning that the coefficient for *large* firms, despite being positive ($\beta_{\text{large}} = 0.415$), does not reach statistical significance ($p > .1$). This discrepancy could stem from the *large* firms' model having fewer observations and higher variability, leading to less robust results compared to the models for other company size segments. While the coefficient for *large* firms is not significant, it depicts the same sign as the other company size segments, suggesting a positive relationship between the logarithm of market share and the likelihood of market share growth.

EBITDA Margin

Next, our analysis extends to explore the relationship between EBITDA margin and likelihood of market share growth, for which Table 10 presents notable positive correlations. For *very small* ($\beta_{\text{very small}}$ = 0.464, $p < .01$), *small* (β_{small} = 0.838, $p < .01$), and *medium-sized* ($\beta_{\text{medium-sized}}$ = 2.853, $p < .1$) firms, there is a statistically significant positive relationship. Specifically, an increase of one percentage point (one hundredth of a unit) in the EBITDA margin is associated with approximately 0.5%, 0.8%, and 3% in odds of market share growth for *very small, small* and *medium-sized* firms respectively.

Conversely, while the analysis indicates a positive trend for *large* firms (β_{large} = 1.330), this association does not reach statistical significance ($p > .1$). This lack of significance could be attributed to the limited number of observations and the presence of noise within the *large* firm model, a point we discussed earlier. Nevertheless, the positive coefficients across all company size segments are notable. They indicate a consistent and positive link between profitability, as measured by EBITDA margin, and the likelihood of market share growth.

R&D per Sale

In analyzing the relationship between innovative activities and the likelihood of market share growth, we turn our attention to the coefficient for R&D per sale, as shown in Table 10. Controlling for other variables, a significant and positive relationship emerges between R&D spending per sale and the likelihood of market share growth for *very small* firms, with a coefficient of $\beta_{\text{very small}}$ = 0.417 ($p < .05$). This implies that a one percentage point increase (or one hundredth of a unit) in R&D per sale is associated with an approximate 0.4% increase in the odds of market share growth for this company size segments.

However, it does not appear that this trend extends to the other company size segments. The coefficients for *small* (β_{small} = -0.087), *medium-sized* ($\beta_{\text{medium-sized}}$ = 1.036), and *large* (β_{large} = 137.23) firms are not statistically significant ($p >.1$). This lack of significance introduces too much uncertainty to assert a definitive relationship. While the coefficients are predominantly positive, except for a slight negative coefficient for *small* firms, the variability among these segments suggests that the relationship between R&D spending and market share growth likelihood is more complex and less uniform across the company size segments.

Cash per Sale

Next, we delve into one of the control variables, cash holdings in relation to sales. From Table 10, we examine the relationship between cash per sale and the likelihood of market share growth and observe notable variations across different firm sizes. For *very small* firms, we observe a negative correlation between cash per sale and market share growth. The coefficient ($\beta_{\text{very small}}$ = -0.089, $p < .01$) indicating that a one percentage point increase in cash per sale (one hundredth of a unit) is associated with a 0.09% decrease in the odds of market share growth. In contrast, *small* firms exhibit a positive relationship between cash per sale and the likelihood of market share growth (β_{small} = 0.321, $p < .1$), suggesting that one percentage point increase in cash per sale is associated with roughly a 0.3% increase in the odds of market share growth. For *medium-sized* and *large* firms, the analysis does not reveal statistically significant links ($\beta_{\text{medium-sized}}$ = -0.084, $p > .1$; β_{large} = 5.675, $p > .1$). The differences in coefficient signs, however, underscores the varied relationship between cash holdings and likelihood of market share growth between company size segments.

Company is in Oslo

The next control variable we consider is the binary variable for companies located in Oslo. Table 10 reveals a non-significant correlation ($p > .1$) between whether a firm's location is in Oslo and its likelihood of market share growth across *very small* ($\beta_{\text{very small}}$ = -0.013), *small* (β_{small} = 0.128), and *medium-sized* ($\beta_{\text{medium-sized}}$ = 0.102) firms. For *large* firms, however, there is a positive and significant association with being located in Oslo and the likelihood of market share growth (β_{large} = 1.360, $p < .1$). This suggests that for *large* firms, being located in Oslo is associated with an increase in the odds of market share growth of roughly 289%. However, this substantial effect could possibly be artificially high, due to the small sample size, and should be interpreted with caution. While other coefficients are not significant, only the coefficient for *very small* reveal a negative association. This trend points to a potential concentration of market power in Oslo, as companies in the city are more inclined to experience market share gains, when other variables are controlled. We do, however, stress the fact that only the large coefficient is significant when adjusting for robust standard errors, which in and of itself must interpreted carefully, due to few observations.

Age of the Company

In contrast to the linear regression models where the age of a company and the logarithm of market share are positively correlated, our logistic regression analysis uncovers a negative relationship between a company's age and its likelihood of market share growth. Notably, our findings in Table 10 indicate that as *very small*, *small*, and *medium-sized* firms age, their likelihood of market share growth tends to decrease ($\beta_{\text{very small}} = -0.047$, $p < .01$; $\beta_{\text{small}} = -0.072$, $p < .01$; $\beta_{\text{medium-sized}} = -0.032$, $p < .05$). In practical terms, this means that an increase of one year in the age of a company is associated with a decrease in the odds of market share growth by approximately 5% for *very small* firms, 7% for *small* firms, and 3% for *medium-sized* firms.

However, for *large* firms, the relationship between age and market share growth is negative but not statistically significant ($\beta_{\text{large}} = -0.015$, $p > .1$). This could be attributed to the presence of noise and a limited number of observations in the model for *large* firms. These findings suggest a possible trend where newer firms experience an initial phase of market share growth, which then tends to plateau or decline as they become more established over time.

Ownership Concentration

Considering the last control variable, concentration of ownership, the logistic regression model uncovers a consistent inverse relationship with the likelihood of market share growth, as evidenced by Table 10. Specifically, *very small*, *small*, and *medium-sized* all exhibit a decrease in the likelihood of market share growth as ownership becomes more concentrated ($\beta_{\text{very small}} = -0.275$, $p < .01$; $\beta_{\text{small}} = -0.558$, $p < .01$; $\beta_{\text{medium-sized}} = -0.950$, $p < .01$), respectively. For *large* firms, however, there is no significant link between the explanatory variable and the response variable ($\beta_{\text{large}} = -2.82$, $p > .1$). Nevertheless, the sign of the *large* coefficients suggests a similar trend to the other company size segments. This trend suggests that firms with less concentrated ownership are more likely to experience growth in market share.

Constant Terms

In the logistic regression models, the constants for different company size segments represent the baseline log odds of market share growth when all other explanatory variables are set to zero. While this scenario is highly theoretical, especially for variables like the logarithm of market share, the constants still provide valuable insights. They highlight inherent differences in baseline market dynamics across these segments, with *medium-sized* firms showing the highest baseline log odds, followed by *small*, *large*, and *very small* firms in that order, revealed by Table 10. This variance in constants suggests that each segment experiences unique market conditions and influences, even when controlling for the effects of other variables in the model. It is however worth mentioning that the constant for *large* firms is not significant (p > .1).

Logistic Regression Models with Robust Standard Errors

Response Variable: Market Share Growth (Binary: 1 = Growth, 0 = No growth)				
	Very Small	Small	Medium-Sized	Large
Log of Market Share	0.218***	0.228***	0.478**	0.415
	(0.017)	(0.056)	(0.208)	(0.764)
EBITDA Margin	0.464***	0.838***	2.853*	1.330
	(0.064)	(0.202)	(1.501)	(9.368)
R&D Per Sales	0.417**	0.087	1.036	137.230
	(0.173)	(0.180)	(2.224)	(122.299)
Cash per Sales	-0.089***	0.321**	-0.084	5.675
	(0.032)	(0.160)	(0.801)	(4.109)
Company is in Oslo	-0.013	0.128*	0.102	1.360*
	(0.035)	(0.073)	(0.205)	(0.824)
Age of the Company	-0.047***	-0.072***	-0.032**	-0.015
	(0.003)	(0.005)	(0.013)	(0.019)
Ownership Concentration	-0.283***	-0.566***	-0.955***	-2.823
	(0.060)	(0.105)	(0.300)	(1.866)
Constant	2.613***	2.950***	3.962***	3.161
	(0.176)	(0.456)	(1.423)	(3.412)
Observations	17,502	3,789	557	99
Akaike Inf. Crit.	22,533.920	4,805.969	708.247	124.168
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Yearly effects are accounted for in all models

Table 10. Logistic Regression Output with Robust Standard Errors *Data source: Regnskapsdatabasen*

5.5.3 Logistic Models' Performance

To assess the performance of the logistic regression models, we evaluate their accuracy in identifying true positives and true negatives. The Receiver Operating Characteristic (ROC) plots the true positive rate against the false positive rate at various thresholds, and a common metric derived from the ROC curve is the Area Under the Curve (AUC), providing an aggregate score of model effectiveness (Dey, 2021). An AUC score of 0.5 suggests that the model's predictive ability is no better than random guessing, while a score of 1 indicates perfect prediction accuracy. The AUC is good for evaluating the model performance also in the presence of disparity in the response variable. However, for all company size segments, the proportions of growth vs. no growth are relatively even.

Model	AUC	N
Very Small Firms	0.660	17,502
Small Firms	0.699	3,789
Medium-Sized Firms	0.720	557
Large Firms	0.855	99

Table 11. Area Under the Curve (AUC)
Data source: Regnskapsdatabasen

Assessing the AUC scores of each model, as shown in Table 11, we see that all models appear better than random guessing (AUC > .5). However, even though the *large* model has a significantly higher AUC score, it should be noted that this segment has substantially fewer observations than the rest, potentially making it more prone to overfitting.

Additionally, the Hosmer-Lemeshow test is used to evaluate the model's goodness-of-fit. This test measures whether there are significant deviations between the observed and the predicted outcomes (Bartlett, 2014). In the Hosmer-Lemeshow test, the null hypothesis is that there is no significant difference between the observed and the expected outcomes. Utilizing the test, we reject the null hypothesis for all models ($p < .01$), which implies poor goodness-of-fit. It should be noted that market share growth is a complex phenomenon, and accurately predicting it using only financial data is unrealistic. However, as previously mentioned, our aim is not necessarily to achieve high model performance, but rather to uncover intricate relationships.

It is also worth mentioning that the Akaike Information Criterion (AIC) is a commonly used statistic to evaluate model performance, considering both the complexity and how well the models fit the data (Bevans, 2023). However, it is most useful when comparing models fitted to the same data, and is therefore not appropriate for comparison between the company size segments but is a useful metric in the sensitivity analysis to be conducted in Section 5.6.2.

5.5.4 Evaluating the Logistic Models

When evaluating the logistic regression models, it's crucial to verify several key assumptions to ensure the validity of the model. These include: (1) the response variable must be binary, (2) the observations should be independent, (3) there should be minimal multicollinearity among predictors, (4) the presence of extreme outliers should be limited, (5) continuous explanatory variables should be linearly related to the log odds of the response variable, and (6) the sample size needs to be sufficiently large (Zach, 2020). Assumptions (1) and (4) have been addressed through the construction of the binary response variable and the exclusion of extreme values as discussed in Section 3.5.3. To mitigate the risk of multicollinearity, explanatory variables were carefully selected, as detailed in Section 5.2.1.

Independent Observations

In assessing the independence of observations for logistic regression, a key challenge is the lack of direct testing methods, unlike in linear regression. The panel data structure, describing organizations over multiple years, does not inherently violate the assumption of independence. However, it introduces complexity, particularly with the potential correlation of a company's data across different years. For example, a company's market share growth in one year might be influenced by its performance in preceding years.

Another aspect of independence in logistic regression relates to the error terms. Repeated measures within the same companies can lead to differing variances over time compared to between companies. Natural clustering, such as by company size or service offerings, may also impact independence, with observations within clusters potentially being more similar than those across different clusters. While our study accounts for size-based clustering, service segment clustering remains unaddressed. Geographic proximity can further complicate this, as companies in similar locations (e.g., Oslo) might exhibit more similarities. The inclusion of a binary variable indicating whether a company is in Oslo partially addresses this.

Despite these challenges, as discussed in Section 5.5.1, our approach of adjusting for robust standard errors helps mitigate potential issues arising from dependencies in the data. This adjustment is a crucial step in ensuring that our model's confidence intervals and standard errors remain valid, even in the presence of non-independence.

Linearity of Logit

Moving on to the analysis of the relationship between each explanatory variable and the response variable, our investigation revealed a noteworthy pattern. While we did not identify any significant non-linear relationships between the log odds of market share growth and the explanatory variables, it's also apparent that these relationships are not perfectly linear. To gain a more detailed understanding of these dynamics, we have included specific plots illustrating the linearity for each variable in Appendix A.3.2. This supplementary material offers a more nuanced view of the relationship patterns.

Sample Size

In evaluating our sample sizes, we encounter significant concerns, especially regarding the representation of *large* firms. The segment for large firms includes just 98 observations over a decade, equating to approximately 10 unique companies. This number is too small for statistically robust conclusions. For *medium-sized* firms, the sample size is slightly better at 541 observations, but this is still on the lower end of acceptability. In contrast, the segments for *very small* and *small* firms, with sample sizes of 17,080 and 3,695 respectively, are robust and adequate. Therefore, it's crucial to interpret the results pertaining to *large* firms with a high degree of caution due to the limited sample size.

5.6 Sensitivity Analysis

While both the linear and logistic regression models have been evaluated, it is important to further assess their rigor. In the field of regression analysis, it is imperative to examine how well the model's predictions hold up when parameters are changed (Frost, n.d.). In this section, we will assess the stability of our regression models by examining the resulting alterations in coefficients and explanatory power when systematically excluding explanatory variables. Ideally, we would analyze all company size segments for both the linear and the logistic regression models. However, to avoid repetition, we will only conduct a sensitivity analysis for the linear and logistic regression models fitted to the company size segment containing *very small* firms. This subset contains the greatest number of observations and is thereby considered most appropriate.

Consequently, we will omit the primary explanatory variables for both the linear regression and the logarithmic models, one by one. Sequentially removing variables, we can observe the shifts in the remaining coefficients and changes in the models fit to discern the relative importance of each variable. This iterative process helps identify what variables are central for the model's explanatory abilities. Note that all coefficients are adjusted for robust standard errors using White's method.

5.6.1 Linear Regression with Fixed Effects

The primary variables we omit in the linear regression model are EBITDA margin, R&D per sale and logarithm of sales per employee. The rationale behind omitting only the forementioned three, is the fact that these are of most relevance to the overall topic of this paper (Section

5.2.1). Variables that are not subject to omission will be kept in the model for the robustness of the models.

As Table 12 illustrates, excluding the EBITDA margin from our regression model does not significantly alter the coefficients of other variables or the overall explanatory power of the model. Although there's a slight decrease in the adjusted R^2, indicating a marginal reduction in the variance explained by the model, the increased F-statistic points to a heightened overall statistical significance. This suggests that the remaining predictors in the model, without the EBITDA margin, still maintain a substantial and even more pronounced collective association with the response variable. By removing the EBITDA margin, we are no longer considering its impact on market share variation. This exclusion implies that any variance in market share previously explained by EBITDA margin is now unaccounted for in the model.

Upon excluding the variable for R&D in relation to sales, we observe in Table 12 that the coefficients and explanatory power of the model remain consistent with the original configuration, as shown in the unchanged adjusted R-squared. A notable change, however, is the increase in the F-statistic, indicating an enhanced overall significance of the model. This increase suggests that the exclusion of R&D per sale does not diminish the model's ability to explain variations in the logarithm of market share. In essence, it appears that, holding other factors constant, R&D spending relative to sales does not significantly influence the logarithm of market share.

Sensitivity analysis of Fixed Effects Linear Regression Models with Robust Standard Errors

	Response Variable: Logarithm of Market Share			
	Complete	No EBITDA	No R&D	No Productivity
EBITDA Margin	0.162^{***}		0.163^{***}	0.674^{***}
	(0.016)		(0.016)	(0.026)
R&D per Sale	0.015	0.045		-0.217^{*}
	(0.090)	(0.092)		(0.114)
Log of Sales per Employee	0.725^{***}	0.761^{***}	0.725^{***}	
	(0.012)	(0.011)	(0.012)	
Cash per Sale	-0.131^{***}	-0.134^{***}	-0.131^{***}	-0.310^{***}
	(0.010)	(0.011)	(0.010)	(0.020)
Ownership Concentration	-0.453^{***}	-0.439^{***}	-0.454^{***}	-0.495^{***}
	(0.057)	(0.057)	(0.057)	(0.074)
Observations	17,080	17,080	17,080	17,080
Adjusted R^2	0.508	0.500	0.508	-0.045
F Statistic	$4,418.417^{***}$ (df = 5; 12650)	$5,373.445^{***}$ (df = 4; 12651)	$5,523.392^{***}$ (df = 4; 12651)	923.166^{***} (df = 4; 12651)
Significance levels				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 12. Sensitivity Analysis of Fixed Effects Linear Model for Very Small Firms *Data source: Regnskapsdatabasen*

We notice in Table 12 that removing the productivity parameter, logarithm of sales per employee, from the model leads to substantial changes, unlike the removal of EBITDA margin and R&D per sale. This exclusion notably alters the size of the coefficients for the remaining variables, and R&D per sale emerges as a significant predictor. A particularly striking change is the model's negative adjusted R^2. A negative value here suggests that the model, without this variable, fits the data worse than a simple horizontal line at the mean of the response variable. This unusual occurrence might, however, be related to the absence of an intercept in the model, a typical outcome when fixed effects are incorporated (Edwards et al., 2008). The drastic decrease in adjusted R^2 implies that the logarithm of sales per employee plays a crucial role in explaining variance in the response variable. Despite this, the model still exhibits a relatively high F-statistic, indicating maintained overall significance, albeit reduced compared to the complete model. The increase in coefficients for the remaining variables suggests that the productivity parameter might have been capturing some of the variance in market share that is now attributed to other variables. Thus, the logarithm of sales per employee emerges as a pivotal variable in our analysis.

5.6.2 Logistic Regression

For the sensitivity analysis regarding the logistic regression model, we apply the same rationale as for the linear regression model, sequentially omitting the primary explanatory variables discussed in Section 5.2.1: EBITDA margin, R&D per Sale, and logarithm of market share. Upon removing the EBITDA margin from our model, we observe in Table 13, minor changes in the significance levels and coefficients of other variables. Notably, cash per sale shifts from being significant at the 1% level to the 5% level. This shift could indicate that the EBITDA margin was partly capturing effects similar to those of cash holdings in relation to sales on market share growth likelihood. In its absence, cash per sale emerges as a more prominent predictor. Additionally, there is a slight increase in both the log likelihood and the AIC, implying a modest reduction in the overall quality of the model. These changes suggest that while the EBITDA margin may not be the most central variable in our model, its presence does contribute to a more nuanced understanding of factors that might be influencing likelihood of market share growth.

Excluding R&D per sale from our model, as shown in Table 13, results in no significant shifts in the significance of other variables and only minor alterations in their coefficients. This stability indicates that R&D per sale may not be a crucial predictor in this context. Additionally, there is a less pronounced increase in both log likelihood and the AIC compared to the removal of EBITDA margin. This implies that the R&D per sale variable, while contributing to the model, does not have as substantial impact on the model's overall fit and predictive accuracy as the EBITDA margin. These findings suggest that R&D per sale plays a relatively minor role in explaining the likelihood of market share growth in our analysis. Its removal, therefore, does not significantly diminish the model's explanatory power, unlike the more notable effects observed with the removal of EBITDA margin.

Sensitivity Analysis of Logistic Regression Model with Robust Standard Errors

	Response Variable: Market Share Growth (Binary: 1 = Growth, 0 = No growth)			
	Complete	No EBITDA	No R&D	No Market Share
EBITDA Margin	0.464***		0.465***	0.661***
	(0.064)		(0.064)	(0.063)
R&D per Sale	0.417**	0.430**		0.506***
	(0.173)	(0.172)		(0.175)
Logarithm of Market Share	0.218***	0.256***	0.220***	
	(0.017)	(0.016)	(0.017)	
Cash per Sale	-0.089***	-0.063**	-0.092***	-0.194***
	(0.032)	(0.030)	(0.032)	(0.032)
Company is in Oslo	-0.013	-0.022	-0.015	0.033
	(0.035)	(0.035)	(0.035)	(0.035)
Ownership Concentration	-0.283***	-0.186***	-0.299***	-0.499***
	(0.060)	(0.059)	(0.060)	(0.057)
Age of the Company	-0.047***	-0.047***	-0.047***	-0.046***
	(0.003)	(0.003)	(0.003)	(0.003)
Constant	2.613***	2.947***	2.647***	0.600***
	(0.176)	(0.170)	(0.175)	(0.081)
Observations	17,502	17,502	17,502	17,502
Log Likelihood	-11,248.960	-11,285.640	-11,251.900	-11,339.170
Akaike Inf. Crit.	22,533.920	22,605.290	22,537.800	22,712.330
Significance levels				*p<0.1; **p<0.05; ***p<0.01

Yearly effects are accounted for in all models

Table 13. Sensitivity Analysis of Logistic Model for Very Small Firms *Data source: Regnskapsdatabasen*

Upon the removal of the logarithm of market share from our logistic regression model, we notice in Table 13 that the R&D per sale coefficient becomes notably significant at a 1% level, compared to a 5% significance in the complete model, indicating a heightened impact of R&D activities on the likelihood of market share growth. Meanwhile, the coefficients size for EBITDA margin and R&D increases, showing their increased influence without market share data. In contrast, cash per sale and ownership concentration become more negatively significant, suggesting that higher cash holdings and ownership concentration are linked to a stronger decline in market share growth likelihood. The most substantial increase in both log likelihood and AIC compared to other variable removals highlights the logarithm of market share's critical role in the model.

Comparing the sensitivity analysis of the logistic regression model to the linear regression model, the logistic regression model seems more stable in its predictors, with only slight variations in model quality.

5.7 Summary of Regression Results

In Chapter 4, our analysis focused on key business metrics including market share, growth, profitability, productivity, and R&D expenditure. We examined each of these parameters aggregated over time to uncover overarching market trends. Chapter 5 shifted the focus to a detailed statistical examination of market share and the likelihood of market share growth, analyzing related factors, both across and within companies. This involved considering variations among firms in the selected parameters, thereby shedding light on the underlying dynamics of the sector.

The goal of the regression analysis was to examine the complex relationships regarding market share growth and current market position. The key aspects we wanted to consider were profitability, productivity, investments in innovative activities and existing market position. Examining models fitted to the different company size segments revealed nuanced insights to the market composition.

Our analysis, comprising both linear models with and without fixed effects and logistic models, offers key insights into market dynamics. The analysis reveals a negative relationship between EBITDA margin and market share in initial linear models. However, when adjusting for fixed effects, the relationship becomes positive for *very small, small,* and *large* firms. Furthermore, in the fixed effects models, R&D spending's impact on market share is inconsistent and non-significant, indicating a complex interaction with market presence. Productivity, however, positively correlates with market share across all company size segments.

The logistic models show a positive, albeit for some company size segments, uncertain correlation between the likelihood of market share growth, and all the explanatory variables: logarithm of market share, EBITDA margin and R&D spend per sale. The current market share and EBITDA margin are significantly correlated for all company size segments except for *large* firms, while R&D per sale is only significant in the *very small* segment.

Overall, incorporating fixed effects in the linear regression models and robust standard errors in both the linear and logistic regression partially addresses issues like heteroskedasticity and autocorrelation. This analysis highlights the role of factors like EBITDA margin, productivity, and market share in understanding these dynamics in the software development sector. As mentioned, several times, it should be noted that the company size segments vary substantially in size. The *medium-sized* and *large* models should therefore be interpreted with additional caution.

Chapter 5's regression analysis offers a detailed examination of variations within and between firms, providing a more nuanced perspective than the broader trends highlighted in Chapter 4. Both chapters, however, converge on the observation of an increasingly fragmented market landscape, which has diluted the average market share of existing companies without significantly impacting their sales or operations. Moving into Chapter 6, we shift from numbers to narratives, presenting survey results from *vendors* and *clientele*. These insights seek to bridge the gap between overarching trends, statistical patterns and the real-world factors influencing software acquisitions, market outlook and innovation.

Chapter 6 - Survey Results

6.1 Findings from the Questionnaires

In this chapter we are moving on from the regression analysis and examine the key findings from the data collected through questionnaires, to get real-world indications of the market attitudes. As mentioned in Section 3.4.1**,** one questionnaire is tailored towards companies whose offering is software developments services (*vendors*), and the other towards companies whose core offering is not that of the *vendors* and are thereby considered potential buyers of such services (*clientele*). We will summarize the findings from both *vendors* and *clientele* without a strict structural distinction, as they often revolve around similar topics. However, we will clearly specify to whom the responses belong. Key findings will be highlighted in graphical representations for clarity and emphasis. For the complete questionnaires, please see Appendix A.4.

The objective for this chapter is to provide valuable insights tailored towards answering Research Question 3.

6.1.1 Respondent Remarks

Both the *vendors* and *clientele* questionnaires got a total of 38 respondents each. Of the 38 responses, the proportion acquired from the CEO of the companies are 71% for *vendors* and 84% for *clientele*. Obtaining responses from the person in charge of the organization is a way to strengthen the validity of the responses, while seeking to avoid potential biases from certain domains of expertise. It is worth noting that 32% and 43% of the responses for *vendors* and *clientele*, respectively, are from companies that are part of conglomerates. The remaining responses are thus from independent companies.

Examining digitalization and IT competence within the *clientele* firms, revealed an interesting pattern. When asked about their own company's digital maturity, they show a tendency towards considering themselves superior to their competitors. Assessing their own digital maturity, 43% rated themselves as mature, while only 14% of competitors were rated as mature. Regarding IT competence, 41% consider their company as above average, while only 22% consider their own competence as above average. Only 8% and 5% consider their company's and personal IT competence, respectively, to be below average, and 51% vs 73%, to be average.

When assessing different business models for *vendors*, we consider Business to Business (B2B), Business to Customers (B2C), and Business to Business to Customers (B2B2C). For the definitions of the service offerings being considered, see Section 1.1. Regarding the business models for the *vendors*, 61% offer B2B services, 50% provide B2B2C services and 19% offer B2C services. We notice that several companies have multiple business models, as well as service offerings, evident by 70% offering Software as a Service (SaaS) solutions, 58% selling IT consulting, and 33% delivering white label (off-the-shelf) solutions. In addition to the top three service offerings, a variety of other services were mentioned, indicating a diverse and heterogeneous market.

Overall, these statistics suggest that the companies responding from both *vendors* and *clientele* are appropriate prospects to consider, and that the individuals answering are competent enough to provide meaningful insight.

6.1.2 Software Status

Considering software status, we consider two types of software: *external* and *internal*. External software refers to software facing the *clientele's* own customer base, while *internal* software refers to software intended for internal operations. When *clientele* companies are asked about their software status, 59% report currently having external software while 92% report having internal software, as revealed in Table 12. Among those currently having external software, 77% plan to acquire more external software, 18% have no plans for additional acquisitions and only 5% are uncertain. Regarding companies currently having internal software, Figure 12 reveals that 56% plan further acquisitions of internal software, 21% do not, and 23% are unsure. Among those with currently no external software, 27% plan to acquire external software in the future, 67% have no such plans and 6% are undecided. For *clientele* with no current internal software, 33% plan to acquire internal software in the future, while 67% have no such plans. This indicates that companies with external or internal software appear more likely to continue acquiring such software in the future, compared to those with no such software.

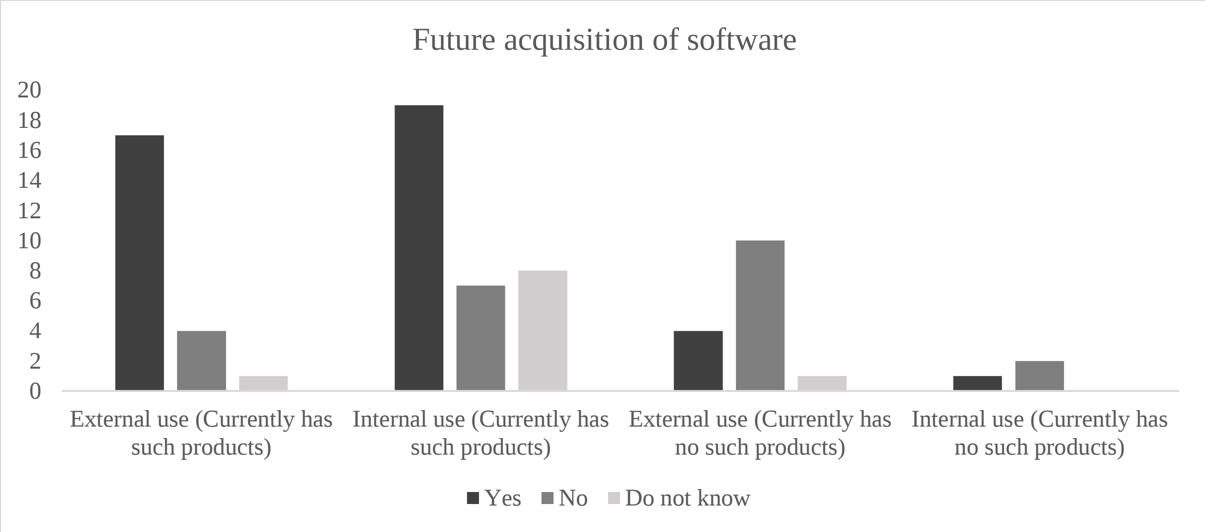


Figure 12. Future Software Acquisition *Data source: Clientele questionnaire*

For acquisitions of internal and external software, 53% of *vendors* identify flexibility and customizability as customers' top priority, as seen in Figure 13, followed by support and maintenance (16%), good communication (13%), quick delivery (9%), low price (6%) and at last scalability (3%). The *clientele* reveals similar findings, with 67% and 54% emphasizing flexibility and customizability as their top priority, for external and internal software, respectively. Regarding scalability, 22% of the *clientele* considers it the most important factor when acquiring external software, dropping to 14% from internal software. Overall, there are many similarities between the most important factors for both external and internal software, with the *vendors* showing a relatively consistent understanding of customer preferences.

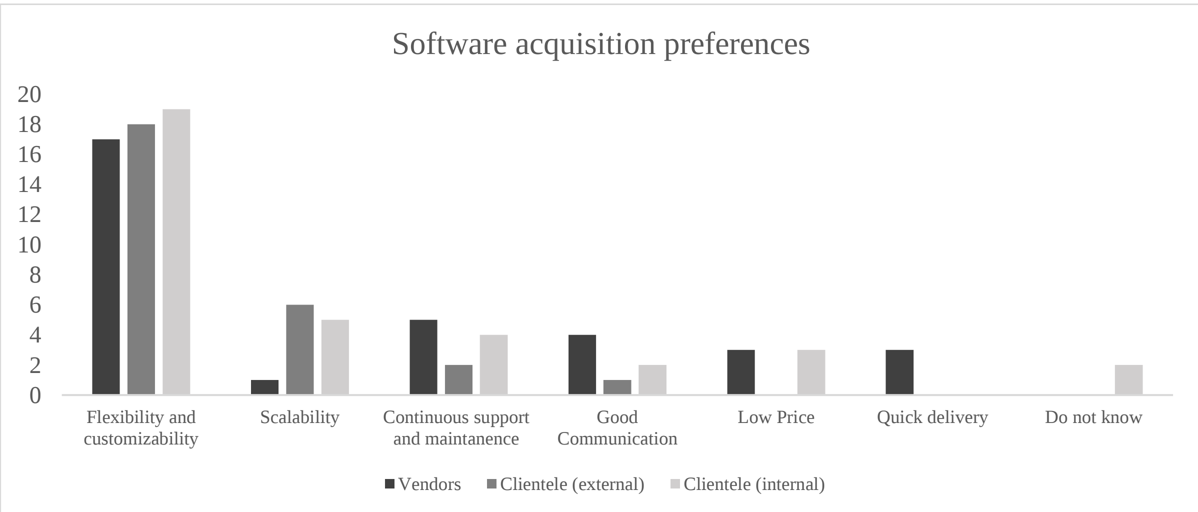


Figure 13. Software Acquisition Preferences *Data source: Vendors questionnaire & Clientele questionnaire*

Regarding overall satisfaction with current software, 64% of *clientele* with external software report being satisfied. For *clientele* companies with internal software, 68% say they are satisfied. This implies a medium to high level of contentment across both external and internal software. These findings suggest that many companies have had a positive experience with their current software.

6.1.3 Acquisition Methods

When asked about how companies have acquired their current software, the *clientele* reveals several prominent methods for both external and internal software. It should, however, be noted that respondents were allowed to answer multiple methods, as often is the case in software acquisition (Argolini, et al., 2022). Revealed by Figure 14**,** white label solutions are used by 45% for external software, and 62% for internal software. For external software, 64% report hiring internal resources, while only 38% report doing so for internal software. 41% reports using SaaS-solutions for external software, compared to 50% for internal software. Consulting services seem to be more popular when acquiring external software than internal software, with a 50% use-rate compared to 35%, respectively. Outsourcing is reported to have been used by 32% regarding external software, and 21% for internal software. The five forementioned methods seem to capture the essence of how software is acquired. Another qualitatively reported method for software acquisition is open-source solutions. Overall, the responses indicate a significant variation in preferences. When acquiring external software, companies tend to rely more on hiring internal resources and consulting services. For internal software acquisitions, white label and SaaS-solutions are more popular. This diversity highlights the varied strategies businesses adopt in the digital landscape.

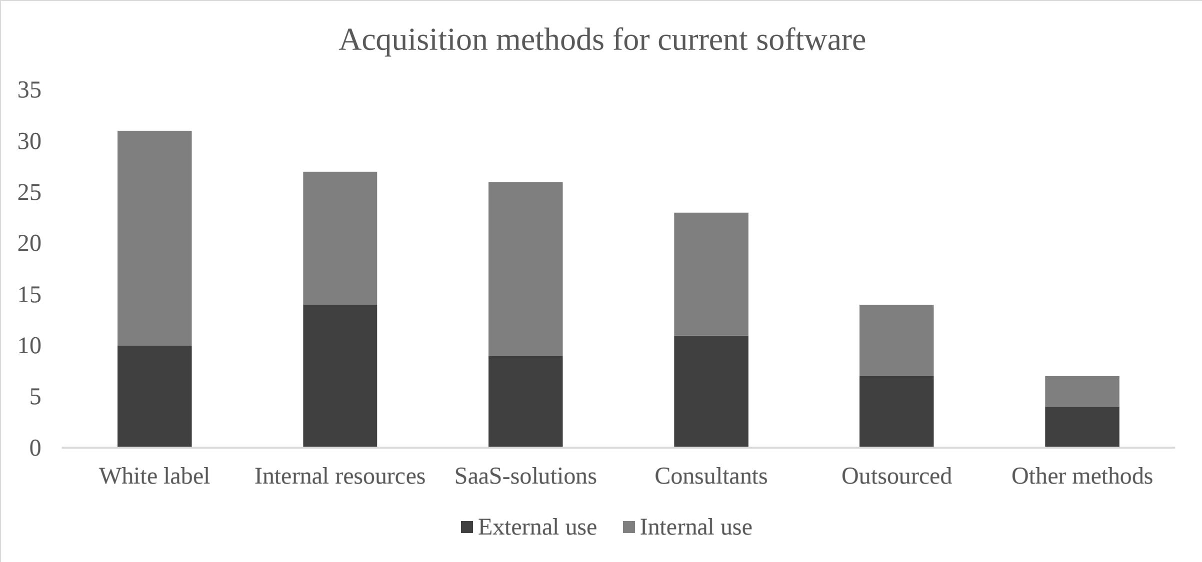


Figure 14. Current Software Acquisition Method *Data source: Clientele questionnaire*

To map out potential changes in preferences, *clientele* companies with external and/or internal software were asked about future acquisition plans. For companies currently having external software, we see in Figure 15 a noteworthy increase in interest in white label solutions, with 60% of companies considering this method for future acquisitions of external software. Additionally, there appears to be a reduced interest in hiring internal resources for external software, decreasing from 64% to 47%. For internal software, hiring internal resources seems more applicable, with an increase of nine percentage points. Only 29% express interest in using SaaS-solutions for external software in the future, marking a decrease of twelve percentage points. The use of consultants appears rather unchanged for external software, but a decrease from 35% to 21% for internal software. It should be noted that while these variations are notable, we have not considered their statistical significance. Nevertheless, the overall preferences regarding acquisition methods appear to have changed somewhat. White label seems more relevant for external software than before, and internal resources and SaaS solutions less relevant. For future internal software acquisitions, using consultants seems less relevant and more companies seem to favor hiring internal resources.

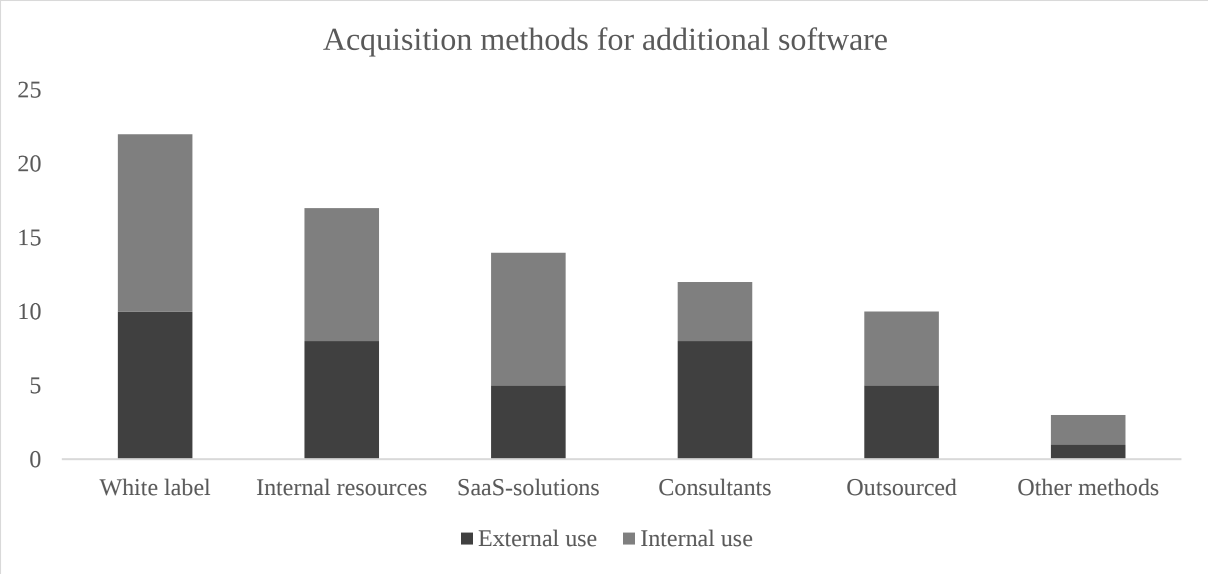


Figure 15. Additional Software Acquisition Methods *Data source: Clientele questionnaire*

6.1.4 Innovation

When asked about innovation in the market they operate in, 52% of *vendors* report they notice competitive pressure from new innovators, while 45% report not feeling pressured, as revealed by Figure 16. 3% were either unsure or had no opinion.

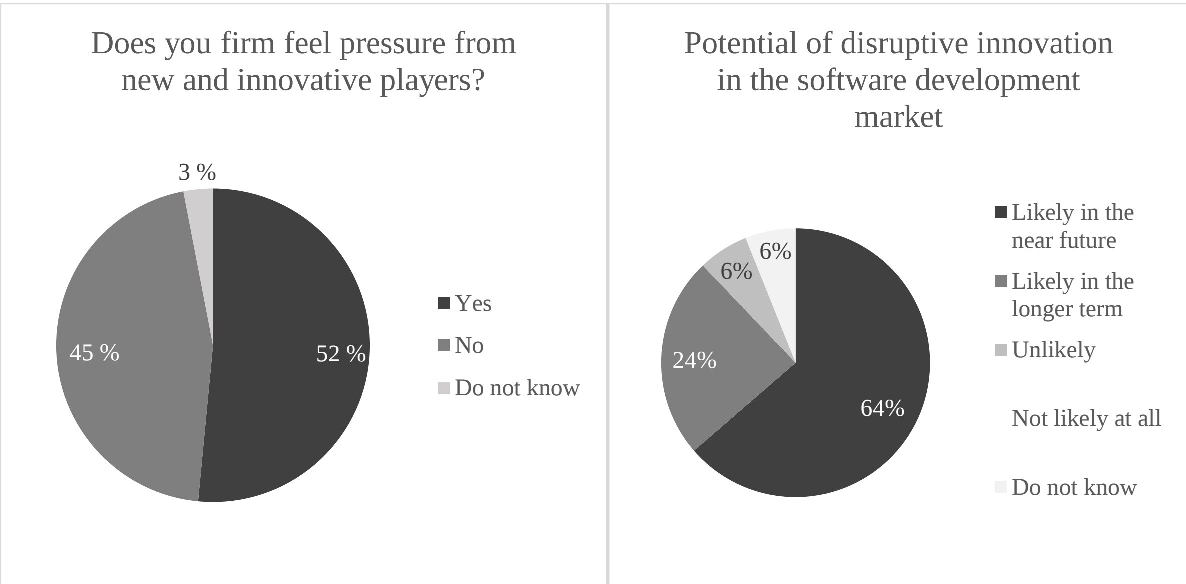


Figure 16. Innovation and Disruption *Data source: Vendors questionnaire*

Furthermore, 64% of *vendors* anticipate disruptive innovations in the near future, while 24% are expecting it in the longer term. Only 6% see disruptive innovation as unlikely, with none completely dismissing the possibility. However, 6% seem to be unsure of the likelihood of disruption. In preparation of innovative competitors and disruptive innovations, 70% are investing in R&D and 61% are monitoring market trends. Only 6% have established a separate innovation unit. An additional 24% are taking other measures. Other qualitatively reported strategies include descaling, outsourcing, customer-driven innovation, seminars, and skill development. Notably, 6% report no active preparations.

Regarding low-code/no-code solutions, technologies that minimize the need for in-depth programming knowledge to develop software products and services, 36% of *vendors* acknowledge an impact on their companies as evident by Figure 17 (Oberoi, 2021). In contrast, 52% report that these solutions have had no impact on their business operations. A rather high 12% seem uncertain about whether low code/no code solutions has affected their company. All of those who report feeling the impact of such solutions consider the impact to be positive. When asked about what parts of their business model has been affected, 75% of those feeling an impact experience changes in their B2B operations and 58% in their B2C operations. Note that respondent could select multiple answers.

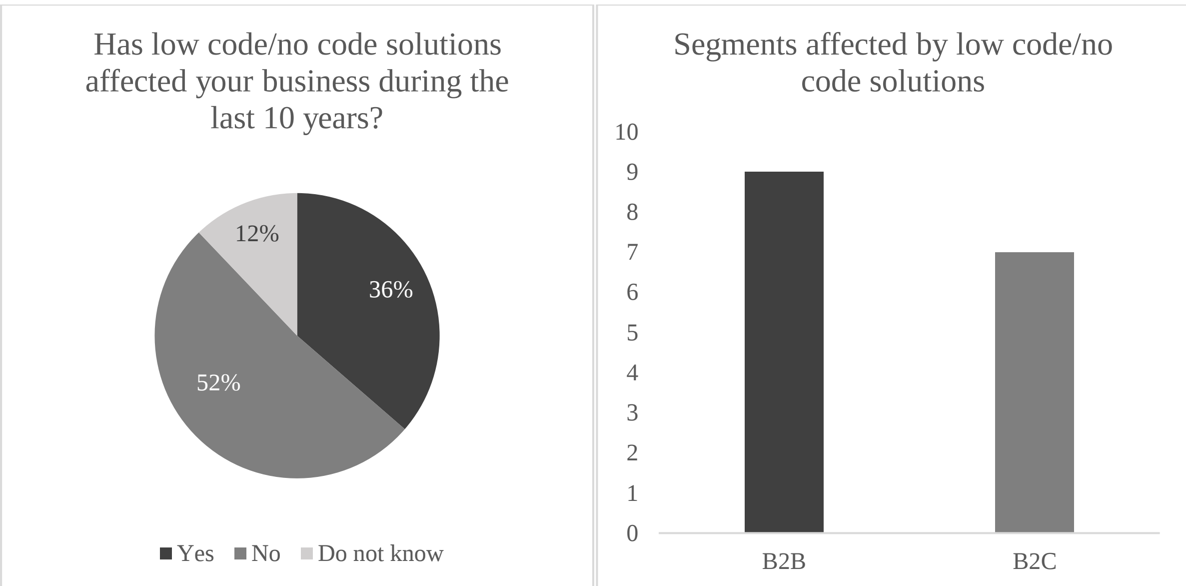


Figure 17. Effects from Low code / No code *Data source: Vendors questionnaire*

Regarding the influence of Artificial Intelligence (AI) on software acquisition, 59% of *clientele* anticipate a significant impact on their external software procurement, 32% foresee some impact, 5% expect a minor impact, and 3% are unsure. Similarly, for internal software, 57% predict a major impact, 35% expect some impact, 5% anticipate a minor impact, and 3% are uncertain about the influence of AI.

Figure 18 reveals that using less reputed distributors for external software is perceived as a high risk by 41% of *clientele*, a moderate risk by 41%, low risk by 14%, and 5% unsure. For internal software, 24% perceived high risk, 51% moderate risk, and 16% low risk. 5% were unsure in both categories.

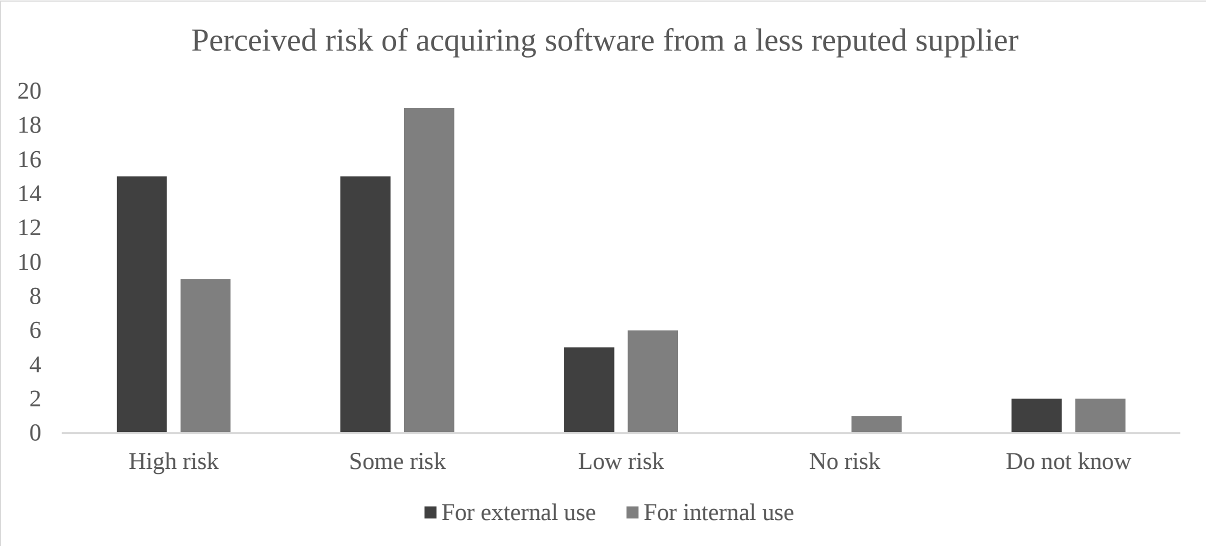


Figure 18. Risk Aversion *Data source: Clientele questionnaire*

6.1.5 Growth and Profitability

During the past decade, 72% of *vendors* report considering the growth in the software development market to be high, 25% consider the growth moderate, and 3% reports low growth, as revealed by Figure 19. The sentiment towards future growth, up to 2030, reveals a slightly more optimistic perspective, with 81% expecting high growth and 19% expecting moderate growth. Regarding the overall competition in the sector, 61% of *vendors* expect the competition to increase with more firms entering the market, 22% foresee no change, and 17% predicts a decrease in competition. Furthermore, 31% anticipate increased profitability in the future, 42% expect it to remain at today's levels, and 28% predicts a decline.

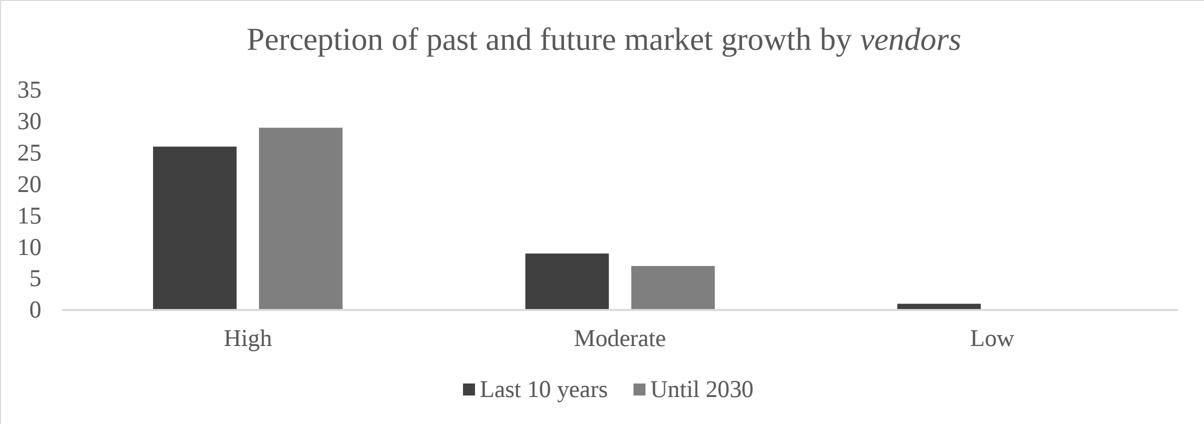


Figure 19. Market Growth Perceptions *Data source: Vendors questionnaire*

When asked about external and/or internal software expenses, the *clientele* reveals no signs of significantly increasing their spending in the future. Roughly half of companies' report having spent less than 1MNOK on acquiring their current solutions, and 59% is spending less than 1MNOK annually for maintenance and improvements. With a similar future spend, the optimistic sentiment from *vendors* regarding future profitability might be somewhat unrealistic.

6.2 Summary of Survey Results

Utilizing the questionnaires to gain insights in both suppliers (*vendors*) and potential customers (*clientele*) of software solution in Oslo, we notice that most of the *clientele* currently utilize both internal and external software solutions and generally report satisfaction with these solutions. Notably, firms that already employ external or internal software are more inclined to continue investing in similar solutions. Until recently, the acquisition of external software has primarily involved hiring internal resources and consultants. However, there's a growing preference for white label solutions for future acquisitions. In contrast, while white label and SaaS-solutions have been popular for internal software, the trend is shifting towards hiring more internal resources.

This evolving landscape is marked by a change in acquisition patterns, a shift further underscored by the expectation that AI-technology will influence future methods of acquisition. Regarding *vendors*, most companies offer SaaS and consulting services. Both *vendors* and *clientele* highlight flexibility and customizability as most important in both internal and external software acquisitions.

Approximately half of the *vendors* in the market are feeling the pressure from emerging, innovative solutions and companies. This supports our findings in Chapter 4, regarding the influx of new companies. A substantial majority of *vendors* foresee the potential for disruptive innovation in the near future. This anticipation aligns with the expectation of a more diverse and competitive market. However, there is a notable hesitance among *clientele* regarding the risks associated with procuring software from less reputed distributors.

Regarding the impact of low code/no code solutions, about 40% of the *vendors* have been affected, with all reporting a positive influence. Looking ahead to 2030, vendors maintain a positive outlook on growth and profitability. Contrastingly, the clientele does not anticipate an increased willingness to pay, suggesting a relative stability in profitability, which aligns with the profitability trends found in Chapter 4.

As Chapter 6 has shed a light on the *vendors* and *clientele* perspectives, it sheds light on the changing acquisition patterns and the expected impact of innovative technologies, highlighting the complex and multifaceted dynamics of the Norwegian software development sector and market. Additionally, it has revealed new insights to that of the findings in Chapter 4 and Chapter 5. While the *vendors* perceptions regarding profitability, growth and market fragmentation roughly aligns with our findings from the previous chapters, there are some notable deviations, for instance regarding R&D investments, that will be discussed in the following chapter.

Chapter 7 - Addressing the Problem Statement

In the previous chapters, we have examined the software development sector and market, utilizing multiple methods. In Chapter 4, we used exploratory analysis to highlight overarching trends, while Chapter 5 reveals statistical insights into more granular market dynamics. The data gathered through questionnaires which is presented in Chapter 6, shed a new light on the data analyzed in the two prior chapters. In this chapter we will use our findings from the previous chapters to address the problem statement by subsequently considering each research question.

7.1 Research Question 1

The first research question sets out to explore recent historical patterns to reveal any positive or negative trends in the secondary data. Utilizing exploratory data analysis, we have uncovered interesting patterns that help address the question.

Research Question 1

What are the recent trends in the Norwegian software development sector and market?

Over the past two decades, our analysis of the Norwegian software development market unveils a landscape of significant transformation and growth. This evolution is characterized by various key trends that together paint a picture of a dynamic and evolving sector.

Firstly, the market has witnessed a remarkable expansion, evident not only in its total sales, but also in the burgeoning number of companies and employees. This growth, however, has been accompanied by a notable shift in the market's structure. The top 5 companies, which once held a substantial portion of the market share, have seen their dominance wane, decreasing from 40% to 14%. This reduction underscores a more fragmented market, with *small* and *medium-sized* firms gaining ground at the expense of larger players. Such a trend towards decentralization suggests a more competitive and diverse marketplace, with a wider range of companies contributing to the sector's growth.

In terms of innovation, there has been a significant upward trend in average CPI adjusted R&D expenditure, particularly noticeable post 2017. This increase in investment in research and development points to a heightened focus on innovation within the sector. The implications of this trend are further explored in Section 7.3.3, in the context of responses from questionnaires.

To summarize, the Norwegian software development market from predominantly 2000 to 2020 has been marked by robust growth in sales, number of companies, and employment. However, this growth has led to a more fragmented and decentralized market structure, with each company generally earning less revenue and employing fewer people over time. Despite these changes, the market's profitability has remained relatively stable, with only minor fluctuations. In more recent years, there's been a discernible increase in R&D spending, indicating a growing emphasis on innovation. Additionally, productivity, as gauged by CPI adjusted sales per employee, has shown a positive trend in the last decade, reflecting the sector's adaptability and resilience in a rapidly evolving technological landscape.

7.2 Research Question 2

While the purpose of Research Question 1 was to get a general overview of the evolution in both the sector and market, Research Question 2 delves more analytically into company specific characteristics regarding market position and growth. To answer Research Question 2, we utilized regression analysis, allowing for an uncovering of statistically significant relationships.

Research Question 2

What are the characteristics of companies that gain market shares, and which aspects indicate current market position?

7.2.3 Gaining Market Shares

In this section, we delve into the characteristics associated with the likelihood of increased market share. Specifically, this exploration sheds light on the nuanced relationships between current market position, profitability, and R&D spending, and their relation to the likelihood of market share growth.

Current Market Share and the Likelihood of Growth

Our analysis has revealed a positive correlation between a company's current market position and its likelihood of achieving market share growth across the sector. However, it's important to note a lower degree of certainty in this assessment for the largest companies, which can be attributed in part to the model’s limitations when dealing with smaller sample sizes. Furthermore, this analysis suggests that for established players in the software development market, maintaining large-scale operations could prove advantageous in securing and expanding their market share. It is crucial to recognize, as highlighted in Chapter 4, that the trend towards market decentralization seems primarily driven by the entry of numerous new companies into the sector. This influx effectively dilutes the market shares of all existing companies. Rather than smaller companies directly taking market shares from larger ones, it appears that the overall market is undergoing an evolution and expansion. This shift is potentially creating new opportunities, which these emerging companies are adeptly exploiting.

Profitability and the Likelihood of Growth

Similar to current market share, we also observe a consistent positive trend across the sector suggesting that higher profitability tends to correlate with a greater likelihood of market share growth. A similar pattern emerges, where the certainty is highest for smaller firms. This effect stems from the disproportion composition of the market, with an abundance and *very small* and *small* firms.

This observation suggests that profitability is a strong predictor of a company’s likelihood of market share growth, regardless of company size. This is relatively intuitive, as profitable companies often enjoy greater economic flexibility, which facilitates easier access to capital and supports the expansion of operations. Moreover, being profitable tends to lower the risks associated with business growth, exploring new ventures, and investing in emerging technologies. In the context of the software development sector, however, this conclusion might initially seem unexpected. The tech market has been known for its startups receiving significant investments for extended periods without achieving profitability, with investors typically prioritizing growth and market consolidation, although this has somewhat changed recently (Guzel & Wei, 2023). Yet, our regression analysis reveals a contrasting trend. It shows that profitability is indeed an important factor in predicting growth and expansion within the software development sector.

R&D Ratio and the Likelihood of Growth

The relationship between R&D spending and the likelihood of market share growth in the software development sector is uncertain. Common intuition could lead to an assumption that an increase in R&D investment, as a driver of innovation, would naturally correlate with likelihood in the rise in market share growth. There exists a slight positive but uneven correlation between R&D spending per sale and the likelihood of market share growth across the sector. However, this relationship holds statistical significance only for *very small* firms. Despite these insights, the value of R&D should not be dismissed outright. It is plausible to consider that the impact of R&D investments may not be immediate and could be overlooked by our logistic regression model. This suggests that the advantages of R&D spending might emerge over a more extended period, and the cumulative effect of sustained R&D investments over time could be necessary to discern a more evident relationship with market share growth. Several factors add complexity to quantifying the effect of R&D spending on market share growth. Firstly, while all correlations observed are positive, their strength varies. Secondly, limitations in the variables within our analysis might conceal the actual impact of R&D spending. Additionally, the possibility exists that there is not a direct causal link between R&D spending and market share growth. In conclusion, although there are minor indications of a positive relationship between R&D spending and market share growth, the overall picture is still not clear, underlining the need for further detailed study.

7.2.4 Current Market Position

Moving on from the assessment of the characteristics of companies that gain market shares, we will now assess how our analysis highlight which aspects indicate current market position.

Productivity and Market Share

Regarding productivity, the fixed effects linear regression model revealed a strong positive correlation between productivity, measured as sales per employee, and the market share. This finding was consistent across the market, but less certain for the *large* company size segment. Nevertheless, productivity emerges as the most significant predictor of market share. A potential explanation for this relationship could be that larger companies might benefit from economies of scale, achieving higher productivity through more efficient systems and resource allocation. This might boost productivity as they could have fewer employees relative to their sales. Furthermore, companies with high sales per employee might be more agile, able to adapt more quickly to new technologies and market changes. This agility could be a crucial factor in their ability to expand operations and secure higher market shares, as it enables them to outpace competitors. Overall, these insights highlight productivity's pivotal role in determining market share within the software development industry.

Profitability and Market Share

In our fixed effects linear regression model, the correlation between EBITDA margin and the logarithm of market share showed significant variation across the sector. For *very small* and *small* firms, there is a notable positive correlation between EBITDA margin and market share. In contrast, for *medium-sized* firms, this relationship becomes negative and statistically insignificant, and for *large* firms, it remains positive but non-significant. These results suggest that the positive link between profitability and market share present for *very small* and *small* firms is not uniform for all company size segments, making generalizations difficult. These findings also highlight the diverse nature of profitability within the sector, which seems to vary significantly, regardless of company size. While EBITDA margin contributes to our model's explanatory power, its influence is less pronounced than the productivity metric. In summary, the data suggests that profitability is a more critical factor in terms of market share for smaller firms.

R&D spend per Sale and Market Share

In our fixed effects analysis of the software development sector, we encountered an unexpected finding: there was no significant correlation between R&D spending per sale and the logarithm of market share. This result is particularly surprising given the prevailing assumptions in an innovation-centric sector like software development. As noted by Ahlawat et al., (2022), variations in relative R&D expenditure were present across different market sizes globally, reflecting the industry's emphasis on innovation and development. However, the lack of a clear pattern in our analysis suggests a more complex picture of the Norwegian landscape. The lack of a common trend indicates that the relative investment in R&D does not consistently correlate with market share, irrespective of company size. It also highlights the diverse approaches to R&D investment across companies of varying sizes, underscoring the sector's heterogeneous nature.

7.2.4 Summary

In addressing the second research question about the traits of growing companies, we found that firms with relatively high market shares and those demonstrating high profitability are most likely to experience growth in market share, a trend consistent across the sector. The impact of R&D spending per sale on market share growth, however, remains ambiguous, as for its relationship with market share. The predominant characteristic of companies holding large market shares is elevated productivity, as indicated by higher sales per employee. This is followed by the importance of high profitability, specifically in the smaller-sized firms' segments.

7.3 Research Question 3

The first two research questions have been answered through analyzing financial data in the sector. Research Question 3 seeks to understand underlying conditions that might drive the trends and characteristics examined in the previous research questions, to get a more nuanced understanding of the sector. Exploring the primary data gathered through questionnaires provides valuable insights to address the research question.

Research Question 3

What are the most common offerings and preferences regarding software solutions, and what are the perceptions and anticipated future trends in this sector and market?

7.3.1 Service Offerings

To explore Research Question 3, we conducted a dual-focused data collection involving both Oslo-based software solution suppliers (*vendors*) and their Oslo-based potential customers (*clientele*). Our findings indicate a general satisfaction among *clientele* with their existing internal and/or external software solutions. Notably, *clientele* using external (or internal) software are more inclined to invest in similar solutions in the future. Traditionally, external software acquisition favored hiring internal resources and consultants, while internal software relied on white label and SaaS solutions. However, a shift is emerging, revealing a growing preference for white label solutions in external software acquisitions and an increased focus on hiring internal resources for internal software needs. This trend suggests a dynamic shift in acquisition patterns, a development also likely to be affected by AI's anticipated impact on future acquisition strategies.

We notice that roughly 2/3 of software development companies in Oslo offer SaaS-solutions, more than 1/2 deliver consultant services and only 1/3 offer white label solutions. However, white label solutions have, by the *clientele*, been the most used acquisition method for internal software, and the most sought-after method for future external software acquisitions. Additionally, the demand for both consultant services and SaaS-services seem to drop, based on what methods the *clientele* plan to use for future acquisitions. Thus, there appears to be a mismatch in what the Oslo's *vendors* offer, and what the *clientele* wants.

This discrepancy in market dynamics can potentially be attributed to two theories. Firstly, there's a possibility that white label solutions, while prevalent, are not significantly profitable for most *vendors*, except for a few market leaders. This economic dynamic could deter many *vendors* from engaging in this segment of the service market. Secondly, it's plausible that the *clientele*'s demand is not exclusively met by local *vendors* but is supplemented by international and other regional providers. These factors collectively could explain the observed market mismatch.

Furthermore, low-code/no-code solutions appear to have affected the operations of software development companies in Oslo (*vendors*). Notably, 38% report that the advent of these solutions has influenced their business operations, with the impact exclusively being positive. This could imply a dual utility of low-code/no-code solutions; that they serve not only as (1) products offered to customers, but also as (2) internal tools that streamline repetitive and standardized tasks within companies (Sap, n.d.). This shift suggests that while general tasks are becoming more standardized, firms could potentially channel their efforts into more specialized areas of expertise. A possible outcome of such efforts could be innovation and/or changes within service offerings.

7.3.2 Preferences

While the *clientele* appears to favor white label solutions, they also highlight flexibility and customization possibilities as their number one priority, and low price as the least important priority, when acquiring both internal and external software. White label solutions, however, tend to be standardized and cheap, not offering flexibility and customizability (Silva et al., 2020). This proposes an indication of a paradox, as the *clientele* gravitate towards acquisition methods that typically don’t satisfy their desires. On the contrary, most of the *clientele* generally appears satisfied with their current software. This could, nevertheless, suggest that there is some kind of gap between what the *clientele* wants, and what they are able to get. Continuing this trail of thought, one could argue that the ideal acquisition method for companies acquiring external and internal software are “out-of-the-box” solutions with the possibilities for customization and flexibility. This does, however, highlight the diverse and complex dynamics of customer preferences and decision-making regarding investing in software services.

7.3.3 Perceptions and Anticipations

The preferences of the *clientele* when acquiring software, and their perceived preferences by *vendors*, mostly align. The vendors seem to correctly identify flexibility and customizability as the clientele’s top priority. However, the second most prominent priority for *clientele* is scalability, a preference the *vendors* do not seem to perceive correctly. Rather, the *vendors* seem to overestimate the importance of continuous support and maintenance, good communication and quick delivery. It is, however, fair to assume that the expected changes in future processes for acquiring software solutions also might affect preferences.

Furthermore, there appears to be a strong anticipation among *clientele* that artificial intelligence will significantly impact how they acquire internal and external software. However, purchasing from a less reputable distributor is perceived as a moderate-to-high risk by most companies. This could imply that more reputable *vendors* could benefit more from the technological change anticipated by AI. However, this assumes that they manage to adopt the technology in an efficient and timely manner.

Among *vendors*, a significant trend of innovative pressure is evident. Nearly half of the *vendors* are already facing competitive pressures from new, disruptive companies and business models, indicating a shift in the market landscape. 60% of *vendors* recognize a high potential for disruptive changes in the near future, with almost universal expectations of substantial market disruption in the medium to long term. This shift is expected to result in a market with a more diverse range of competitors and intensified competition.

In response to this pressure, 70% of *vendors* are increasing their investments in research and development (R&D). However, the descriptive statistics in Section 5.2.2, which compiles data over the past decade, reveals a noticeable discrepancy. Despite the reported increase in R&D investment, only a few companies, spanning various sizes, are actually dedicating substantial funds to R&D. This raises questions about the accuracy of R&D expenditures reported in financial statements, as they may not truly reflect the actual investments in innovative activities. Moreover, the *vendor* sample examined may not completely represent the entire sector, or they could exaggerate their R&D investments.

Nonetheless, Figure 8 from Section 4.2 presents an encouraging sign. It shows an increase in average CPI-adjusted R&D expenditures from 2017 to 2020. When compared with the responses from the questionnaire, a trend of escalating innovative activities becomes apparent. This suggests a recent surge in R&D investments, potentially aligning earlier data with current perceptions of innovation among *vendors* in the sector.

Regarding the market outlook, the perception of the *vendors* is predominantly optimistic. Most companies consider the market to have grown during the past decade, and anticipate this trend to either continue or increase towards 2030. This fits well with the recent market trends discovered in Chapter 4. The alignment of sector perceptions and descriptive data enhances the credibility of the primary data.

As evident in Figure 6 in Section 4.2, there has been a stable level of weighted average EBITDA margins in the last decade at about 6-8%, suggesting a trend of continued profitability in the sector. Most of the *vendors* expect profitability towards 2030 to remain at the current levels or increase. There does not, however, seem to be a clear tendency towards increased spending on software solutions among the *clientele*. Given the intensifying competition and *clientele* indications of maintaining their current spending levels, the sector's future profitability may be impacted. Consequently, for the 31% of *vendors* anticipating increased profitability, it may become necessary to focus on either reducing costs or increasing the volume of sales. Innovative solutions and new technology, such as AI and no code/low code, could, however, benefit the *vendors* in this regard, potentially increasing the firms' productivity and hence facilitating opportunities for sales growth or cost cuts.

7.3.4 Summary

In addressing the research question concerning conventional offerings, preferences, and anticipated future trends in the software development market, our analysis uncovers a complex and evolving landscape. A key trend identified is a shift in preferences towards white label solutions, particularly for external software acquisitions. This shift indicates a departure from traditional acquisition methods like hiring internal resources or consultants. Due to this inclination towards white label solutions, there exists a paradox, since most customers prioritize flexibility and customization, traits not typically associated with standardized white label offerings. This discrepancy suggests a potential market gap, signifying an opportunity for solutions that blend white label's cost-effectiveness with customization and flexibility.

The sector is under strong innovative pressure, with many companies expecting disruptive changes and increased competition. This anticipation is driving investments in R&D and other innovative practices, even though traditional financial metrics may not fully capture these innovation efforts. The optimistic market outlook, with expectations of growth and profitability, is tempered by the recognition of increasing competition and the need for strategic adaptations. This dynamic market, characterized by evolving preferences and technological advancements, presents a landscape where companies must be agile and responsive to maintain relevance and profitability in the coming years.

Chapter 8 - Conclusion & Future Research

8.1 Conclusion

This thesis reveals the Norwegian software development sector and market as dynamic and evolving. Marked by significant growth, the sector has seen a notable increase in companies, employees, and sales, coupled with changes in market structure and company dynamics.

Three key questions guided the research, each uncovering distinct aspects of the market's transformation, to address the following overarching question:

Problem Statement

What are the recent and expected characteristics of the Norwegian software development sector and market?

The first research question highlighted a considerable rise in total sales and number of companies over the past decades, increased productivity, stable profitability, and increased spending on innovative activities, paralleled by market fragmentation and decentralization. The second research question revealed that higher grossing firms are more productive, and that higher grossing firms and more profitable firms are more likely to grow their market share. Interestingly, the link between R&D spending and both market share and the likelihood of market share growth was complex and ambiguous. The third research question investigated market offerings, preferences, and possible future trends, identifying an inclination towards white label solutions, despite a demand for customization and flexibility. This highlights a market gap and an opportunity for innovative, adaptable solutions. The sector anticipates disruptive changes, prompting strategic adaptations and investments in R&D.

In conclusion, the Norwegian software development market is characterized by strong growth and evolving market dynamics. The sector faces a future where agility and innovation are crucial for success. This thesis offers insights into the interplay of market growth and company dynamics, providing a comprehensive overview.

8.2 Future Research Directions

Our comprehensive analysis of the Norwegian software development market opens several pathways for further exploration, each promising to deepen the understanding of this diverse and complex sector.

Segment-Specific Analysis

The diversity of the Norwegian software development sector suggests that focused studies on specific market segments could reveal interesting insights. This approach could unveil nuanced trends in areas such as small-scale startups, large corporations, or niche service providers, like those offering solutions to for example the construction industry or oil sector. Such targeted research could provide a deeper understanding of the unique dynamics within these segments.

Technology-Specific Studies

Investigating specific technologies, such as AI, low code/no code, or a deep dive into specific service offerings, presents another promising area for research. Detailed analysis of these areas would offer a closer look at the intricacies of their importance within the Norwegian software development market.

Exploring Innovation

Exploring how major Norwegian software firms navigate disruptive innovation through the lens of established theories such as Christensen's (1997) The Innovator's Dilemma, could offer intriguing perspectives. Delving into case studies or detailed analyses of a specific company's approach to innovative challenges can provide valuable insights into their strategic planning and adaptive tactics, as well as the future composition of the market.

References

Advania. (2023). *Hvilke IT-tjenester trenger du?* Retrieved from Advania: https://www.advania.no/kunnskap/artikler/it-tjenester?fbclid=IwAR2S6flfTIRzv_SISq0nzYGAcycbMJCP7G0oDSoy3Bclt7dAeqwXOFDrKrE

Ahlawat, P., O'Niell, C., Mohan, A., & Wiles, T. (2022). *Solving Software's R&D Conundrum*. Retrieved from BCG: https://www.bcg.com/publications/2022/software-companies-tackling-research-and-development-conundrum

Akbar, R., & Safdar, S. (2015). A short review of Global Software Development (GSD) and latest software development trends. *2015 International Conference on Computer, Communications, and Control Technology (I4CT)* (pp. 314-317). Kuching, Malaysia: IEEE. doi:https://doi.org/10.1109/I4CT.2015.7219588

Argolini, R., Pellegrinelli, S., Burkacky, O., Kapellos, C., Deichmann, J., Johnston, S., . . . Bode, C. (2022). *Reaching excellence in software procurement*. Retrieved from McKinsey & Company: https://www.mckinsey.com/capabilities/operations/our-insights/reaching-excellence-in-software-procurement

AxiomQ. (2022). *Most Important Aspects Involved in Software Development*. Retrieved December 2023, from AxiomQ: https://axiomq.com/blog/most-important-aspects-involved-in-software-development/

Bartlett, J. (2014). *The Hosmer-Lemeshow goodness of fit test for logistic regression*. Retrieved from The Stats Geek: https://thestatsgeek.com/2014/02/16/the-hosmer-lemeshow-goodness-of-fit-test-for-logistic-regression/

Baruch, Y., & Holtom, B. C. (2008). Survey Response Rate Levels and Trends in Organizational Research. *Human Relations, 61*(8), 1139-1160. doi:10.1177/0018726708094863

Berli, H. L., & Hundhammer, A. (2020). Bergen: Norges Handelshøyskole. Retrieved from https://openaccess.nhh.no/nhh-xmlui/handle/11250/2678751

Bevans, R. (2023). *Akaike Information Criterion | When & How to Use It (Example)*. Retrieved from Scribbr: https://www.scribbr.com/statistics/akaike-information-criterion/

Brønnøysundsregistrene. (2023). *Industrial codes*. Retrieved from The Brønnøysund Register Centre: https://www.brreg.no/en/business-2/industrial-codes/?nocache=1702372816521

Christensen, C. M. (1997). *The innovators Dilemma when new technologies cause great firms to fail.* Harvard Business Review.

Croissant, Y., & Millo, G. (n.d.). *Panel data econometrics in R:*. Retrieved from cran.r-project: https://cran.r-project.org/web/packages/plm/vignettes/A_plmPackage.html

Dallhammer, E., Derszniak-Noirjean, M., Gaugitsch, R., Hans, S., Zillmer, S., Gaupp-Berghausen, M., . . . Lüer, C. (2019). *The impacts of metropolitan regions on their surrounding areas : local and regional dimension.* doi:10.2863/35077

Dey, V. (2021). *Understanding the AUC-ROC Curve in Machine Learning Classification*. Retrieved from Analytics India Mag: https://analyticsindiamag.com/understanding-the-auc-roc-curve-in-machine-learning-classification/

Diansyah, A. F., Rahman, M. R., Handayani, R., Cahyo, D. D., & Utami, E. (2023). Comparative Analysis of Software Development Lifecycle Methods in Software Development: A Systematic Literature Review. *International Journal of Advances in Data and Information Systems*, 97-106. doi:10.25008/ijadis.v4i2.1295

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, a. O. (2008). *An R^2 Statistic for Fixed Effects in the Linear Mixed Model*. Retrieved from National Library of Medicine: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2587505/

Frost, J. (n.d.). *Making Predictions with Regression Analysis*. Retrieved from Statistics By Jim: https://statisticsbyjim.com/regression/predictions-regression/

Golding, J. L. (2023). *The Biggest Tech Hubs in the World & How They're Changing*. Retrieved from techpoint: https://techpoint.org/the-biggest-tech-hubs-in-the-world/

Guzel, B., & Wei, B. (2023). *Startups must embrace sustainable scaling strategies*. Retrieved from techcrunch: https://techcrunch.com/2023/12/15/startups-must-embrace-sustainable-scaling-strategies/

Hahs-Vaughn, D. L. (2023). *Pearson Correlation Coefficient*. Retrieved from ScienceDirect: https://www.sciencedirect.com/topics/social-sciences/pearson-correlation-coefficient?fbclid=IwAR1M34fSlQkSlkSKa3JgLNT8SZyShzZ8vpX6cT2l7NrGZzlTzdKYv50TWcY

Hassan, M. (2023a). *Regression Analysis – Methods, Types and Examples*. Retrieved from researchmethod.net: https://researchmethod.net/regression-analysis/

Hassan, M. (2023b). *Reliability – Types, Examples and Guide*. Retrieved from researchmethod.net: https://researchmethod.net/reliability/

Hassan, M. (2023c). *Construct Validity – Types, Threats and Examples*. Retrieved from researchmethod.net: https://researchmethod.net/construct-validity/

Hassan, M. (2023d). *External Validity – Threats, Examples and Types*. Retrieved from researchmethod.net: https://researchmethod.net/external-validity/

Holliday, M. (2021). *9 Productivity Metrics and KPIs That Matter*. Retrieved from Oracle Netsuite: https://www.netsuite.com/portal/resource/articles/human-resources/productivity-metrics.shtml

Huntington-Klein, N. (2023). *Chapter 16 - Fixed Effects*. Retrieved from The Effect Book: https://theeffectbook.net/ch-FixedEffects.html

IBM. (n.d.). *Think: What is software development?* Retrieved from IBM Corporation Web site: https://www.ibm.com/topics/software-development

Irascu, D. (2023). *The Norwegian ICT Industry: Tech Companies & Performance*. Retrieved from Tech Behemoths: https://techbehemoths.com/blog/the-norwegian-ict-industry-in-2021-tech-companies-performance

Javeed, F., Shahzad, B., Shaikh, A., Reshan, M. S., Ahmad, S., Alshahrani, H., & Rajab, K. (2022). Strengthening Software Make vs. Buy Decision: A Mixed-Method Approach. *Intelligent Automation and Soft Computing*, 193-205. doi:10.32604/iasc.2022.021769

Jørgenrud, M. B. (2010). *Atea kjøper Umoe IKT*. Retrieved from digi: https://www.digi.no/artikler/atea-kjoper-umoe-ikt/291647?fbclid=IwAR3eOpRsbfT83N35n-YF1Uq9BaHMqW2T0QCu3zMEiT0Xx9VHiIjiraMKwqc

Krämer, W. (2014). *Durbin–Watson Test.* Springer, Berlin, Heidelberg. doi:https://doi.org/10.1007/978-3-642-04898-2_219

Kuiken, S. V. (2022). *Tech at the edge: Trends reshaping the future of IT and business*. Retrieved December 2023, from McKinsey: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-at-the-edge-trends-reshaping-the-future-of-it-and-business

Kulkarni, G., Gambhir, J., & Palwe, R. (2012). Cloud Computing-Software as Service. *International Journal of Cloud Computing and Services Science*, 11-16. Retrieved from https://www.researchgate.net/publication/274481004_Cloud_Computing-Software_as_Service

Kumar, A., Grover, P., Kar, A. K., & Pani, A. K. (2017). IT Consulting: A Systematic Literature Review. *Digital Nations – Smart Cities, Innovation, and Sustainability* (pp. 474-484). Springer, Cham. doi:10.1007/978-3-319-68557-1_42

Laato, S., Mäntymäki, M., Islam, A. K., Hyrynsalmi, S., & Birkstedt, T. (2022). Trends and Trajectories in the Software Industry: implications for the future of work. *Information Systems Frontiers*, 929–944. doi:https://doi.org/10.1007/s10796-022-10267-4

LaBerge, L., O'Toole, C., Schneider, J., & Smaje, K. (2020). *How COVID-19 has pushed companies over the technology tipping point—and transformed business forever*. Retrieved from McKinsey & Company: https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/how-covid-19-has-pushed-companies-over-the-technology-tipping-point-and-transformed-business-forever

Lebo, H. (2019). *The Internet At 50: How the Dot-Com Bubble Burst*. Retrieved from HNN News Network: https://historynewsnetwork.org/article/173359

Lundh, Y. (2018). A Slice of Norway's Computing History. *IEEE Annals of the History of Computing* (pp. 62-72). IEEE Computer Society. doi:https://doi.org/10.1007/978-3-642-23315-9_28

Malato, G. (2023). *An Introduction to the Shapiro-Wilk Test for Normality*. Retrieved from builtin: https://builtin.com/data-science/shapiro-wilk-test

Mansournia, M. A., Nazemipour, M., Naimi, A. I., Collins, G. S., & Campbell, M. J. (2020). *Reflection on modern methods: demystifying robust standard errors for epidemiologists.* International Journal of Epidemiology. doi:https://doi.org/10.1093/ije/dyaa260

Mjøs, A., & Selle, S. F. (2022). *Regnskapsdatabasen - Norwegian Corporate Accounts.* Bergen: Centre for Applied Research at NHH.*Regnskapsdatabasen - Norwegian Corporate Accounts.* Bergen: Centre for Applied Research at NHH. Retrieved from https://snf.no/media/durjpedn/r04_22.pdf

Moayed, V. (2021). *The Rise Of Global Tech Entrepreneurship*. Retrieved from Forbes: https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2021/07/12/the-rise-of-global-tech-entrepreneurship/?fbclid=IwAR2_UKS643jUJCUIT2s0MpotIwAK_nGcXBAYk6zqWX11w7utEP-7wo63DA8

Murgia, M. (2023). *Transformers: the Google scientists who pioneered an AI revolution*. Retrieved from Financial Times: https://www.ft.com/content/37bb01af-ee46-4483-

982f-ef3921436a50?fbclid=IwAR1VI6G9EYh-z_GD7WP8rjd-6HxZGeulnv0QKKYYNmkoXHD3X-6MOVKjFQQ

Oberoi, A. (2021). *Why Low-code or No-code is more Distruptive Now than Ever*. Retrieved from daffodil: https://insights.daffodilsw.com/blog/why-low-code-or-no-code-is-more-disruptive-now-than-ever

Osterweil, L. J. (2018). What is software? In *The Essence of Software Engineering* (pp. 59-76). Springer, Cham. doi:https://doi.org/10.1007/978-3-319-73897-0_4

Peterdy, K. (n.d.). *Pestel Analysis*. Retrieved from CFI: https://corporatefinanceinstitute.com/resources/management/pestel-analysis/

Pinzon, E. (2022). *Heteroskedasticity robust standard errors: Some practical considerations*. Retrieved from The State Blog: https://blog.stata.com/2022/10/06/heteroskedasticity-robust-standard-errors-some-practical-considerations/

Porter, M. E. (2008). *The Five Competitive Forces That Shape Strategy*. Retrieved from Harvard Business School: https://www.hbs.edu/faculty/Pages/item.aspx?num=34522&fbclid=IwAR1nNapWl21erTKYDYJRJ9VuEm1h_cCticWHAVeaAGouRxbbnJhRUHobFIE

Proff Forvalt. (n.d.). *Om Proff Forvalt*. Retrieved from Proff Forvalt: https://forvalt.no/Om/om-proff-forvalt/om-forvalt?fbclid=IwAR1Xw8QyLLi1J2eCsCoe8OPnNQUCyyGrufSndIvb6Cu4thNzXC6RcSoxmdg

Quirt, B. (2022). *SMBs: The next growth opportunity for high tech*. Retrieved from Accenture: https://www.accenture.com/us-en/blogs/high-tech/smbs-the-next-growth-opportunity-for-high-tech

Regjeringen. (2019). *SMB-Definisjon.* Retrieved from Regjeringen: https://www.regjeringen.no/contentassets/1bdef920d6cc4245b2e09a5902cb4374/vedlegg1.pdf?fbclid=IwAR2uofzETMQhilfW4R22JslGPrRyA2SezCXhWeGfxb0Xp_ZoYh04mUqmnzI

Riveros, J. M. (n.d.). *Wooldridge Serial Correlation Test for Panel Data using Stata.* Retrieved from MS Research Hub: https://blog.ms-researchhub.com/2021/02/24/wooldridge-serial-correlation-test-for-panel-data-using-stata/

Rosenberg, S. (2018). *How the Great Recession teed off tech's long boom*. Retrieved from Axios: https://www.axios.com/2018/09/15/dotcom-bust-tech-growth-great-recession?fbclid=IwAR0O0s_MOVcm4BIbdAAo_cDE1sCqF9Y3McoRdx2LKOIKVTvC5KEHAFgwrv0

Sadowski, M., & Naborshchikov, S. (n.d.). *In-House vs. Outsourcing Software Development*. Retrieved from Mobile Reality: https://themobilereality.com/blog/in-house-vs-outsourcing-software-development

Sap. (n.d.). *What is low-code/no-code application development?* Retrieved from Sap: https://www.sap.com/products/technology-platform/low-code/what-is-low-code-no-code.html

Saunders, M. N., Lewis, P., Adrian, & Thornhill. (2019). *Research Methods For Business Students.* Harlow: Pearson Education Limited.

Schutz, M. (2018). *Software as a Service vs Software as a Product*. Retrieved December 2023, from Bynder: https://www.bynder.com/en/blog/software-as-a-product-vs-software-as-a-service/

Sikt. (n.d.). *Sikt – Norwegian Agency for Shared Services in Education and Research*. Retrieved from Sikt: https://sikt.no/en/about-sikt

Silva, F., Machado, I., & Souza, R. M. (2020). Taming and Unveiling Software Reuse opportunities through White Label Software in Startups. *Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* (pp. 302-305). Portoroz, Slovenia: IEEE. doi:10.1109/SEAA51224.2020.00057

Soetewey, A. (2020). *Outliers detection in R*. Retrieved from Stats and R: https://statsandr.com/blog/outliers-detection-in-r/

Sparebank 1. (n.d.). *AS eller ENK - hva skal du velge?* Retrieved from Sparebank1: https://www.sparebank1.no/nb/bank/bedrift/tips-og-rad/starte-bedrift/forskjell-aksjeselskap-enkeltpersonforetak.html

SSB. (2009). *Standard for næringsgruppering (SN)*. Retrieved from Statistisk Sentralbyrå: https://www.ssb.no/klass/klassifikasjoner/6?fbclid=IwAR1ZF7y4yf2MdFuV2-t_mR3G97cfEtBl1LD6FnG_wnoG07DySYVey8N6iyw

SSB. (2023). *Konsumprisindeksen*. Retrieved from Statistisk Sentralbyrå: https://www.ssb.no/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen

Stammann, A., Heiss, F., & McFadden, D. (2016). *Estimating Fixed Effects Logit Models with Large Panel Data.* Retrieved from ECONSTOR: https://www.econstor.eu/bitstream/10419/145837/1/VfS_2016_pid_6909.pdf

Steine, T. O. (2010). *The Founding, Fantastic Growth, and Fast Decline of Norsk Data AS.* doi:https://doi.org/10.1007/978-3-642-23315-9_28

Taboga, M. (2021). *Hetroskedasticity-robust standard errors*. Retrieved from StatLect: https://www.statlect.com/glossary/heteroskedasticity-robust-standard-errors

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). *Attention Is All You Need*. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Wahab, A. A., & San, T. P. (2018). A Systematic Literature Review on Risk Factors in Software Development Outsourcing. *International journal of engineering & technology*, 76-81. doi:10.14419/ijet.v7i3.20.18984

Westberg, P. (2023). *EBITDA Margin: A Common Profitability Metric*. Retrieved from quatr: https://quartr.com/insights/investing/ebitda-margin-a-common-profitability-metric

Wojtecka, M. (2023). *Navigating Norwegian Accounting Standards: Your Comprehensive Guide to Gaap and Reporting Requirements*. Retrieved from NOVUM: https://www.novum.no/en/blog/navigating

Zach. (2020). *The Breusch-Pagan Test: Definition & Example*. Retrieved from Statology: https://www.statology.org/breusch-pagan-test/

Zach. (2022). *How to Calculate Robust Standard Errors in R*. Retrieved from Statology: https://www.statology.org/robust-standard-errors-in-r/

Zebrabi. (2023). *Research and Development (R&D) Expense Ratio*. Retrieved from Zebrabi: https://zebrabi.com/guide/research-and-development-rd-expense-ratio/

Appendix

A.1 Variable description

Variable label	Description	Data type	Use case(s)
Logarithm of Market Share	Logarithm of a company's sales in relation to the sum of sales from all companies in the sector Calculated as: $\log\left(\frac{Sales\ of\ the\ given\ firms}{Sum\ of\ sales\ from\ all\ firms}\right)$	Numeric	Response (Linear) Explanatory (Logistic)
Market Share Growth	A binary variable (0, 1) where 1 indicates growth in market share between current year and last year, and 0 indicates no growth	Logical	Response (Logistic)
EBITDA Margin	A company's earnings before interest, taxes depreciation and amortization in relation to its total income Calculated as: $\frac{Ebitda}{Total\ income}$ (Mjøs & Selle, 2020)	Numeric	Explanatory (Linear) Explanatory (Logistic)
R&D per Sale	A company's research and development (R&D) expenses in relation to its sales Calculated as: $\frac{R\&D\ Expenses}{Sales}$	Numeric	Explanatory (Linear) Explanatory (Logistic)
Log of Sales per Employee	The logarithm of a company's sales per employee, measuring productivity Calculated as: $\log\left(\frac{Sales}{Employees}\right)$	Numeric	Explanatory (Linear)
Cash per Sale	A company's cash holdings in relation to its sales Calculated as: $\frac{Cash\ holdings}{Sales}$	Numeric	Explanatory (Linear) Explanatory (Logistic)
Age of the Company	How old a company is. Calculated as: $Current\ year - Year\ founded$	Numeric	Explanatory (Linear) Explanatory (Logistic)
Ownership Concentration	Concentration of ownership calculated as Herfindahl Index (HHI). Calculated as: $\sum_{i=1}^{n} s_{ij}^2$ were s_{ij} is the ownership of shareholding i in company j as a fraction (Mjøs & Selle, 2020)	Numeric	Explanatory (Linear) Explanatory (Logistic)
Year	Current year	Categorical	Explanatory (Linear) Explanatory (Logistic)

A.2 Correlation matrices

Very Small Firms

	log_mshare	is_oslo	ebitdamarg	cash_inn	aar	fou_inn	log_inn_per_emp	age	aksj_hhi	mgrowth_binary
log_mshare	1	0.075	0.236	-0.253	-0.217	0.081	0.770	0.007	-0.279	0.123
is_oslo	0.075	1	-0.008	0.011	0.008	-0.021	0.099	-0.024	-0.026	0.010
ebitdamarg	0.236	-0.008	1	0.010	0.007	0.006	0.380	0.024	0.108	0.078
cash_inn	-0.253	0.011	0.010	1	-0.009	-0.065	-0.183	0.131	0.109	-0.068
aar	-0.217	0.008	0.007	-0.009	1	0.021	-0.034	-0.065	0.075	0.050
fou_inn	0.081	-0.021	0.006	-0.065	0.021	1	0.004	-0.008	-0.124	0.038
log_inn_per_emp	0.770	0.099	0.380	-0.183	-0.034	0.004	1	-0.043	-0.017	0.100
age	0.007	-0.024	0.024	0.131	-0.065	-0.008	-0.043	1	-0.016	-0.131
aksj_hhi	-0.279	-0.026	0.108	0.109	0.075	-0.124	-0.017	-0.016	1	-0.056
mgrowth_binary	0.123	0.010	0.078	-0.068	0.050	0.038	0.100	-0.131	-0.056	1

Small Firms

	log_mshare	is_oslo	ebitdamarg	cash_inn	aar	fou_inn	log_inn_per_emp	age	aksj_hhi	mgrowth_binary
log_mshare	1	0.025	0.196	-0.148	-0.299	-0.178	0.754	0.121	0.047	0.011
is_oslo	0.025	1	-0.085	0.069	0.059	-0.044	0.066	-0.113	-0.007	0.051
ebitdamarg	0.196	-0.085	1	-0.091	0.006	-0.123	0.279	0.076	-0.008	0.075
cash_inn	-0.148	0.069	-0.091	1	0.080	0.006	-0.126	-0.010	-0.107	0.027
aar	-0.299	0.059	0.006	0.080	1	0.086	-0.078	0.029	0.058	0.080
fou_inn	-0.178	-0.044	-0.123	0.006	0.086	1	-0.230	-0.021	-0.134	0.009
log_inn_per_emp	0.754	0.066	0.279	-0.126	-0.078	-0.230	1	0.115	0.071	0.037
age	0.121	-0.113	0.076	-0.010	0.029	-0.021	0.115	1	-0.042	-0.207
aksj_hhi	0.047	-0.007	-0.008	-0.107	0.058	-0.134	0.071	-0.042	1	-0.073
mgrowth_binary	0.011	0.051	0.075	0.027	0.080	0.009	0.037	-0.207	-0.073	1

Medium-sized Firms

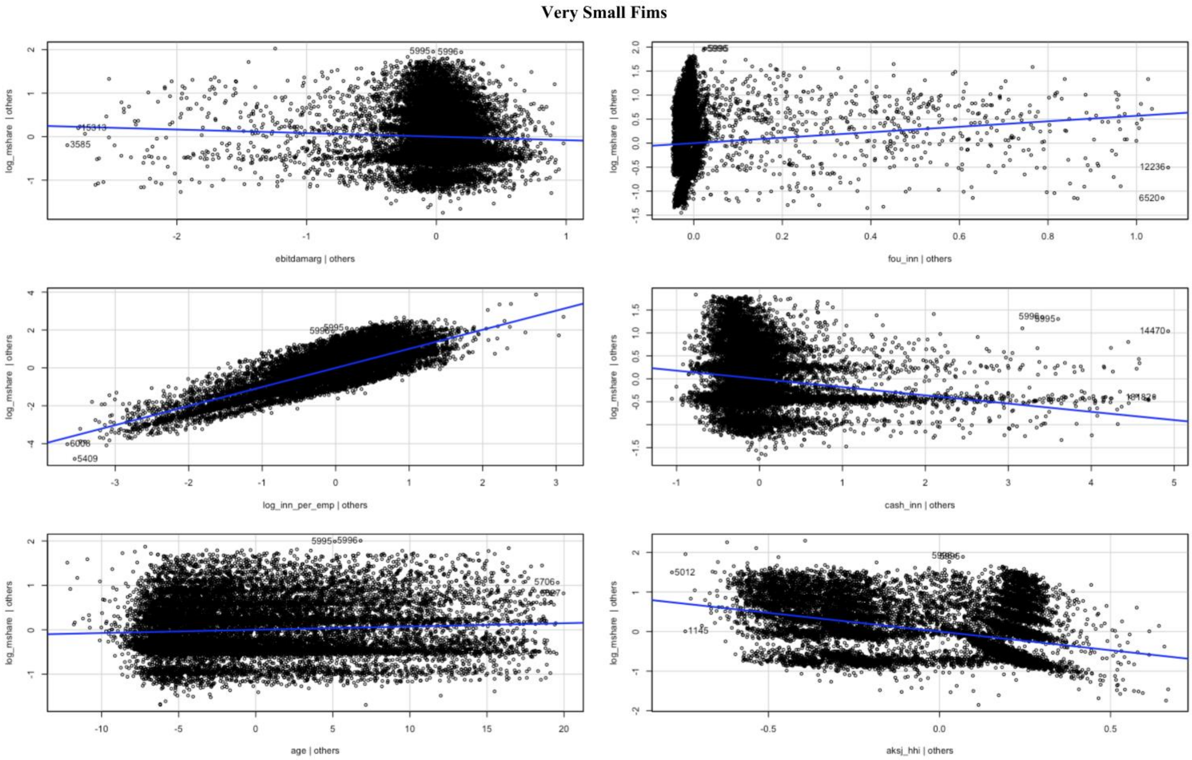
	log_mshare	is_oslo	ebitdamarg	cash_inn	aar	fou_inn	log_inn_per_emp	age	aksj_hhi	mgrowth_binary
log_mshare	1	0.164	-0.039	-0.087	-0.326	-0.020	0.556	0.112	0.300	0.039
is_oslo	0.164	1	0.060	0.144	-0.088	-0.196	-0.069	-0.156	0.103	0.033
ebitdamarg	-0.039	0.060	1	0.286	0.082	-0.047	-0.055	-0.038	0.015	0.082
cash_inn	-0.087	0.144	0.286	1	-0.111	0.002	-0.080	0.011	0.066	-0.007
aar	-0.326	-0.088	0.082	-0.111	1	0.099	0.040	0.125	0.021	0.023
fou_inn	-0.020	-0.196	-0.047	0.002	0.099	1	0.042	0.195	-0.078	-0.011
log_inn_per_emp	0.556	-0.069	-0.055	-0.080	0.040	0.042	1	0.181	0.107	0.067
age	0.112	-0.156	-0.038	0.011	0.125	0.195	0.181	1	0.076	-0.110
aksj_hhi	0.300	0.103	0.015	0.066	0.021	-0.078	0.107	0.076	1	-0.109
mgrowth_binary	0.039	0.033	0.082	-0.007	0.023	-0.011	0.067	-0.110	-0.109	1

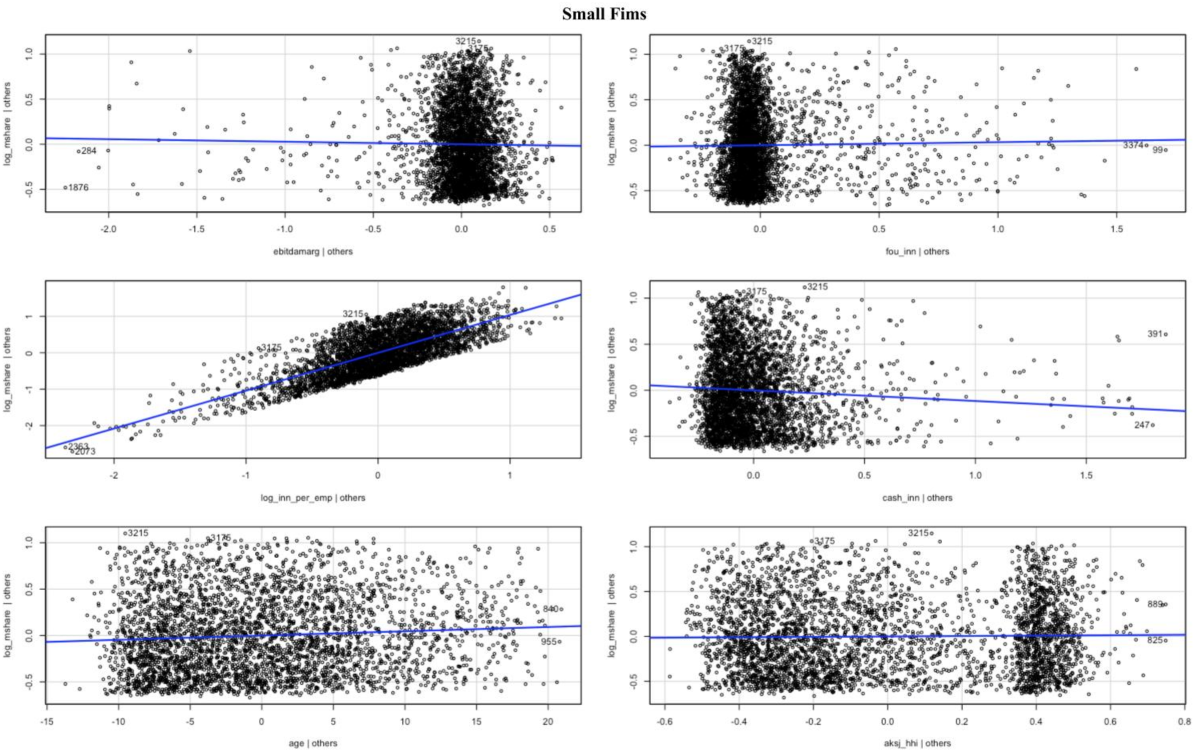
Large Firms

	log_mshare	is_oslo	ebitdamarg	cash_inn	aar	fou_inn	log_inn_per_emp	age	aksj_hhi	mgrowth_binary
log_mshare	1	-0.126	-0.347	-0.085	-0.350	-0.138	0.446	0.361	-0.029	-0.030
is_oslo	-0.126	1	0.374	0.038	0.147	0.188	-0.227	0.100	0.296	0.262
ebitdamarg	-0.347	0.374	1	0.282	0.271	0.144	-0.244	0.014	0.140	0.160
cash_inn	-0.085	0.038	0.282	1	-0.079	0.039	-0.290	0.031	0.103	0.151
aar	-0.350	0.147	0.271	-0.079	1	0.110	-0.073	-0.128	0.198	0.045
fou_inn	-0.138	0.188	0.144	0.039	0.110	1	-0.176	0.234	0.070	0.150
log_inn_per_emp	0.446	-0.227	-0.244	-0.290	-0.073	-0.176	1	0.253	0.079	-0.357
age	0.361	0.100	0.014	0.031	-0.128	0.234	0.253	1	0.066	-0.044
aksj_hhi	-0.029	0.296	0.140	0.103	0.198	0.070	0.079	0.066	1	-0.049
mgrowth_binary	-0.030	0.262	0.160	0.151	0.045	0.150	-0.357	-0.044	-0.049	1

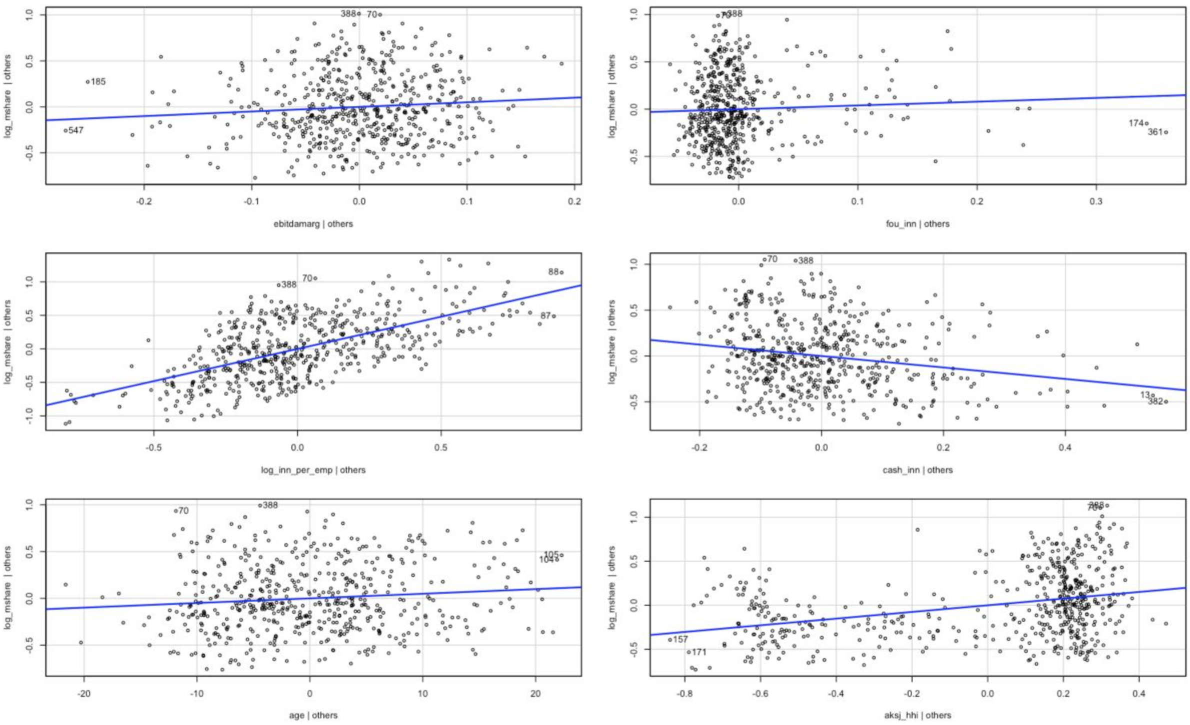
A.3 Linearity plots

A.3.1 Linear Regression

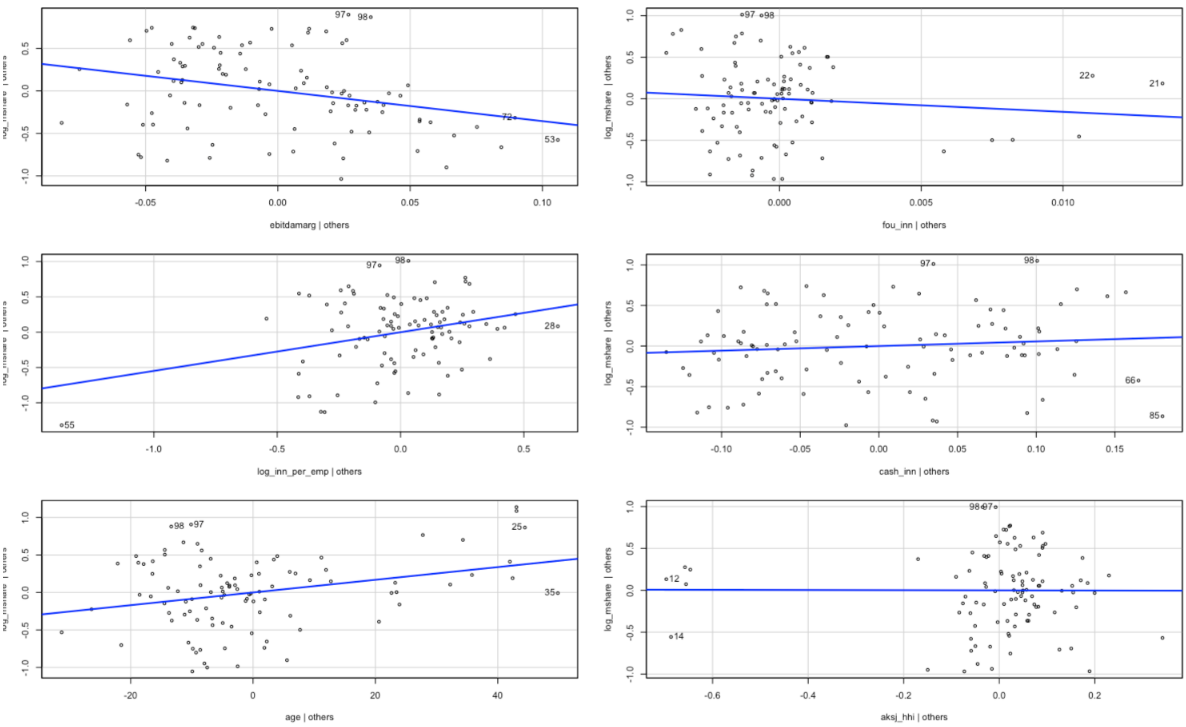




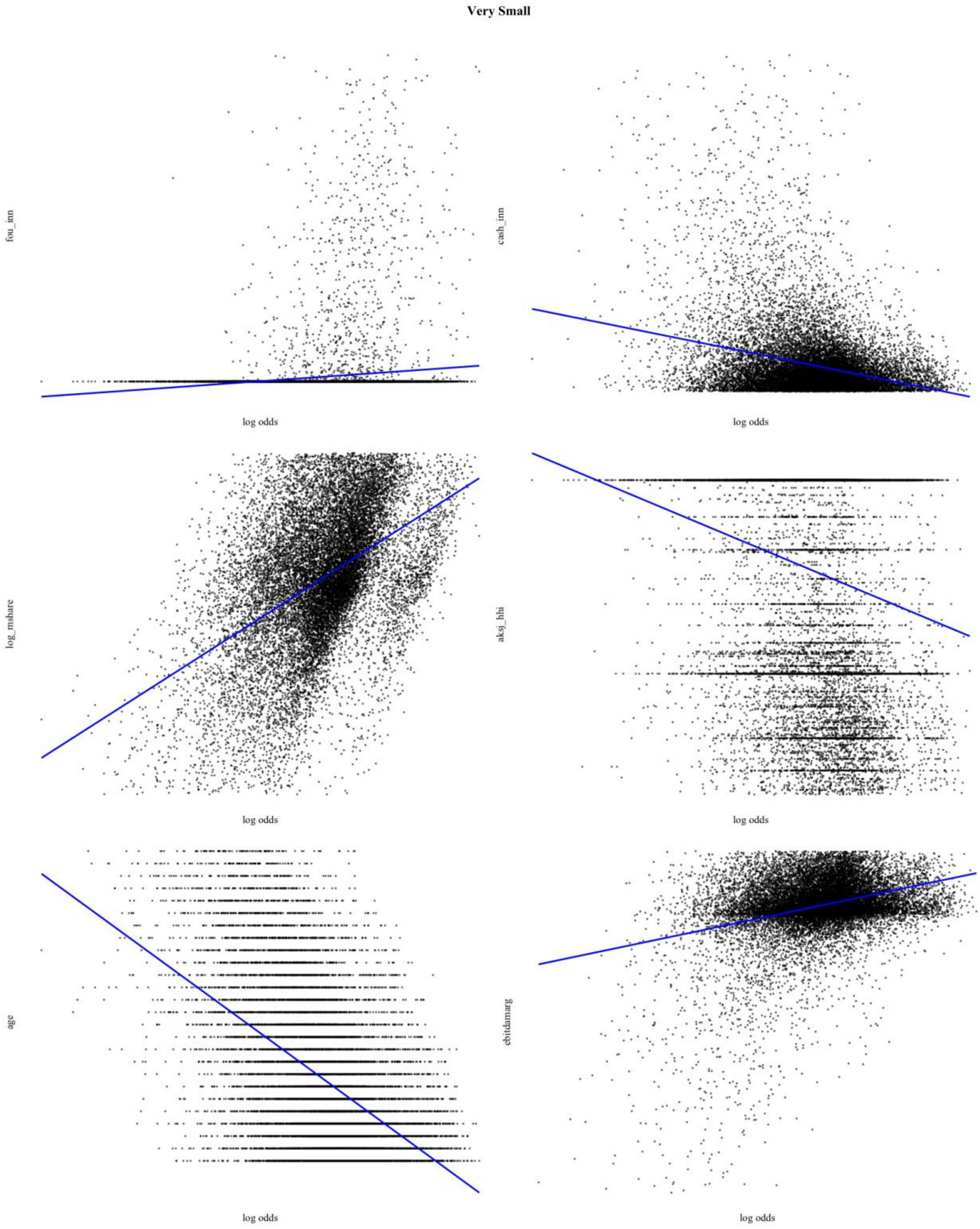
Medium-Sized Firms



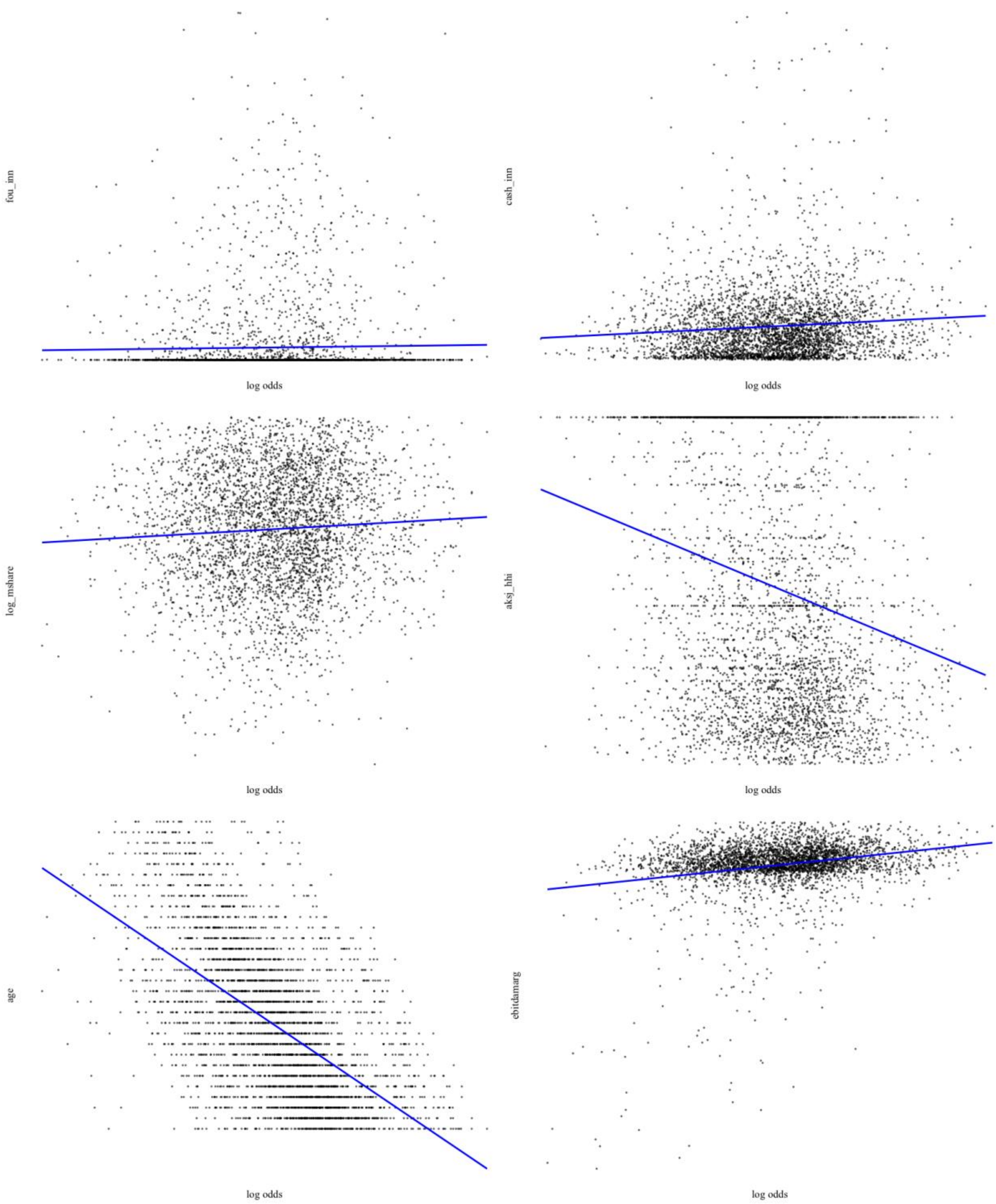
Large Firms



A.3.2 Logistic Regression

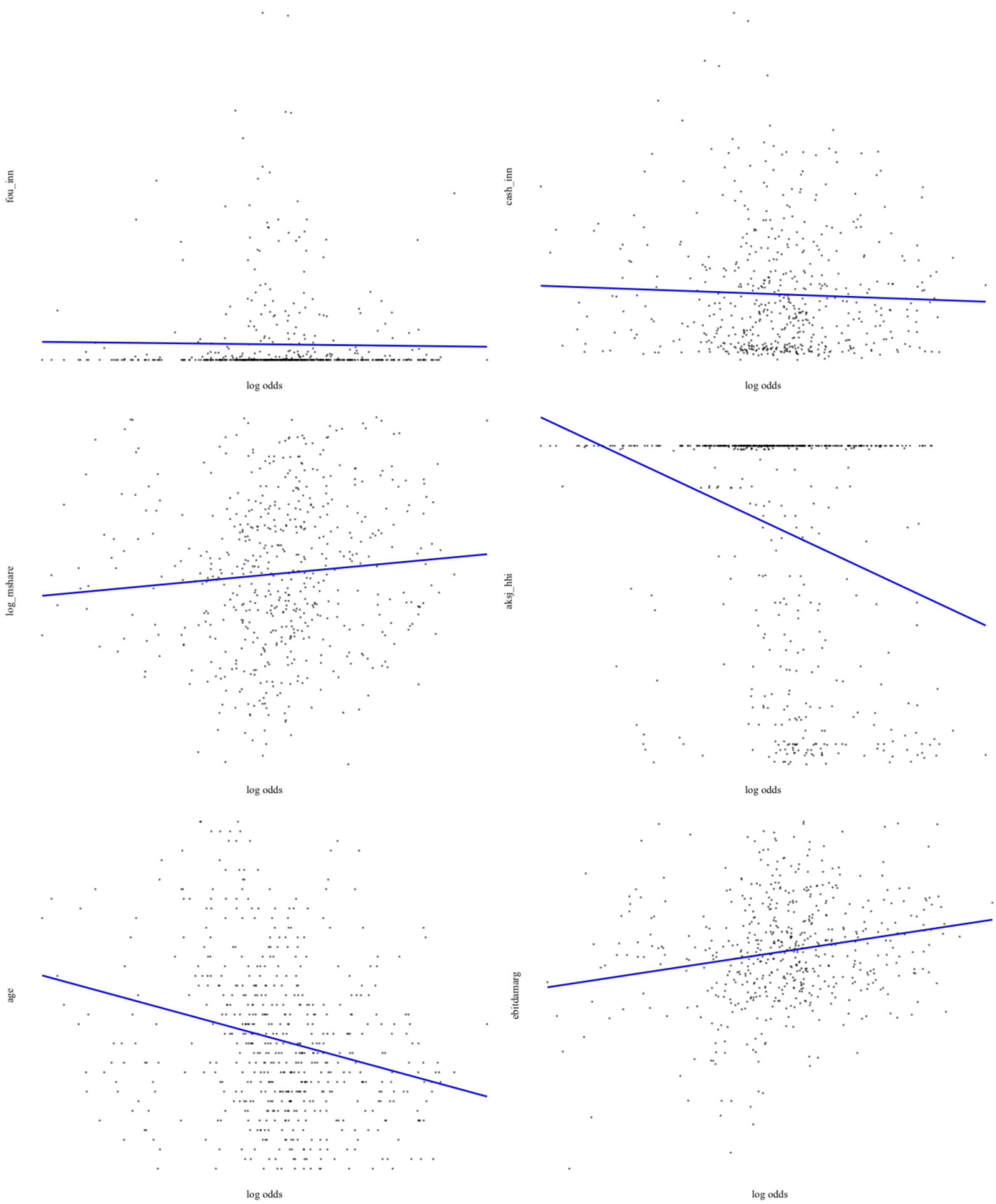


Small

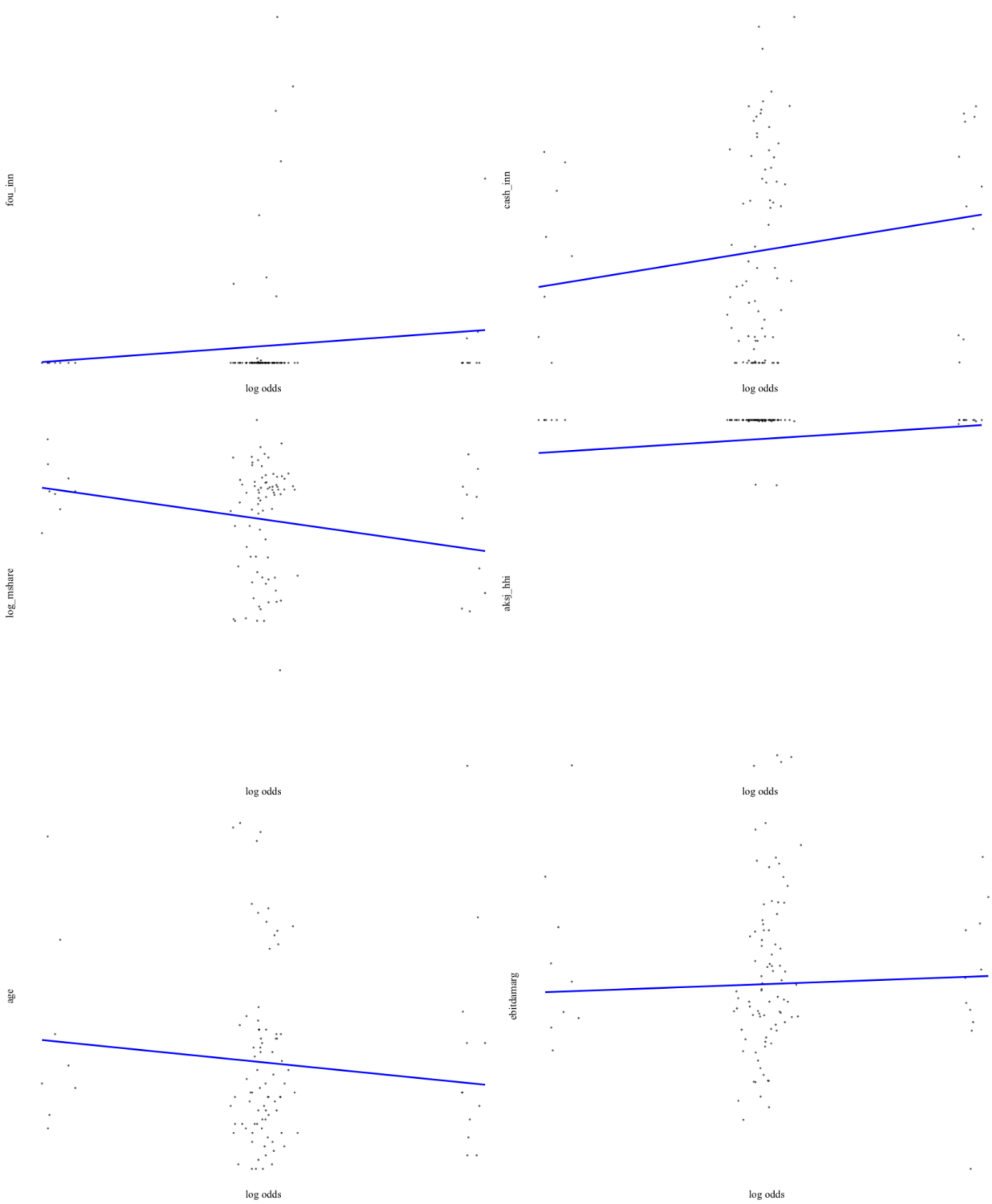


fou_inn
log odds
cash_inn
log odds
log_mshare
log odds
aksj_hhi
log odds
age
log odds
ebitdamarg
log odds

Medium-Sized



Large


fou_inn
log odds
cash_inn
log odds
log_mshare
log odds
aksj_hhi
log odds
age
log odds
ebitdamarg
log odds

A.4 Questionnaires

* Please note that both questionnaires are written in Norwegian

( ) = Radio buttons

[ ] = Text inputs

A.4.1 Vendors

Spørreundersøkelse blant tilbydere av digitale produkter og tjenester

Denne undersøkelsen er utformet av Magnus Nadheim og Halvard Haugse, med formål om å samle inn supplerende data til deres masteroppgave ved Norges Handelshøyskole (NHH) høsten 2023.

Svarene dine vil bli brukt til å kartlegge tilbudet av digitale produkter og tjenester i Norge.

All informasjon du deler med oss blir anonymisert i henholdt til taushetserklæringen på neste side.

Taushetserklæring

Alle svar behandles konfidensielt, forblir anonyme og kan ikke spores tilbake til deg eller din bedrift. Ved å trykke "Jeg godtar" i denne undersøkelsen godtar du at svarene dine brukes for å fremstille data i et samlet format.

Eksempelvis:

- *"80% av bedrifter forventer at veksten i markedet for digitale produkter og tjenester skal være høy frem mot 2030".*
- *"CEOs i små foretak synes å være mer åpne for bruken av kunstig intelligens enn CEOs i store foretak"*

Vennligst informer om du ønsker å delta eller ikke.

( ) Jeg godtar
( ) Jeg godtar ikke

Spørsmål 1
Vennligst fyll inn informasjonen

Bedriftens navn: [ ]
Din stilling i bedriften: [ ]

Spørsmål 2
Er bedriften din en del av et større konsern og/eller en norsk filial av en utenlandsk virksomhet?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 3
Digitale produkter og tjenester refererer i denne sammenhengen til software (programvare)-applikasjoner, verktøy og tjenester, ikke hardware (maskinvare).

Driver bedriften din med - eller hjelper andre bedrifter med - utvikling av digitale produkter og tjenester?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 4
Vennligst velg den/de beskrivelsen(e) som best beskriver din bedrifts forretningsmodell (Mulig å velge flere)

( ) B2B2C: Vi hjelper andre bedrifter med utvikling av digitale løsninger og/eller ferdige digitale løsninger rettet mot deres sluttkunder
( ) B2B: Vi hjelper andre bedrifter med utvikling av digitale løsninger og/eller ferdige digitale løsninger for intern bruk i bedriften
( ) B2C: Vi utvikler digitale produkter og/eller tjenester rettet mot våre egne sluttkunder
( ) Vet ikke

Spørsmål 5
Hva er det bedriften din primært tilbyr innenfor digitale produkter og/eller tjenester? (Mulig å velge flere alternativer)

Begreper
- *Hyllevare-løsninger er standardprogramvare utviklet for å dekke generelle behov med lite eller ingen tilpasning.*
- *SaaS-løsninger er skybaserte programvaretjenester tilbudt på abonnementsbasis, tilgjengelig via internett uten lokal installasjon.*

( ) Konsulenttjenester
( ) Saas-løsninger (Software as a service)
( ) Hyllevare-løsninger (White-label)
( ) Annet
( ) Vet ikke

Spørsmål 5.1
Vennligst beskriv hva annet bedriften din tilbyr innen digitale produkter og/eller tjenester
[ ]

Spørsmål 6
Hvilken faktor opplever bedriften din som viktigst blant kunder som går til anskaffelse av digitale produkter og/eller tjenester?

( ) Lav pris
( ) Egnet for å skalere
( ) Fleksibilitet og mulighet for tilpasning
( ) Kontinuerlig støtte og vedlikehold
( ) God kommunikasjon
( ) Hurtig leveranse
( ) Vet ikke

Spørsmål 7
Føler bedriften din press fra nye, innovative aktører innen utvikling av digitale produkter og/eller tjenester?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 8
Hvordan vurderer du potensialet for disruptive innovasjoner i markedet for utvikling av digitale produkter og tjenester?

( ) Sannsynlig i nær fremtid
( ) Sannsynlig på lengre sikt
( ) Lite sannsynlig
( ) Ikke sannsynlig
( ) Vet ikke

Spørsmål 9
Hvilke forberedelser gjør bedriften din for å raskt kunne tilpasse seg nye, disruptive innovasjoner? (Mulig å velge flere alternativer)

( ) Investering i forskning og utvikling
( ) Skilt ut en egen innovasjonsenhet med selvstendig drift
( ) Overvåkning av markedstrender
( ) Andre forberedelser
( ) Ingenting
( ) Vet ikke

Spørsmål 9.1
Vennligst beskriv hvilke andre forberedelser bedriften din gjør
[ ]

Spørsmål 10
Har low-code/no-code-løsninger påvirket din bedrift de siste 10 årene?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 10.1
Hvordan har low-code/no-code-løsninger påvirket din bedrift?

( ) Positivt
( ) Negativt
( ) Vet ikke

Spørsmål 10.2
Hvilke digitale produkter og/eller tjenester har blitt påvirket av low-code/no-code?

( ) **B2B:** Til intern bruk i kundebedrifter
( ) **B2C:** Til kundebedrifters sluttkunder
( ) Vet ikke

Spørsmål 11
Hvordan vurderer bedriften din veksten i markedet for utvikling av digitale produkter og tjenester ...

	Høy	Moderat	Lav	Vet ikke
... de siste 10 årene?	( )	( )	( )	( )
... frem mot 2030?	( )	( )	( )	( )

Spørsmål 12
Tror bedriften din at det vil være flere eller færre tilbydere i markedet for utvikling av digitale produkter og tjenester i 2030 enn i dag?

( ) Flere
( ) Omtrent det samme som i dag
( ) Færre
( ) Vet ikke

Spørsmål 13
Hvordan vurderer bedriften din lønnsomhetsutviklingen i markedet for utvikling av digitale produkter og tjenester frem mot 2030?

( ) Mer lønnsomt enn i dag
( ) Like lønnsomt som i dag
( ) Mindre lønnsomt enn i dag

Tusen takk!
Vi setter veldig stor pris på at du tok deg tiden til å hjelpe oss ☺
Svarene dine er sendt inn, men du og bedriften din forblir anonyme.

A.4.2 Clientele

Spørreundersøkelse om etterspørselen etter digitale produkter og tjenester

Denne undersøkelsen er utformet av Halvard Haugse og Magnus Nadheim for om å samle inn supplerende data til deres masteroppgave ved Norges Handelshøyskole (NHH) høsten 2023.

Svarene dine vil bli brukt til å kartlegge etterspørselen etter digitale produkter og tjenester i Norge.

All informasjon du deler med oss blir anonymisert i henhold til taushetserklæringen på neste side.

Taushetserklæring

Alle svar behandles konfidensielt, forblir anonyme og kan ikke spores tilbake til deg eller din bedrift.
Ved å trykke "Jeg godtar" godtar du at svarene dine brukes for å fremstille data i et samlet format.

Eksempelvis:

- *"80% av norske bedrifter bryr seg mest om lav pris ved kjøp av digitale produkter og tjenester"*
- *"CEOs i små foretak synes å være mer åpne for bruken av kunstig intelligens enn CEOs i store foretak"*

Undersøkelsen starter dersom du godtar taushetserklæringen

( ) Jeg godtar
( ) Jeg godtar ikke

Spørsmål 1
Vennligst fyll inn informasjonen nedenfor.

Bedriftens navn: [ ]
Din stilling i bedriften: [ ]

Spørsmål 2
Er bedriften din en del av et større konsern og/eller en norsk filial av en utenlandsk virksomhet?
( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 3
Med IT-kompetanse referer vi til kunnskap og ferdigheter i forbindelse med bruk, forståelse og utvikling av teknologiske systemer og programvare.

Hvordan vil du vurdere ...

	Over gjennomsnitt	Gjennomsnittlig	Under gjennomsnitt	Vet ikke
... IT kompetansen i selskapet ditt?	( )	( )	( )	( )
... din egen IT kompetanse?	( )	( )	( )	( )

Spørsmål 4
Digital modenhet kan defineres som en organisasjons evne til effektivt å adoptere og utnytte nye digitale prosesser, programvare og teknologier.

Hvordan vil du vurdere ...

	Moden	Litt moden	Umoden	Vet ikke
... din egen bedriftsdigitale modenhet?	( )	( )	( )	( )
... lignende bedriftersdigitale modenhet?	( )	( )	( )	( )

Spørsmål 5
Digitale produkter og tjenester refererer i denne sammenhengen til software (programvare) -applikasjoner, -verktøy og -tjenester og ikke hardware (maskinvare).

Har bedriften din digitale produkter og/eller tjenester rettet ut mot kundene deres?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 6
Har bedriften din planer om å anskaffe eller utvikle digitale produkter og/eller tjenester som skal rettes ut mot kundene deres?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 6.1
Nedenfor er noen mye brukte metoder for anskaffelse av digitale produkter og tjenester.

Vennligst velg de metodene dere planlegger å bruke for å anskaffe digitale produkter og tjenester som skal rettes ut som kundene deres. (Mulig å velge flere)

Begreper

- *Outsourcing innebærer å delegere utviklings- og/eller vedlikeholdsoppgaver til en ekstern organisasjon eller frilanser.*

- *Hyllevare-løsninger er standardprogramvare utviklet for å dekke generelle behov med lite til ingen grad av tilpasning.*
- *SaaS-løsninger er skybaserte programvaretjenester tilbudt på abonnementsbasis, tilgjengelig via internett uten lokal installasjon.*

( ) Ansette interne ressurser for å utvikle det
( ) Leie inn konsulenter for å utvikle det
( ) Outsource utvikling
( ) Kjøpe hyllevare-løsning (whitelabel)
( ) Lisensiere SaaS-løsning (Software as a Service)
( ) Andre metoder

Spørsmål 6.2
Hvilke andre metoder planlegger dere å bruke for å anskaffe digitale produkter og/eller tjenester som retter seg ut mot kundene deres?
[ ]

Spørsmål 5.1
Nedenfor er noen mye brukte metoder for anskaffelse av digitale produkter og tjenester.

Vennligst velg de metodene dere har brukt for å anskaffe de digitale produktene og tjenestene dere har som retter seg ut mot kundene deres. (Mulig å velge flere)

Begreper

- *Outsourcing innebærer å delegere utviklings- og/eller vedlikeholdsoppgaver til en ekstern organisasjon eller frilanser.*
- *Hyllevare-løsninger er standardprogramvare utviklet for å dekke generelle behov med lite til ingen grad av tilpasning*
- *SaaS-løsninger er skybaserte programvaretjenester tilbudt på abonnementsbasis, tilgjengelig via internett uten lokal installasjon.*

( ) Ansatt interne ressurser for å utvikle det
( ) Leid inn konsulenter for å utvikle det
( ) Outsourcet utvikling
( ) Kjøpt hyllevare-løsning (whitelabel)
( ) Lisensiert SaaS-løsning (Software as a Service)
( ) Andre metoder
( ) Vet ikke

Spørsmål 5.2
Hvilke andre metoder har dere brukt for å anskaffe digitale produkter og/eller tjenester som retter seg ut mot kundene deres?
[ ]

Spørsmål 7
Hvilken faktor opplever du som viktigst for bedriften din ved utvikling eller anskaffelse av digitale produkter og/eller tjenester som skal rettes ut mot kundene deres?

( ) Lav pris
( ) Egnet for å skalere
( ) Fleksibilitet og mulighet for tilpasning
( ) Kontinuerlig støtte og vedlikehold
( ) God kommunikasjon
( ) Hurtig leveranse
( ) Vet ikke

Spørsmål 8
Opplever du bedriften din som fornøyd med deres nåværende digitale produkter og/eller tjenester som retter seg ut mot kundene deres?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 9
Planlegger bedriften din å gå til anskaffelse av flere digitale produkter og/eller tjenester som skal rettes ut mot kundene deres i fremtiden?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 9.1
Vennligst velg de metodene dere planlegger å bruke for å anskaffe nye digitale produkter og/eller tjenester som skal rettes ut mot kundene deres. (Mulig å velge flere)

( ) Ansette interne ressurser for å utvikle det
( ) Leie inn konsulenter for å utvikle det
( ) Outsource utvikling
( ) Kjøpe hyllevare-løsning (whitelabel)
( ) Andre metoder
( ) Vet ikke

Spørsmål 9.2
Hvilke andre metoder planlegger dere å benytte for å anskaffe digitale produkter og/eller tjenester som skal rettes ut mot kundene deres?
[ ]

Spørsmål 10
Har bedriften din digitale produkter og/eller tjenester for intern bruk?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 11
Har bedriften din planer om å anskaffe eller utvikle digitale produkter og/eller tjenester for intern bruk?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 11.1
Nedenfor er noen mye brukte metoder for anskaffelse av digitale produkter og/eller tjenester.

Vennligst velg de metodene dere planlegger å bruke for å anskaffe nye digitale produkter og/eller tjenester for intern bruk. (Mulig å velge flere)

Begreper

- *Outsourcing innebærer å delegere utviklings- og/eller vedlikeholdsoppgaver til en ekstern organisasjon eller frilanser.*
- *Hyllevare-løsninger er standardprogramvare utviklet for å dekke generelle behov med lite til ingen grad av tilpasning*
- *SaaS-løsninger er skybaserte programvaretjenester tilbudt på abonnementsbasis, tilgjengelig via internett uten lokal installasjon.*

( ) Ansette interne ressurser for å utvikle det
( ) Leie inn konsulenter for å utvikle det
( ) Outsource utvikling
( ) Kjøpe hyllevare-løsning (whitelabel)
( ) Lisensiere SaaS-løsning (Software as a Service)
( ) Andre metoder
( ) Vet ikke

Spørsmål 11.2
Hvilke andre metoder planlegger dere å benytte for å anskaffe digitale produkter og/eller tjenester for intern bruk?
[ ]

Spørsmål 10.1
Nedenfor er det listet opp kjente metoder for anskaffelse av digitale produkter og/eller tjenester.

Vennligst velg de metodene dere har brukt for anskaffelsen av digitale produkter og/eller tjenester for intern bruk (Mulig å velge flere)

Begreper
- *Outsourcing innebærer å delegere utviklings- og/eller vedlikeholdsoppgaver til en ekstern organisasjon eller frilanser.*
- *Hyllevare-løsninger er standardprogramvare utviklet for å dekke generelle behov med lite til ingen grad av tilpasning*
- *SaaS-løsninger er skybaserte programvaretjenester tilbudt på abonnementsbasis, tilgjengelig via internett uten lokal installasjon.*

( ) Ansatt interne ressurser for å utvikle det
( ) Leid inn konsulenter for å utvikle det
( ) Outsourcet utvikling
( ) Kjøpt hyllevare-løsning (whitelabel)
( ) Lisensiert SaaS-løsning (Software as a Service)
( ) Andre metoder
( ) Vet ikke

Spørsmål 10.2
Hvilke andre metoder har dere brukt for å anskaffe digitale produkter og/eller tjenester for intern bruk?
[ ]

Spørsmål 12
Hvilken faktor opplever du som viktigst for bedriften din ved utvikling eller anskaffelse av digitale produkter og/eller tjenester for intern bruk?

( ) Lav pris
( ) Egnet for å skalere
( ) Fleksibilitet og mulighet for tilpasning
( ) Kontinuerlig støtte og vedlikehold
( ) God kommunikasjon
( ) Hurtig leveranse
( ) Vet ikke

Spørsmål 13
Opplever du bedriften som fornøyd med deres nåværende digitale produkter og/eller tjenester for intern bruk?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 14
Planlegger bedriften din å gå til anskaffelse av flere digitale produkter og/eller tjenester for intern bruk i fremtiden?

( ) Ja
( ) Nei
( ) Vet ikke

Spørsmål 14.1
Hvordan planlegger bedriften din å anskaffe nye digitale produkter og/eller tjenester for intern bruk? (Mulig å velge flere)

( ) Ansette interne ressurser for å utvikle det
( ) Leie inn konsulenter for å utvikle det
( ) Outsource utvikling
( ) Kjøpe hyllevare-løsning (whitelabel)
( ) Lisensiere SaaS-løsning (Software as a Service)
( ) Andre metoder
( ) Vet ikke

Spørsmål 14.2
Hvilke andre metoder planlegger dere å benytte for å anskaffe digitale produkter og/eller tjenester for intern bruk?
[ ]

Spørsmål 15
Hva er ditt estimat på kostnadene i bedriften din for ...

	< 1 MNOK	1-5 MNOK	5-15 MNOK	15-50 MNOK	>50 MNOK	Vet ikke	Ønsker ikke å oppgi
... anskaffelse av eksisterende digitale løsninger?	( )	( )	( )	( )	( )	( )	( )
... årlig vedlikehold og videreutvikling av eksisterende digitale løsninger	( )	( )	( )	( )	( )	( )	( )

Spørsmål 16
Hva er ditt estimat på kostnadene i bedriften din for ...

	< 1 MNOK	1-5 MNOK	5-15 MNOK	15-50 MNOK	>50 MNOK	Vet ikke	Ønsker ikke å oppgi
... anskaffelse av nye digitale løsninger?	( )	( )	( )	( )	( )	( )	( )
... årlig vedlikehold og videreutvikling av nye digitale. løsninger?	( )	( )	( )	( )	( )	( )	( )

Spørsmål 17
Hvordan vurderer du risikoen forbundet med bruken av mindre kjente leverandører ved anskaffelse eller utvikling av digitale produkter og/eller tjenester som ...

	Høy risiko	Noe risiko	Lav risiko	Ingen risiko	Vet ikke
... rettes mot slutt kunder?	( )	( )	( )	( )	( )
... kun er til internbruk?	( )	( )	( )	( )	( )

Spørsmål 18
I hvilken grad tror du at kunstig intelligens vil påvirke hvordan man anskaffer digitale produkter og/eller tjenester som ...

	I stor grad	I noe grad	I liten grad	I ingen grad	Vet ikke
... rettes mot slutt kunder?	( )	( )	( )	( )	( )
... kun er til internbruk?	( )	( )	( )	( )	( )

Tusen takk!
Vi setter veldig stor pris på at du tok deg tiden til å hjelpe oss ☺
Svarene dine er sendt inn, men du og bedriften din forblir anonyme.