Copenhagen Business School

Master Thesis

Innovation and firm performance

A longitudinal study of the Innovation Survey in Denmark

Authors:
Håkon Ottesen (116313)
Martine Holt Hasle (116310)

Supervisor:
Thomas Einfeldt

A thesis submitted in fulfillment of the requirements for the degree of MSc. Applied Economics and Finance

Total pages: 92
Total characters: 187,863
Submission date: May 15, 2019

Abstract

Enterprises in all sectors face extensive competitive and changing environments, and as a consequence, ought to have a modern and innovative mindset to compete. Research has shown that investments in specific innovation activities, depending on the objective, has a substantial effect on firm performance. This study aims to determine what the relationship between innovation input, innovation output and firm performance for Danish enterprises is, namely, which factors are vital given the different objectives. This relationship has been studied mainly by looking at expenditures to innovation activities and cooperation, implemented innovation types, and firm performance.

Building on existing literature concerning R&D and innovation studies, it was found appropriate to conduct a quantitative analysis utilizing longitudinal estimation techniques, including both linear and quasi-likelihood methods to adapt the different characteristics of the variables. The econometric models will be estimated on a balanced data panel covering nine industries from 2009 to 2016, mostly based on data from the Innovation Survey by Statistics Denmark. The results of the regression analyses indicate several strong and meaningful relationships. Of the innovation activities, that is R&D and non-R&D activities, only expenditures to intramural and consultancy services provided a significant effect on sales given a two years delay, and all innovation activities except acquisition of external rights suggest a substantial influence on the implementation of either one of the different types of innovation. Out of these innovation types, the findings suggest that marketing innovations play a key role in generating income short-term.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

There is a consensus among scholars that a firm's power to utilize and improve their innovative abilities is positively related to performance and competitive advantage (Bettis and Hitt, 1995; Helfat and Peteraf, 2003). The need for new processes and inventive organizing has become imperative for all establishments, including those already engaged in research. Firms must generate continual innovations in order to gain market shares or overcome oppositions to endure the fiercer competition and declining product life cycles. According to Statistics Denmark, companies are spending more on innovative activity than before, which could be the result of increasing levels of competition that forces firms to innovate in order to stay operational (Porter, 1990).

Innovation could lead to the retaining of existing customers that further enables them to obtain new markets with a substantial amount of growth in sales. However, innovation is a comprehensive process, and it is necessary to identify specific indicators that explain innovation implicitly (Fagerberg, Mowery, Nelson, et al., 2005). Moreover, significant innovation inputs are required in order to achieve innovation output. These inputs may be a series of activities that firms invest in, aiming at developing specific innovations as output that could further lead to increased sales and value for the stakeholders (OECD/Eurostat, 2005). Hence, it is vital to identify the innovation activities that generate the preferred type of innovation, to succeed in today's modern and competitive environments.

The initiation of innovation does not only involve expenditures to R&D activities, but equally important, expenditures to non-R&D activities. According to the Innovation Survey from Statistics Denmark in 2016, enterprises spent most resources on systematic R&D activities. The R&D expenditures amounted to 87 per cent of the total sum spent, whereas non-R&D activities merely represented 13 per cent. Denmark has expressed a goal to have strong innovative capabilities as innovation and restructuring-abilities represents one of the most vital sources of increased wealth creation. Figure 1.1 shows the total amount in billions invested in R&D and non-R&D activities from the nine main industries specified later on. There is a downward trend in the two first years, whereas a clear upward trend is present until 2016. The expenditures have increased overall by approximately 10 billion Danish kroner, from 60 billion in 2009 to around 70 billion in 2016.

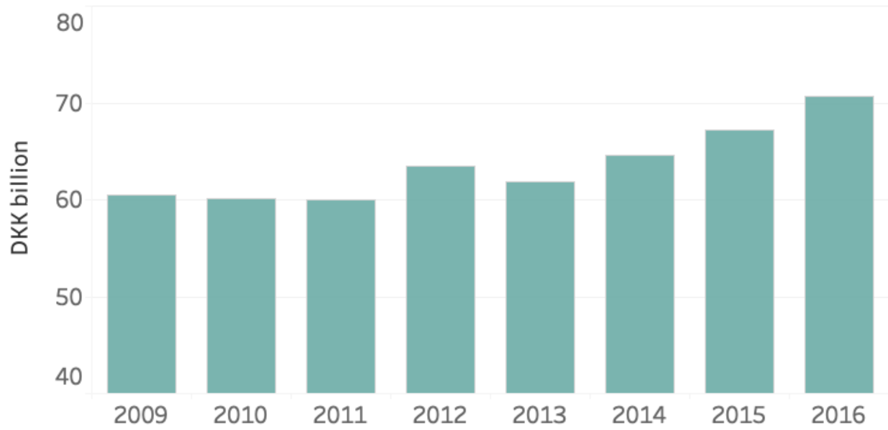

Figure 1.1: Total amount invested in R&D and non-R&D activities (Statistics Denmark)

Specific innovation activities and types that have proven significant for firm performance is the main focus of this thesis. The innovation types are extended from product and process innovation as put forward by existing literature (Li, Chen, & Shapiro, 2010), to the updated definition of innovation proposed by the Organization for Economic Co-operation and Development (OECD), that is, product, process, organizational, and marketing innovation (OECD/Eurostat, 2005). Figure 1.2 shows the development of the mentioned innovation types from 2009 to 2016, and relatively how much each innovation is implemented overall. In general, organizational and marketing innovation, are slightly more common than product and process innovations.

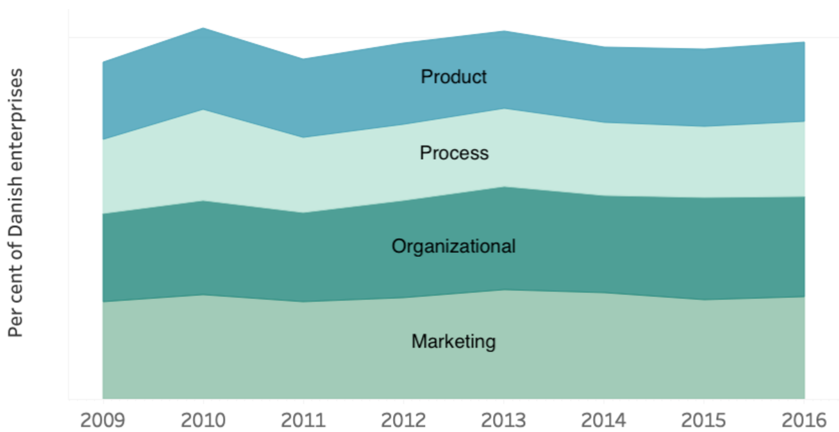

Figure 1.2: Development of the OECD innovation types (Statistics Denmark)

This thesis contributes to innovation studies by exploring the innovation practices of Danish firms, specifically; their innovation activities, types of innovation and cooperation, as well as the effect on firm performance. These aspects will make grounds for a quantitative study of innovation with secondary data from a large portion of domestic companies. The purpose of this thesis is to add to the existing scholarship and to provide a new perspective on the Danish market.

1.1 Motivation

The motivation for this thesis emerges from the ongoing discussion on the effects of innovation in today's economic environment. Innovation is a key tool for firms to be more productive and more adaptable to change, and have become an integral part of firms' strategies to achieve a competitive advantage. R&D and other non-R&D activities are recognized as fundamental aspects of innovation, and strategic decisions concerning these activities are, therefore, a critical first step of the innovation process (OECD/Eurostat, 2005).

There is a clear trend among academics to study patents application, R&D expenses and government funding in the light of firm performance. However, limited research exists on non-R&D activities and the OECD innovation types. This paper will contribute to the existing literature by focusing on several aspects of the innovation procedure based on the Innovation Survey. More specifically, by focusing on expenditures related to both R&D and non-R&D activities, product, process, organizational and marketing innovation, and sales, this paper gives a whole new perspective toward the progression of innovation in the Danish market.

In a competitive economy, firms invest large sums in innovation to be innovative and to gain the upper hand. Adopting the right strategic decisions are a vital factor in implementing successful innovations. By analyzing innovation measures in the light of innovation inputs and innovation outputs, this paper will highlight areas of great importance in the Danish business sector. The firms' ability to introduce new or improved products, processes, organizational practices or marketing methods will be considered to provide a unique perspective on this highly-debated topic. Further, exploring innovation activities and innovation types directly to firm performance adds to the understanding of innovation. Thus, it is found relevant to conduct a quantitative research on the relationship between innovation input, innovation output, and firm performance.

1.2 Problem statement

This thesis aims to investigate, by using a quantitative approach, how innovation will increase the firms' competitive advantage. More specifically, how innovation activities and cooperation affects the implementation of product, process, organizational, and marketing innovations, and subsequently, how this affects firm performance. The problem statement is formulated as follows:

What is the relationship between innovation input, innovation output and firm performance for Danish enterprises?

An illustration of the research framework is presented in figure 1.3 and the following sub-questions are introduced to support the main research question further:

1. How do innovation activities and innovation types differ between industries?
2. How do innovation activities affect firms' sales?
3. How do innovation activities and cooperation affect types of innovation?
4. How do types of innovation affect sales?

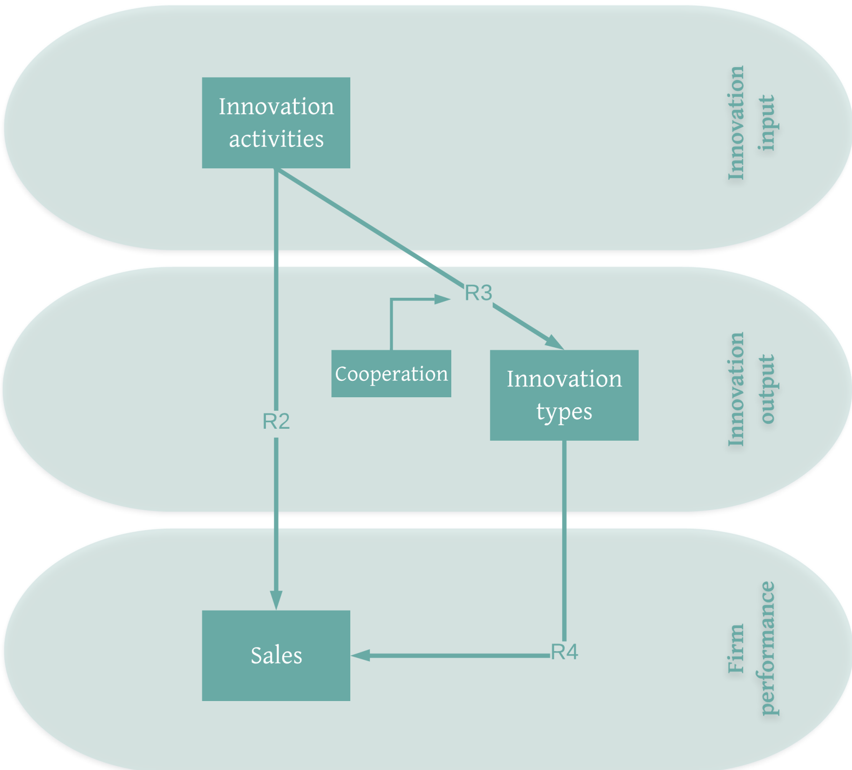

Figure 1.3: Research framework

1.3 Delimitation

The concept of innovation is a broad subject. As stated, the purpose of this research is to highlight the effect of innovation activities and types of innovation on firm's performance and to bring a new perspective to existing research. However, as this study is part of the two-year program M.Sc. in Applied Economics and Finance at Copenhagen Business School, the paper will be limited by scope. It will be limited to the Danish business sectors' innovation activities and performance collected from Statistics Denmark, as the aim of the paper is to add to the complex dynamics of innovation and performance in the business sector. The primary data used is based on the Innovation Survey, supplied with data from Purchases and Sales by Enterprises. Both statistics exclude the public sector, and hence will not be examined in this thesis. The collected data is from the period 2009 to 2016 and will set the time limit of the study.

1.4 Sources and validity

In order to conduct accurate and sound research that can answer the research questions, the data has to be both reliable and valid. It has to be reliable in the sense that the results supplied are consistent, whereas the validity refers to the extent it measures the intended objective (Carmines & Zeller, 1979). A combination of reliable and valid sources of the data used is optimal and determines the quality of the research.

All data analyzed is provided by Statistics Denmark, which is accountable for collecting, producing and providing relevant and reliable statistics of high quality (DST, 2016d; DST, 2016a). Research communities, public administration, and international organizations regularly use the publicly available statistics, as a basis of strategic decisions and further research (DST, 2016e). In other words, Statistics Denmark is a highly recognized source, and the collected data is therefore deemed to be reliable and of good quality. Furthermore, only acknowledged and reliable academic literature is used in order to ensure the quality of the references.

The Innovation Survey performed yearly by Statistics Denmark serve as the main source of data. The survey provides information concerning innovation input, the innovation process, and the innovation output via questionnaires for a representable sample of Danish business enterprises. The electronic questionnaires have a high response rate, with limited bias due to integrated and manually detailed controls performed by Statistics Denmark (DST, 2016c). The content of the questionnaires is performed following the guidelines of data collection and reporting in the Oslo Manual proposed by Eurostat and OECD (DST, 2016b).

1.5 Thesis outline

This paper is divided into eight chapters. Chapter 2 is the literature review and gives justification and arguments for central components, both regarding theory and key concepts. Chapter 3 will introduce and explain the key topics to fully understand the underlying assumptions and concepts discussed in this thesis. The research design is presented in chapter 4, including variables description, data information, and theory from the methods used in the regression analysis. A descriptive and econometric analysis will be conducted in chapter 5 and 6, respectively before the findings are discussed in chapter 7. Chapter 8 concludes and adds up the findings and the preceding discussion.

Chapter 2

Literature review

2.1 Innovation input and firm performance

Many scholars have studied the topic innovation, and widespread academic literature exists on the innovation-performance relationship. Many scholars find that investing in innovation activities ultimately enhance a firm's performance (Cohen and Levinthal, 1989; Dosi, 1988; Baldwin, Business, and Group, 1994; Wieser, 2005; Kafouros, Buckley, Sharp, and Wang, 2008; Sharif, Baark, and Lau, 2012). For example, Branch (1974) used a distributed lag model with the cross-section data of 111 firms within seven different industries and found that R&D expenditures have a positive effect on future profits and sales growth.

However, according to Jaruzelski, Dehoff, and Bordia (2005) R&D activities do not necessarily increase profits. R&D expenditures are risky and increase the volatility of future operative performance (Pandit, Wasley, & Zach, 2011). In addition, R&D activities may not always be able to achieve their set goals (Baker & Freeland, 1975). Mitchell and Hamilton (1988) found that R&D expenditure boosts a firm's innovation activity, but that not all R&D projects are necessarily beneficial for a firm's performance as some projects might be unsuccessful. On the other side, Klomp and Van Leeuwen (2001) studied the relationship between innovation output and firm performance and found that process innovation is both strongly and positively associated with sales, productivity and employment growth. A negative relationship between product innovation and employment growth were found; however, this connection proved to be insignificant.

To some extent, there seems to be shared opinions on the innovation-performance relationship (Kafouros et al., 2008). In this context, Wieser (2005) conducted a meta-analysis of earlier research with mixed results and found on average a strong and positive relationship between R&D expenditures and firm performance. Indeed, Kafouros et al.

(2008) argue that the reason for the variation across previous studies may be due to the lack of an understanding of the factors influencing the innovation-performance relationship.

2.2 Innovation output and firm performance

Innovation is a broad topic that several researchers have tried to define for decades, resulting in a diversity of definitions and approaches. Some scholars look at the direct relationship between R&D expenditures and performance, whereas others use intermediate outcomes, such as patent numbers and innovative sales, as measures for a firm's innovativeness with R&D expenditures as the innovation input. Numerous scholars argue that it is innovation output that increases productivity, rather than innovation input (see for instance Crépon, Duguet, and Mairessec, 1998; Lööf, Heshmati, Asplund, and Nåås, 2001; Pandit et al., 2011; Mohnen, Mairesse, and Dagenais, 2006). Crépon et al. (1998) looked at how investing in R&D activities influenced process or product innovations and productivity, thus relating innovation input and output with economic performance. Using a cross-section of French firms in the manufacturing sector, they found a positive effect of R&D activity on innovation output measured by numbers of patents and a positive and significant effect on firm productivity. The model combines the output of innovation with a production function (Parisi, Schiantarelli, & Sembenelli, 2006) and is referred to as the Crepon-Duguet-Mairesse (CDM) model.

The appointed CDM model has later been an inspiration to a growing literature studying the innovation-performance relationship. Mairesse and Mohnen (2003) linked R&D as input to innovation output and productivity, using firm-level data from the second Community Innovation Surveys (CIS2) of France, Germany, Spain, and the United Kingdom. They found a positive effect of R&D intensity on the process and product innovation, measured by patent numbers and innovative sales, respectively. Further, they found a positive effect of patents and innovative sales on labor productivity, whereas only innovative sales had a significant effect. Using the same CIS2 data of manufacturing firms in Finland, Norway, and Sweden, Lööf et al. (2001) found a positive relationship between product and process innovation and firm productivity. Parisi et al. (2006) found that R&D input was positively associated with the process and product innovation, which in turn had a significant impact on firm productivity of Italian firms. However, it was proved that process innovation had a much stronger effect.

Measures of firm performance

In addition to a wide range of definitions on innovation, scholars differ in the ways performance are measured, but it is usually related to firm profitability, revenue (Kafouros et al., 2008) and productivity (e.g., Crépon et al., 1998; Lööf et al., 2001; Mairesse and Mohnen, 2003). Pandit et al. (2011) combined inputs and outputs by examining how the relationship between R&D expenses and the future operating performance is better understood by including patent numbers and citations as a proxy for innovation performance. By conducting a longitudinal study of 272 firms across 35 industries, Artz, Norman, Hatfield, and Cardinal (2010) found evidence of a positive relationship between R&D expenditures and process innovation, measured by the number of patents granted. However, they further found that patents are negatively associated with both ROA and sales growth, whereas product innovation measured by a number of new product announcements has a positive impact.

Despite differences in measures of firm performance, there seems to be a general agreement of the positive and significant relationship between innovation output and firm performance. This is aligned with the positive and significant result Lööf and Heshmati (2006) found when conducting a sensitivity analysis on papers with different performance measures. In this regard, sales are used as a proxy for a firm's performance in this research.

2.3 Innovation types as outputs

Innovation output is, however, not only associated with new products or processes, but also with a company's marketing and organization (Gunday, Ulusoy, Kilic, & Alpkan, 2011). Schumpeter (1934) identified five types of innovation: (1) new products, (2) new production processes, (3) the exploitation of a new market, (4) new sources of supply and (5) carrying out a new organizational structure. In line with Schumpeter's theory of innovation, the Oslo Manual (OECD/Eurostat, 2005) defines four types of innovation: (1) product innovation, (2) process innovation, (3) marketing innovation and (4) organizational innovation.

In accordance to the four innovation types proposed by the OECD, Gunday et al. (2011) explored the effects of the product, process, marketing and organizational innovations on different features of performance, including innovative, production, marketing, and financial performance. By conducting an empirical study of 184 Turkish manufacturing firms, they found a direct and positive impact of the product, marketing and organizational innovations on innovative performance, measured as, among other, new patents, new product announcements, and new projects. However, only process innovations were found to have an indirect effect on product innovations. Also, Gunday et al. (2011) found that marketing and organizational innovations influence innovative performance through product innovations and that there was a positive and indirect association between innovative performance and financial performance.

Karabulut (2015) considered the same four innovation outputs when studying the innovation-performance relationship using different performance measures, such as financial performance and learning and growth performance. He found that the results were all significant and positive, except for marketing innovation that had a negative impact on learning and growth performance. Shaukat, Nawaz, Naz, et al. (2013) found a positive relationship between products, process, marketing and organizational innovations and performance in Pakistani manufacturing firms. Similar results were found for low and medium-low-technology industries in Europe using data from the Fourth Community Innovation Survey (CIS4) (Heidenreich, 2009). Furthermore, Jiménez-Jiménez and Sanz-Valle (2011) and Ortt and van der Duin (2008) found evidence that organizational innovation has a positive impact on performance.

Following previous research, the Innovation Survey provided by Statistics Denmark will be used as a fundamental source to identify and classify the different innovation outputs. That is, product, process, organizational, and marketing innovation will be the measures of a firms' innovativeness in this paper.

2.4 Innovation activities as inputs

Sharif et al. (2012) explored the relationship between innovation activities (i.e., inputs) and innovation performance (i.e., outputs), using the Fourth Community Innovation Survey (CIS4) of 492 companies in Hong Kong. The inputs are associated with innovation activities, sources of innovation and expenditures, whereas product, process, marketing and organizational is related to the innovative outputs. They found evidence that intramural and extramural R&D spending, acquiring new machinery as well as cooperation with an external partner are all vital for the innovation outputs. Indeed, various academics argue that firms operate and conduct the innovation activities in collaboration with external organizations (Freeman, 1991, Harland, 1996, Gulati, Nohria, and Zaheer, 2000). Sharif et al. (2012) claim that to innovate in interaction with partners such as customers, suppliers, public research institutes or other organizations, may provide firms with required external inputs.

Conclusively, this paper seeks to extend previous research and thus examine the innovation-performance relationship, concerning innovation inputs, innovation outputs and firm performance based on the Innovation Survey. The relationship between innovation activities (i.e., input) and the four innovation types product, process, organizational and marketing innovation (i.e., output) is studied, followed by an examination in which inputs and outputs directly affect firm performance in terms of sales. Cooperation will be included as an interaction term in the relationship between innovation activities and innovation types, to test whether innovation activities done in collaboration with partners is essential to the implementation of the four innovation types.

Chapter 3

Key concepts

3.1 The concept of innovation

Innovation provides the foundation for new companies, new workplaces, and productivity growth and hence is an essential driver for economic growth. Innovation can help address the social and global issues associated with demographic changes, resource shortage, and the shifting climate, all at a lower cost (OECD, 2010). Furthermore, innovation is a key tool for firms to gain competitive advantage and to survive (Karabulut, 2015). Innovative firms are proving to be more productive, more robust to shock and more adaptable to change, that may, as a result, improve their performances, overcome competitors and provide value to their stakeholders (OECD, 2010).

Several definitions exist in the academic literature, but the term "innovation" is commonly accepted as the exploitation of new ideas that may be regarding a new process, a feature or an outcome. Schumpeter (1934) first introduced five types of innovation:

1. The launch of new products or new product qualities;
2. The introduction of new methods of production or sales of a product
3. The exploitation of new markets
4. The acquiring of new sources of supply; and
5. The creation and application of a new way to organize business

The OECD defines innovation as:

"Innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations" (OECD/Eurostat, 2005, p.46).

This broad definition contains several possible innovations, and the OECD/Eurostat (2005) is further classifying innovation as four innovation types: product innovation, process innovation, organizational innovation, and marketing innovation.

Innovations are about implementing something new with the aim of value creation. That is, in order to be qualified as an innovation, two requirements must be met: (1) the product, process, marketing method or organizational method must be either new or significantly improved to the firm, (2) and it must be implemented in practice. The firm may either develop the innovation itself or adopt it from other firms or institutions (OECD/Eurostat, 2005). Further, it follows from the OECD/Eurostat (2005) that a new or enhanced product is implemented when it is launched, while new processes, marketing approaches or organizational approaches are implemented when they are put into use in the firm's business.

A firm that has implemented an innovation in the form of product, process, marketing method or organizational method during the observation period, is referred to as an innovative firm. It is not a requirement for an innovation to be of commercial success, as many innovations fail during the attempt (OECD/Eurostat, 2005).

Furthermore, the term "innovation" can be used in various contexts to refer to either a process or an outcome. Aligned with the OECD/Eurostat (2018), the term "innovation activities" will in this paper refer to the input of the process while the term "innovation" is restricted to outcomes.

3.2 The OECD Oslo Manual

What is the Oslo Manual?

The Oslo Manual developed by Eurostat and the OECD is the main international provider of guidelines for the collection and use of innovation data and creates a platform for research and testing on innovation measurement. Today, the importance of innovation measures is acknowledged by several countries and international organizations, which are in a continuous process of collecting innovation data following the guidelines. The Oslo Manual is primarily directed at supporting producers of innovation statistics, such as offices for national statistics, but is also suitable for users of the innovation information OECD/Eurostat (2018).

Purpose

Innovation is an important driver of the economic growth that benefits individuals, institutions, and the economy as a whole. It is therefore important for decision-makers with a sound measurement and those using the innovation data in research, to fully understand economic and social developments, assess the contribution of innovation to social and economic goals, and monitor and evaluate the efficiency of their policies (OECD/Eurostat, 2018).

The Oslo Manual aims to provide a directory with common terminology, applied conventions, and principles agreed upon to collect and report innovation data. It has since 1922 been the international guidebook for measuring innovation, providing a base for discussing innovation, activities supporting innovation as well as innovation outcomes at a universal language (OECD/Eurostat, 2018).

New edition

The economic and social factors, the nature and origin of innovation, and the exchange of experience among experts are constantly in change. Besides, there exists a political demand for empirical evidence concerning innovation and a positive trend in social awareness of innovation, increasing the interest in new potential objects that need to be measured. This makes it important to continually develop how innovation is measured and allow for further research to improve and expand innovation data. Since 1922, the Oslo Manual has been revised three times to account for the continual evolution (OECD/Eurostat, 2018).

In 2018, a third edition was published based on experiences from the ongoing assembling of innovation data from the 1990s in OECD and non-member countries. This version takes the role of digitized information into account and provides a better guide on understanding the link between digitalization and innovation. The newest edition recognizes and proposes new potential innovation activities and other competencies related to data and digital platforms, to be measured (OECD/Eurostat, 2018).

The second edition of the Oslo Manual introduces the four innovation types product, process, organizational, and marketing innovation, whereas the third edition distinguishes between product and business process innovation. However, as the period of the extracted innovation data is based on the second edition of the Oslo Manual, the definitions and particularities of the four innovation types as stated in this edition will be used.

3.3 The Innovation Survey

The most two common measures of innovation are R&D expenditures and patents applications. Registrations of patents record back to as early as the nineteenth century using national patent offices and institutions to document different intellectual property rights. R&D is usually collected annually with R&D surveys since the 1950s, according to the OECD (2002). However, there is a third source of innovation indicators that have increased significantly in popularity, namely innovation surveys. The reason for this is that actual R&D quantities are only inputs in the innovation process, while patents only concern new innovations which may or may not be successful. Therefore, innovation surveys offer qualitative and quantitative data on different innovation actions and types in a given economy. This type of survey is widely used among statisticians and policymakers to examine innovation performance, in addition to predicting and analyzing the relationship and effects on other economic factors (Mairesse & Mohnen, 2003). Innovation surveys measure companies' abilities to innovate and restructure within areas of product, process, organizational practices, and marketing. It also provides evidence to framework conditions for these activities and how they are managed (DST, 2016d).

There are numerous countries today around the world that performs innovations surveys at different time intervals toward their enterprise segments. The surveys are commonly known as Community Innovation Surveys (CIS) in Europe and are produced at fixed intervals. In general, innovation surveys exist under different names in several other OECD countries, as well as in emerging and developing economies. Nevertheless, the surveys have the same framework and questions concerning innovation; however, differences in content, formulation, and organizing are natural to encounter, also regarding the CIS. From the initiation in 1992, carried out by the European Union in collaboration with the OECD, the survey was produced every four years starting with CIS1 (1990-1992) to CIS4 (2002-2004). A reduction of the time interval from four years down to two years was introduced from 2007 and onwards. As the previous surveys, they were named by the last year in the interval, namely the reference year, and covered a three years period. The first survey in 1992 included only manufacturing firms, and from the second survey (CIS2), service firms were also counted in. Today, the CIS is carried out in all European Union member states.

Structure and content

The Innovation Survey provides several important characteristics about firms' innovation development. Mairesse and Mohnen (2003) summarizes this with three key arguments:

1. It presents indicators of innovation output. That is, specific innovation types (discussed in chapter 3.4), the percentage of sales due to products new to the firm or new to the market, and the shares of products at various stages of the product's life cycle.

2. A better variety of innovation expenditures or activities rather than only R&D expenditures is provided. For instance, information about patents and licenses granted, products design, training of employees and marketing analysis.

3. It arranges for important information about how innovation advances, such as knowledge sources, motives, obstacles, implementation capacity, and partners.

Data is collected on both innovators and non-innovators. "Innovators" are companies that over a period, most likely three years, have presented new products or processes. "New" is referred to as significantly improved or completely new, whereas "new products" are both registered as new to the firm and new to the market. In other words, enterprises are required to give information about inputs, outputs and in general all aspects of their innovative activities.

The questions, which is in accordance with the guidelines from the Oslo Manual, is presented in table 3.1 to give a brief outline. Slight modifications in questions from one CIS to another is normal, and the different series often modernize or introduces new questions. This is because the relevance of some questions expires and new creations provide the need for variation and contemporary focus. For instance, different appropriation procedures were gradually discarded after the CIS2, while environmental and sustainability questions subsequently emerged. Even though the CIS exists in numerous European countries, not all questions are entirely consistent across borders. All countries have their distinctiveness in the questionnaire, as additional questions, different orders or altered formulations of more or less the same question. However, the outline and fundamental questions are practically equivalent throughout the surveys.

Table 3.1: Schematic questionnaire of an innovation survey

1. General information
Independent or part of a group?
Domestic or foreign group?
Country of location
Main industry affiliation
Number of employees (level and growth)
Turnover (level and growth)
Exports (level and growth)
Mother, daughter or sister enterprise
Significant changes in turnover
Newly established
Merger affected turnover for more than 10%
Closure affected turnover for more than 10%
Most significant market: national or international, nearby, or distant
Number of employees with higher education, female, expected increase
Gross investment in tangible goods
Geographic markets in which goods and services are sold
2. Innovator (yes/no)
Introduced new to the firm but not new to the market products in the last 3 years?
If yes: Who developed the new products?
Introduced new to the market products in the last 3 years?
Introduced new process in the last 3 years?
If yes: Who developed the new process?
New to the market?
Unfinished or abandoned innovative projects?
3. Categorical data for innovators
Sources of information for innovation
Objectives of innovation
Effects of innovation
Means of transferring technology
Effectiveness of appropriation mechanisms
4. Dichotomous data for innovators
R&D
R&D done continuously
R&D in cooperation with partners
Most valuable cooperation partner
Government support for innovation from various sources
Applied for a patent?
5. Continuous data for innovators
R&D expenditures (intra- & extramural)
R&D personnel
Innovation expenditures (+ subitems)
Estimated share of products in different phases of life-cycle
Share in total sales of products new to the enterprise but not new to the market
Share in total sales of products new to the market

6. Data on all firms (innovators or not)
Factors hampering innovations
Possession of valid patents
If yes: the number of valid patents
The share of patent-protected sales
Strategic and organizational changes
Use of various IP protection methods
Introduction of organizational innovations
Importance of organizational innovations
Introduction of marketing innovations
Importance of marketing innovations
Introduction of any innovation with environmental benefits
Determinants of environmental innovations
Procedures in place to identify and reduce environmental impacts

List of questions based on the Oslo Manual guidelines, as in the Community Innovation Survey waves. Reproduced from Mairesse and Mohnen (2003)

The questionnaire is usually divided into six parts, as illustrated in table 3.1. Section 1 asks general questions of the individual firms, while section 2 examines if the company was involved in innovative activities during the period. If the company identifies as an innovator in section 2, they will have to continue the questionnaire to the end of section 5, namely answering specifics about innovation, divided into categorical (section 3), dichotomous (section 4), or continuous (section 5) data. The last and sixth section includes several questions that all respondents must complete.

Characteristics of data

The data from innovation surveys is qualitative, subjective and censored (Mairesse & Mohnen, 2003). Following the Oslo Manual, all data is collected from stratified samples. The strata are usually expressed as size, industry, and region. The series of the Innovation Survey comes in the form of a panel with cross-sectional data on firm-level; nevertheless, not all firms are included in every survey. As mentioned, there are a few changes between the surveys, both in time and place. These characteristics provide certain complications which provide the need for special treatment when managing variables and running econometric analysis.

First of all, it is critical to note the qualitative nature of the questionnaire. According to the glossary of statistical terms from OECD (2004), qualitative data describes the attributes or properties that an entity has. The properties are categorized into groups which could be given numeric values. Qualitative data is thus providing less information than quantitative data. However, qualitative data opens up for whole new interpretations and possibilities to analyze subjects that are typically not so easily quantified. Furthermore, qualitative data is less prone to measurements errors, miscalculations, and mathematical inaccuracies. Several adequate econometric methods are developed to cope with this specific type of data. The methodology in section 4.3 will examine the different opportunities concerning the methodology, by looking at models specially designed to handle qualitative data, such as binomial, multinomial and logistic regressions.

The innovation surveys clearly distinguish between innovators and non-innovators. Consequently, the innovating firms are asked questions from section two till five in table 3.1, while the non-innovators are not. This leads to several variables being censored and only assembled for specific companies out of the entire sample. Thus, several observations in the finishing data set could be set at zero, such as expenditures incurred because of new products. However, in some cases, the censoring has to be adjusted to prevent potential selection biases, since the results could essentially have no meaning. A way of reducing the bias is to use sample selection models including a regression for the censored variable with a selection equation (Mairesse & Mohnen, 2003). Nevertheless, it is evident from table 3.1 that there is limited information about firms not exercising innovation activities, and by using this kind of technique and merging data with the Innovation Survey data, it does leave less room to adequately distinguish between innovator and non-innovator to correct for potential selectivity biases. However, with an aggregate perspective for countries or industries, there is enough data to give acceptable indications of significant relationships between whole entities (Mairesse & Mohnen, 2003).

Another critical characteristic of this particular data is the subjective feature. The respondents own personal judgment, perceptions and beliefs could largely influence the individual answers. Some variables would probably make more sense to treat as categorical variables instead of continuous, as they are known to be rounded. This proves the data's subjective nature along with the fact that some definitions in the Oslo Manual are interpreted differently due to their ambiguity, thus leaving room for inaccuracies.

3.4 Innovation types

3.4.1 Product innovation

Product innovation relates to the introduction of a new or significantly improved good or service, where the term "product" is used to cover both goods and services. Product innovations can be linked to the utilization of new knowledge or technologies, or it can be related to new uses or combinations of existing knowledge or technologies. New products refer to goods and services that either differs substantially in their features or the intended use compared to the firm's previous produced products. Even products that have only minor changes to its technical characteristics is referred to as product innovation, as long as the development includes a new intended use (OECD/Eurostat, 2005).

Existing products can have significant improvements in the materials, components or other attributes contributing to improvements in the product's performance. In services, product innovations can be significant enhancements in the way services are provided, new features that are added to existing services, or the implementation of services that are entirely new to the firm. However, as the changes have to be significant, product innovations exclude minor changes or improvements as well as routine upgrades and regular changes (OECD/Eurostat, 2005).

3.4.2 Process innovation

Process innovation refers to the application or introduction of a new or significantly improved production or delivery method. It can be applied with the intention of decreasing the production or delivery costs per unit, improving the quality of an existing product, or implementing new or improved products. Production methods include routines, equipment, and software solutions, whereas delivery methods involve the firm's logistics and involve equipment, software solutions and routines used in the supply chain and delivery system (OECD/Eurostat, 2005).

It may further include new or significantly improved routines, equipment, and software related to the creation and provision of services as well as in ancillary support activities, such as auditing, maintenance, and procurement (OECD/Eurostat, 2005). Fagerberg (2004) argued that while the introduction of new products is commonly assumed to have a strong and positive effect on the growth of income and employment, process innovation may have a more vague effect due to its cost-cutting nature.

3.4.3 Organizational innovation

The implementation of a new organizational method in a firm's business practices, external relations or workplace organization, is referred to as organizational innovation. Organizational innovations are often aimed at increasing a firm's performance by reducing costs such as administrative or transaction costs, enhance the labor productivity (and thus improve workplace satisfaction), gaining access to non-tradable assets, or reducing supply costs. Organizational innovations differ from other organizational changes in that the implementation of the organizational method is a result of strategic decisions taken by the management, and are, therefore, completely new to the firm.

An example would be the introduction of practices for codifying internal knowledge by, for instance, establishing databases of best practices, making the knowledge more easily available to others within the firm (OECD/Eurostat, 2005). Hence, organizational innovations are strongly linked to administrative efforts of introducing new organizational practices, procedures, and systems, aiming to encourage teamwork, information sharing, training and innovativeness (Gunday et al., 2011).

3.4.4 Marketing innovation

The implementation of a new marketing approach with significant changes in a product's design, packaging, placement, promotion or pricing, is referred to as marketing innovation. These innovations are intended to address customer needs better, expanding into new markets, newly positioning a firm's products on the market to increase the sales of the firm. However, for a change to be a marketing innovation, the implementation of the marketing method has to be part of a new approach that differs from that of a firm's existing marketing methods. Hence, the new marketing approach has to be significantly new to the firm and can be applied to both new and existing products (OECD/Eurostat, 2005).

Examples would be significant changes in the design of a product line of furniture to expand its appeal, or the first use of an entirely different media or practice, such as product placement in movies or television (OECD/Eurostat, 2005). Thus, marketing innovations are related to the four P's of marketing: product or service properties, pricing strategies, product placement and promotion activities Kotler, Armstrong, and Harris (2016).

3.5 Innovation activities

Innovation is a process that may occur across a series of activities initiated by a company aimed at developing specific innovations and may be related to science, technology, organization, development, finance, and business (OECD/Eurostat, 2018).

The innovation process involves not only activities included in R&D , which is a crucial step in the process, but also other key components such as development activities, support activities, capital purchases, and other current innovation-related expenditures. Being able to identify such factors, referred to as innovation activities, is critical for firms to innovate and improve their ability to innovate. These innovation activities are considered as investments, as they are meant to enhance the innovation performance and ultimately provide future returns to the firm. Indeed, such results tend to exceed the specific innovation to which the activity is directed (OECD/Eurostat, 2005).

The intended innovation can either be a continuous, delayed or neglected innovation, and may, thus, need dedicated resources and involvement in certain activities. In addition, how the innovation activities are organized may vary notably between firms. For some firms, an innovation represents an intermediate or final milestone of well-defined innovation projects or programs with specified innovation activities. Other firms aim to make continuous improvements to their products and processes by integrating their innovation activities into the firm's operations, whereas some are engaged in innovation activities mainly on an ad-hoc basis (OECD/Eurostat, 2018).

Innovation activities may be performed within the company, or it may involve the purchase of goods, services or knowledge from external sources. The acquisition of external knowledge or technology, regardless of where the activity is performed, may lead to a specific innovation itself or it may be necessary for the implementation of other innovations. This includes elemental research activities that are not directly associated with the development of a particular innovation (OECD/Eurostat, 2005).

Innovation activities are divided into activities related to R&D and non-R&D , distinguished in the way that R&D activities "result in new knowledge or use of knowledge to devise new applications" (OECD, 2002, §146). Providing quantitative measures of the expenditures spent on each innovation activity is essential to measure the level of innovation activity in firms. Along with output measures, these expenditures are used to calculate returns on innovation activities as proposed in the third edition of the Oslo Manual.

3.5.1 R&D activities

R&D is one of a range of activities that can develop innovations (OECD/Eurostat, 2018). R&D activities include creative work undertaken in a systematical order to increase the knowledge base, including knowledge of the humanity, culture, and society, and to develop new applications of knowledge already available. R&D may have the purpose of attaining either specific or general objectives but is always aimed at new discoveries. Hence, for an activity to be a R&D activity, the discovery has to be creative and original, uncertainty about its outcome has to be present, and it must be a systematic activity where the results could be either freely transferred or traded in a marketplace (OECD, 2015).

When these systematically R&D activities are performed within the firm, it is referred to as intramural R&D, whereas extramural R&D activities refer to the same activities only purchased from external research organizations or other enterprises, including the firm's own division (OECD/Eurostat, 2005).

3.5.2 Non-R&D activities

Innovation activities do not need to be related to R&D in order to generate innovations. Activities that are not systematically initiated to develop an innovation but still contribute to innovation are referred to as non-R&D activities. These activities all aim to increase the firm's performance, which can be done by enhancing the capabilities that allow the development and the implementation of innovations, or the ability to successfully introduce innovations initiated by other firms. The firms' intention of these activities can be to develop or implement new products or processes, new methods of marketing, to sell its products, or to change the organizational method in the firm's practices and structure. That is, non-R&D activities can lead to the company introducing a product, process, marketing, or organizational innovation. However, only activities not already included in R&D , are referred to as non-R&D activities (OECD/Eurostat, 2005).

Non-R&D activities that are related to the development and implementation of product and process innovations include the acquisition of knowledge and technology, preparations as well as training. A firm can purchase external knowledge, such as the rights to use patents and non-patented inventions, licenses, brands, expertise or other forms of knowledge from other enterprises, institutions or consultancy services. External knowledge and technology may also be in the form of capital goods used in the implementation of product and process innovations. Examples of such capital goods are the acquisition of machinery, equipment and computer hardware or software (OECD/Eurostat, 2005).

Non-R&D based activities, associated with marketing and organizational innovations, contribute to the development and implementation of new marketing methods, the introduction of change, or the application of new organizational structures. A firm can, for instance, develop new approaches of marketing and selling of its products and services, or it can reconstruct its policies, process and procedures used, or even its overall business activities (OECD/Eurostat, 2005).

3.5.3 Innovation cooperation

Cooperation refers to the coordinating and implementation of activities in an innovation process, where two or more contributors agree to share information and the responsibility of the process. Two firms cooperate if one of the firms provides a detailed description of what it may need, and acquires any ideas or inputs in return

(OECD/Eurostat, 2018).

The activities an innovative firm engage in depends on the range and the network of its links to the sources of information, technologies, expertise and financial and human resources. These linkages are what relating the innovative firm to other innovative actors, such as laboratories ruled by the government, educational institutions, governmental departments, customers, suppliers, and competitors. Further, these relations may be passive sources, suppliers of knowledge or technology, or cooperative partners and may be associated with product, process, organizational, or marketing innovations (OECD/Eurostat, 2005).

A firm's type of relations is often dependent on the market of the enterprise and the nature of the firm. Mature firms operating in a stable sector are typically driven by the costs and turnover of the inputs and have the suppliers and customers' market signals as their primary connections. Firms that are in a more volatile environment may need a variety of connections to rapidly get as much information and knowledge as possible to adapt to the changing environment (OECD/Eurostat, 2005).

The firms' relations differ by the source of the linkage, the total required investment, and the interaction between the firms. External sources such as, for instance, patent departments or publications regularly provide information at a low expense, while other sources, such as consultancy services, are commonly expensive (OECD/Eurostat, 2005).

3.6 Summary key concepts

Figure 3.1 displays an overview of the key concepts used in this paper in which the research questions are based on and illustrates the link between innovation input, innovation output, and firm performance. The innovation inputs are represented by innovation expenditures related to R&D and non-R&D activities. Further, the innovation outputs are represented by the four innovation types; product, process, organizational, and marketing innovation. Finally, domestic sales are included to serve as a measure of firm performance.

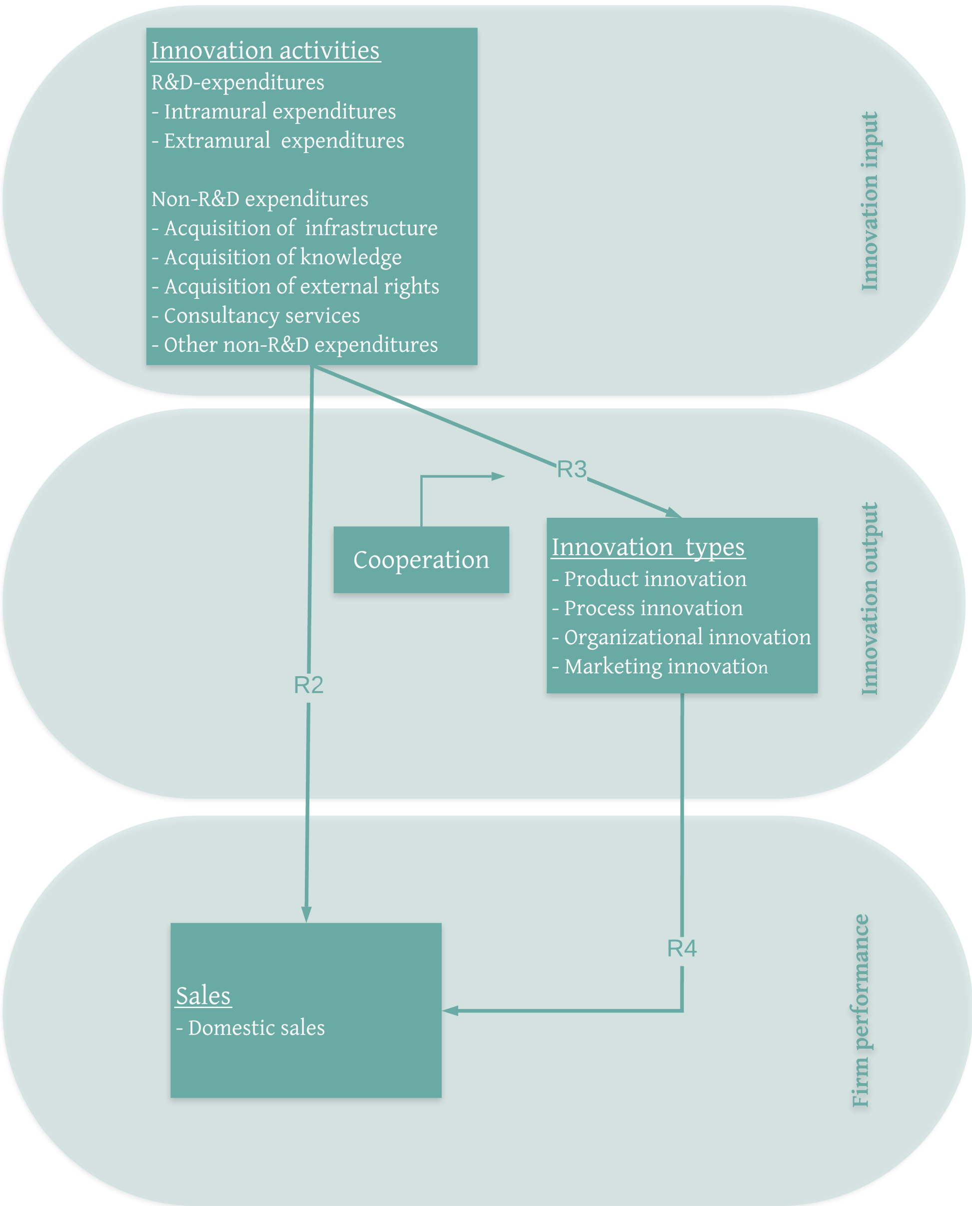

Figure 3.1: Descriptive research framework (OECD/Eurostat, 2005)

Figure 3.2 shows the possible outcomes related to each of the four innovation types. That is, to be classified as a product, process, organizational or marketing innovation, the subsequent outcomes has to be new or significantly improved.

Product innovation		Process innovation			Organizational innovation					Marketing innovation					
Goods	Services	Manufacturing processes	Logistics, delivery, distribution	Auxiliary functions for enterprise's processes	Business procedures	Knowledge management systems	Organization of product development	Workplace organization	External relationships	Design	Packaging	Promotional techniques	Marketing strategies	Sales channels and product placement	Pricing methods

Figure 3.2: Outcomes related to the four innovation types. Reproduced from Danish Ministry of Economic Business Affairs (2008)

Chapter 4

Research design

This chapter is divided into three sections. The first section gives an overview of the variables used in the econometric analysis, along with a discussion of expectations. The second section describes the source and the manipulation process of the data, followed by the third section that presents the relevant theory concerning the econometric analysis.

4.1 Variables description

This section will present all variables included in the econometric analysis in chapter 6, and the explanation will be divided between independent and dependent variables, regardless of the research question's order. The section will provide as a supplement and reference list to clarify the variables used in all regressions, to assist interpretations and discussions altogether.

4.1.1 Independent variables

Innovation types

Product innovation refers to new or significantly improved products or services that are launched by firms during the period under review (OECD/Eurostat, 2005). As discussed under the literature review (chapter 2), numerous scholars have found that product innovation had a positive effect on sales (see for instance; Artz et al., 2010), and it is therefore expected to find a positive relationship in this study.

Process innovation refers to new or significantly improved production or delivery methods that are introduced or applied by firms during the period of analysis. This type of innovation is usually intended at decreasing the unit costs of production or transport, improving the quality of existing products or contribute to the implementation of new or significantly improved products (OECD/Eurostat, 2005). Therefore, a positive relationship with sales is expected, aligned with the positive and significant result found by Klomp and Van Leeuwen (2001).

Organizational innovation refers to the application of new organizational design in the firm's business activities, workplace organization and external relations (OECD/Eurostat, 2005). This type of innovation is often aimed at increasing the enterprise's profitability by reducing administrative or transaction costs (OECD/Eurostat, 2005). Karabulut (2015) found that organizational innovation has a positive and significant effect on various performance measures, such as financial performance, and learning and growth development. However, due to costs not being included in the independent variable concerning domestic sales, a positive but insignificant relationship is expected.

Marketing innovation refers to the application of a new or significantly changed marketing method, related to a product's layout, packaging, promotion or pricing (OECD/Eurostat, 2005). Marketing innovations often aim to increase a firm's sales by better meeting customer needs, by expanding into new markets, or by newly placing a firm's products on the market. Supported by Karabulut (2015), that found a positive and significant relationship on various performance measures; marketing innovation is expected to affect a firm's sales positively.

Innovation activities

Intramural R&D expenditures refers to all expenditures related to R&D activities that are performed within the enterprise. This includes R&D that is directed at the development and application of product, process, organizational, and market innovations, but also R&D that is not targeting a particular innovation (OECD/Eurostat, 2005). In addition, Sharif et al. (2012) found evidence that intramural R&D expenditures are important for all innovation outputs, whereas Artz et al. (2010) found a positive relationship between R&D expenditures and process innovation. Based on this, a positive effect of intramural R&D expenditures on all four innovation types is expected. Further, a positive relationship between intramural R&D expenditures and sales are expected (aligned with Branch (1974)).

Extramural R&D expenditures refers to all expenditures related to R&D activities that are purchased from external organizations. According to Jha and Bose (2016), such R&D activities are usually proceeded by firms that have limited resources or abilities to undertake R&D projects within the firm. As with intramural expenditures, extramural expenditures may be directed toward product, process, organizational or market innovations. It is therefore expected to find that extramural expenditures positively affect the four innovation types (Sharif et al., 2012; Artz et al., 2010), and domestic sales (Branch, 1974).

Acquisition of machinery, equipment, software etc. refers to expenditures related to the acquisition of capital goods such as land and buildings, machinery, equipment and instrument, and software that is not included in R&D expenditures (i.e., intramural and extramural) (OECD/Eurostat, 2005). This is expenditures that are invested with the intention of developing product and process innovations (OECD/Eurostat, 2005), and it is therefore expected to find that this variable positively affects the product and process innovation output. It is further expected to have a positive effect on sales, as innovation activities are aimed at increasing the firm's performance (OECD/Eurostat, 2005). The variable is further referred to as acquisition of machinery.

Acquisition of external rights refers to expenditures associated with the acquisition of external rights that are not related to R&D, with the purpose of developing and applicate innovations. This includes patents, non-patented inventions, brands, licenses, patterns, and designs (OECD/Eurostat, 2005). Such activities are directed toward product and process innovations, and it is therefore expected to find a positive relationship between these variables. As with the acquisition of machinery, a positive effect on firms' sales is expected.

Acquisition of other external knowledge refers to expenditures related to the acquisition of other external knowledge that is not R&D. Other external knowledge may involve computer services, technical or scientific services that are directed toward the implementation of product and process innovations, with the aim of increased sales (OECD/Eurostat, 2005). Thus, the acquisition of other external knowledge is expected to have a positive effect on product innovation, process innovation, and sales. The variable is further referred to as the acquisition of knowledge.

Acquisition of consultancy services refers to the purchase of external professionals that provide the firm with advice within different areas. Consultancy services are often acquired to increase firm performance by developing and implementing product and process innovations (OECD/Eurostat, 2005). Therefore, this variable is expected to have a positive effect on these innovation types as well as on sales.

Other non-R&D activities refers to operating expenditures for innovations that are not R&D , such as activities and technical preparations related to the introduction of new products, workflows, new marketing approaches or production processes (OECD/Eurostat, 2005). Specifically, other non-R&D activities include all expenses related to preparations for product, process, organizational, and marketing innovations. Hence, a positive relationship with each of the four innovation types is expected. Further, as with the other innovation activities, firms undertake other non-R&D activities with the aim of increasing their performance, and a positive effect on sales is therefore expected.

Innovation Cooperation

Cooperation refers to the coordinating and implementing of innovation activities in collaboration, and may include partners like suppliers, customers, universities or private R&D institutes (OECD/Eurostat, 2005). Innovation cooperation is included as an interaction term with each of the innovation activities, to test whether innovation expenditures affect innovation output when the activities are done in collaboration with others.

4.1.2 Dependent variables

Sales

For financial firm performance, domestic sales in Danish enterprises is used. Based on the discussion under the literature review (chapter 2), it is expected to see that innovation (i.e., innovation activities and innovation types) will have a positive effect on sales figures for most of the enterprises. As such, it is interesting to examine whether different innovation activities will affect the enterprises' sales in the years following from such activities is of the essence. Further, there is a general agreement among scholars that a positive and significant relationship between innovation types and firm performance exist, regardless of the performance measure. Therefore, this paper examines whether the implementation of the different types of innovation will affect firms' sales the following year.

Innovation types

The next focus is the four innovation types as dependent variables, rather than independent variables. Thus, product, process, organizational, and marketing innovation will be used to examine whether innovation activities will affect firms to be more innovative. More precisely, whether investing in innovation activities will increase the likelihood of firms introducing one or more innovations in the following year, is considered. In lines with the discussion in the literature review (chapter 2), a positive and significant relationship are expected to be found between several innovation activities and the four innovation types.

4.2 Data

This section is divided into three sub-sections. The first section describes the source of the data used, before the manipulation process of the data is presented. In the third section, summary statistics of the variables related to firms' performance, innovation types and innovation activities are presented. Finally, the limitations regarding the data will be discussed.

4.2.1 Data source

Data used in this thesis is retrieved from Statistics Denmark, which is responsible for collecting, producing and providing relevant and reliable statistics of high quality (DST, 2016d; DST, 2016a). The publicly available statistics and analyzes are commonly used by various research communities, public administration, and international organizations, as a basis of strategic decisions and further research (DST, 2016e). Thus, Statistics Denmark is a highly recognized state institution, which means that the collected data will be considered both reliable and of good quality.

The Innovation Survey

The greater part of the data used comes from the Innovation Survey performed yearly by Statistics Denmark. The survey aims to analyze the extent, type and outcome of the business enterprises innovation and provides information concerning the innovation input and output (DST, 2018). The statistics are obtained via electronic questionnaires and are conducted in accordance with Eurostat and OECD guidelines for innovation surveys as described in the Oslo Manual [1]. Data on individual firms are obtained and then aggregated by the Statistics Denmark into regions, size class and industries defined in the Danish Industrial Classification 2007 (DB07) (DST, 2018).

The Statistics Denmark conducts comprehensive validation of the data through several steps to ensure statistics of good quality. First, the responses are controlled to see whether there are any significant level changes from the previous years. The data is further validated through computer-aided processes, where outliers and errors are controlled for, and the most significant errors are reviewed manually. The data is then compared and validated against information from The Central Business Register and from public accounts of the enterprises. Finally, the last part of the validation process includes imputations and calibrated weighting of missing responses from enterprises above a certain size (DST, 2018).

[1] See sections 3.2 and 3.3 for an explanation of the Oslo Manual and the Innovation Survey

Purchases and Sales by enterprises

The data retrieved from Purchases and Sales by enterprises by Statistics Denmark is used as an indicator for firm performance, complementary to the Innovation Survey data. This Statistics aims to monitor trends and economic activity in the Danish business sector. It provides information on purchases and sales of all firms that are covered by the Danish VAT (Value Added Tax) system. Further, the statistics consist of all enterprises in the business sector with an annual turnover of 50,000 Danish kroner or more, or enterprises that are voluntarily VAT registered. The information is submitted by firms to the Central Customs and Tax Administration in connection with the payment of VAT and is further obtained and published by the Statistics Denmark. The data is then aggregated before it is published by region, size, and industries (DB07) (DST, 2019).

Before the statistics are acquired and published by Statistics Denmark, it is thoroughly validated. First, the data are based on clearly defined VAT declarations reported to the Danish Tax Agency. As this information is important for the VAT payments of the firms and is also inspected by the Danish Tax Agency, the data is considered of high quality. Finally, after the Statistics Denmark have collected the data, it is checked for errors at both industry and enterprise level, and missing data is imputed before it is published (DST, 2019).

Time period

The time period examined in this thesis is from 2009 to 2016. There are several practical reasons for this. First, the statistics from the Innovation Survey in its current form is only comparable from the year 2007, as it was implemented severe quality changes in the survey from the year before. The two primary changes were the response rate and the design, whereas the response went from 47 per cent in 2006 to 90 per cent in 2007, along with a significant improvement in design and usability. Thus, improved quality changes increased the reliability of the statistics. Second, some variables from the Innovation Survey were only available from 2009 due to changes in the estimation process to reduce the measurement uncertainty, hence excluding the two first years (2007 and 2008) from this study (DST, 2018).

4.2.2 Data manipulation

Longitudinal data set

To study the dynamics of innovation, the data from the two statistics, the Innovation Survey and Purchases and Sales by enterprises, were transformed to panel data sets.

In both statistics, a panel of firms was followed over a more extended period, making it appropriate to convert to longitudinal data sets. A motivation for using panel data sets is that it makes it possible to examine the time lags in the explanatory variables. Another motivation is the ability to control for firm-specific effects, that is, individual heterogeneity. Besides, a panel helps to deal with the complications that arise if firms enter and exit over time, or radically change their organizations through rationalizations, mergers, and acquisitions (Mairesse & Mohnen, 2003). The benefits of using panel data are discussed further in the methodology of the analysis (section 4.3).

However, to examine the effect of innovation, little can be done with the Innovation Survey data alone as there are not enough variables collected for all firms. Therefore, to measure the firms' innovation outcome, the two statistics Innovation Survey and Purchases and Sales by enterprises are merged. According to Mairesse and Mohnen (2003), this will contribute to the variables relevance and explanatory power as more independent variables are included in the models. Both statistics are drawn initially from individual firms and then aggregated into industries following Danish Industrial Classifications (DB07) before published by the Statistics Denmark. Finally, the statistics are distributed by industry, size class, and region.

Variable manipulations

The dependent variables are related to the four innovation types and firm performance. The model used to address both the second and the fourth research question, includes a measure of firm performance, **sales**, that refers to the annual domestic sales made by firms in the period 2009 to 2016 and are measured in million Danish kroner. The continuous variable is log-transformed in the statistical analysis [2]. The model related to the third research question includes variables indicating whether the enterprise had launched a product, process, organizational, or marketing innovation in the reference period, noted **prod**, **proc**, **org**, and **mrk**, respectively. The variables concerning the innovation types initially took the value 1 or 0 depending on whether the firm declared it to be innovative or not. The variables are summed up to what is called a dichotomous variable. The final measure if a fractional response variable since the dichotomous value were divided by total enterprises.

To examine the effect of innovation among Danish firms, data on independent variables related to innovation activities and innovation types as input, are gathered. The model considering the second research question includes the expenditures related

[2]The natural logarithm of sales is used to address the problems with skewed data in alignment with Changyong et al. (2014).

to innovation activities, **intra**, **extra**, **nonrd**, **aqknow**, **aqmachin**, **exrights** and **consult**. These variables are firms' estimate of expenditures allocated to various innovation activities, and the continuous variables are measured by million Danish kroner after a log transformation [3]. In the model regarding the third research question, interaction terms with cooperation and innovation activities are included. Cooperation is measured as a percentage of innovative enterprises that cooperated on innovation activities in the reference period. This is also a fractional response variable based on how many companies launched an innovation in collaboration with external partners. Lastly, the model addressing the fourth research question, considers the four innovation types **prod**, **proc**, **org** and **mrk**, however, as innovation input rather than output.

All independent variables are lagged, addressing the potential problems of reverse causality and simultaneity of the data (Phelps, 2010). Note that the innovation expenditures as input are lagged by two periods in the model addressing the second research question, while lagged by only one period in the model considering the third research question. When considered as innovation input, the four innovation types are lagged by one period. The variables are lagged as it takes time for the innovation activities to transform into innovation output and to generate sales. Similarly, the innovation output needs time to turn into sales. Further, lags among expenditures related to innovation activities and performance are commonly seen in innovation research (e.g., Artz et al., 2010; Ho, Tjahjapranata, and Yap, 2006; Hirschey and Weygandt, 1985; Hsu, Chen, Chen, Wang, et al., 2013). This is discussed further in 4.3.

Finally, the model concerning research question three includes control variables to account for individual and time effects. Industry differences are controlled for by using eight industry dummies, following the Danish Industrial Classification 2007 (DB07). The firms are arranged into nine industries, that is, *manufacturing*, *construction*, *wholesale and retail trade*,*transport*, *hotels and restaurants*, *information and communication*, *financial and insurance activities*, *business activities* such as consultancy and travel agencies, and *other industries.* The latter includes companies not covered by the former classifications, for instance, art and entertainment. The industry wholesale and retail trade is further referred to as *trade*, and information and communication is referred to as *communication.* Moreover, seven year dummies are included to account for year specific effects. Finally, table 4.1 provides a description of all treated variables used throughout the analysis.

[3]Gurmu and Pérez-Sebastián (2008) states that a log-transformation of innovation expenditures as explanatory variables addresses the skewed nature of innovation and R&D data.

Table 4.1: Description of variables

Variable	Description
Dependent variables	
Firm performance:[a,c]	
sales	Domestic sales
Innovation types:[d]	
prod	Product innovation
proc	Process innovation
org	Organizational innovation
mrk	Marketing innovation
Independent variables	
Innovation types:[b,d]	
prod	Product innovation
proc	Process innovation
org	Organizational innovation
mrk	Marketing innovation
Innovation activities:[a,b,c]	
intra	Intramural expenditures
extra	Extramural expenditures
nonrd	Other non-R&D activities
aqknow	Acquisition of knowledge
aqmachin	Acquisition of machinery
exrights	Acquisition of external rights
consult	Consultancy services
Innovation cooperation:[b,e]	
coop	Cooperation
Control variables	
id	Firm classification (dummy)
year	Year (dummy)

[a]The natural logarithm of the variables are taken

[b]The variables are lagged by one period

[c]The units of measurement are Danish million in current prices

[d]The units of measurement are per cent of total enterprises

[e]The units of measurement are per cent of total innovative enterprises

Sample

The Innovation Survey is based on a sample of about 4,500 enterprises randomly drawn from the frame population provided by the Business Register, which in 2016 consisted of 13,779 enterprises. Over the period 2009 to 2016, 26 per cent of the targeted business firms in Denmark were on average covered. All enterprises with more than 100 full-time employees are included in the survey each year. The enterprises with less than 100 full-time employees are randomly drawn with an increased likeliness of being chosen in line with the number of employees (DST, 2016c).

Table 4.2 present an overview of the pool of respondents drawn from the targeted statistical population over the eight years, with the correspondingly share of the total observations. The table shows a relatively proportional sample, with a total of 37,243 observations. The share of firms is somewhat stable over time, with the range of 11.6-13.5 per cent of the total observations. However, of the eight years, 2015 was found to have a somewhat higher number of firms, whereas 2010 had the least number of firms.

Table 4.2: Sample size by year (Statistics Denmark)

Year	Number	Share (%)
2009	4545	12.2
2010	4322	11.6
2011	4424	11.9
2012	4698	12.6
2013	4787	12.9
2014	4901	13.2
2015	5044	13.5
2016	4522	12.1
Total	37,243	100

Table 4.3 provides the longitudinal pattern of the Danish enterprises from which the data is collected, distributed by industry and year. The data reveals that the industries are not equally distributed in the sample. However, according to Statistics Denmark, the probability of being selected is higher for enterprises with more full-time employees and more R&D-intensive activities, which will naturally cause some industries to be more represented in the sample. Therefore, even though the industries are not equally distributed, the sample is deemed appropriate. Of the total 37,243 observations, the *manufacturing* industry has 9,643 observations, accounting for 26 per cent of the total observations. The industry *business activity* is also well represented with 8,470

observations, accounting for 23 per cent of the total observations, followed by firms engaged in *trade* and *communication*, with the total number of 6,892 (19%) and 5,036 (14%) observations, respectively.

Moreover, a closer examination of the industries discloses that the *hotel and restaurant* industry had the fewest number of observations each year compared to others, except for 2009 where the construction had fewer observations (81 firms). With a total of 703 observations, the *hotel and restaurant* industry accounted for the lowest share among the industries, with 2 per cent of the total observations. Followed is the *construction* industry that accounts for a total of 1,022 observations, or 3 per cent, over the eight years. Then, companies engaged in *transport* follow with the total number of observations at 1,486 (4%), then firms involved in *financial and insurance*, which accounts for 1,553 total observations (4%), and finally the classification *other industries* such as arts and entertainment with 2,458 observations (7%).

Table 4.3: Sample size by year and industry (Statistics Denmark)

	2009	2010	2011	2012	2013	2014	2015	2016	Obs	Share (%)
Industry										
Business activity	1,038	984	1,029	1,045	1,162	1,156	1,201	855	8,470	23
Communication	793	594	577	585	623	621	673	570	5,036	14
Construction	81	105	97	133	131	139	152	164	1,002	3
Financial and insurance	220	194	197	185	187	193	195	182	1,553	4
Hotels and restaurants	121	67	72	83	78	86	96	100	703	2
Manufacturing	1,014	1,219	1,182	1,233	1,216	1,257	1,263	1,259	9,643	26
Other industries	294	197	295	317	304	332	335	384	2,458	7
Trade	835	808	809	918	885	917	925	795	6,892	19
Transport	149	154	166	199	201	200	204	213	1,486	4
Sample (No. of firms)	4,545	4,322	4,424	4,698	4,787	4,901	5,044	4,522	37,243	100

4.2.3 Descriptive statistics

In panel data, simple descriptive summaries cannot be used as the observations are clustered within entities over time and are not independent of one another (Flint, 2012). Variables can potentially change over time and entities, and the variation that exists for a given entity over time is called "within" variation, whereas the variation between entities is named "between" variation (Cameron & Trivedi, 2010). To exploit the valuable information about the change that exists within and between entities, a method that is specifically designed to summarize panel data (the "xtsum" command in Stata) was used (Flint, 2012). Further, estimators vary in their use of "within" and "between" variation, which makes the distinction between the two types of variation important. In particular, when using the fixed effects model for panel data, it is essential that there exists some within variation in the regressors for the coefficient to be precisely estimated (Cameron & Trivedi, 2010).

Table 4.4 at the end of this section presents a detailed descriptive overview of the variables as they appear in the econometric analysis, separated by dependent and independent variables. The dependent variables are reported in the current period and are divided into firm performance and the four innovation types (i.e., product, process, organizational, and marketing innovation). The independent variables are further divided into innovation types, innovation activities, and innovation cooperation. When reported as independent variables, the four innovation types are lagged by one period, the innovation activities are lagged by both one and two periods, and cooperation is lagged by one period. Only innovation types that are reported in the current period and innovation activities lagged by one period will be commented.

The most critical insight from the summary statistics is the dynamics of the variables. The overall standard deviation describes the variation of all observations, regardless of how they are nested within the industries. This is the standard deviation that would have been reported using a simple summarize command not designed for panel data. "Between" standard deviation refers to the variation across all industries in a given period, while "within" standard deviation refers to the variation within the same industry over time. If the variable is time-independent, its "within" standard deviation is equal to zero, and the statistic is thus a measure to which extent the variable differ over time.

Dependent variables

Looking at the first dependent variable, **sales**, the statistics show an overall low standard deviation relative to its mean of 11.848. Most of the variation in the coefficient is due to a higher variation between industries, suggesting that there is a great deal of persistence.

Innovative firms that had introduced a product innovation in the period had a mean of 17.4 per cent, and firms implemented a process innovation had 21.2 per cent, whereas 29 and 24.3 per cent implemented an organizational and marketing innovation, respectively. The summary statistics of the four innovation types shows that all variables contain quite some variation both between and within the industries. **prod** and **mrk** vary more between industries than over time, while **proc** and **org** varies somewhat more over time than across the industries. This indicates that **prod** and **mrk** are more dependent on the industry, which is expected when looking at the features of the different industries. Some industries are more product and marketing reliable, so that the distribution of product and marketing innovations among Danish enterprises will, naturally, be dispersed. Further, process and organizational innovations are to some extent more dependent on time, which is also expected considering most firms, regardless of the industry, would benefit from these types of innovation. For instance, manufacturing is usually a more product dependent industry, and it is expected to see a higher share of product innovations among the innovative firms.

Independent variables

Looking at means of the independent variables innovation activities, intramural activities, **intra**, has the highest mean of 6.513 followed by extramural activities, **extra**, with mean 5.357. This indicates that firms invest more heavily in R&D activities than in other innovation activities. The innovation activities with the lowest means are the acquisition of knowledge, **aqknow**, and the acquisition of external rights, **exrights**, with 2.204 and 2.489, respectively, indicating that firms invested the least in these types of activities. Further, all variables related to innovation activities vary more between industries than over time, which is expected taking into account that different industries will invest in different innovation activities depending on the targeted innovation. For instance, as it is expected to see more product innovations among the enterprises of the manufacturing industry, it will be natural to expect these firms to invest more in innovation activities that are aimed for product innovations, such as intramural activities and the acquisition of machinery.

Finally, firms that collaborated on their innovation activities had a mean of 0.325. As with the innovation activities, this variable is more dependent on the industry than on time, indicating differences among the innovative firms. The dynamics and the industry differences between the four innovation types and the innovation activities will be further elaborated in the descriptive analysis in chapter 5.

Table 4.4: Summary statistics of all variables (Statistics Denmark)

Variable	Obs	Mean	Overall Std. Dev	Between Std. Dev	Within Std. Dev
Dependent variables					
Firm performance[a]:					
$sales_t$	72	11.848	1.071	1.125	0.083
Innovation types:					
$prod_t$	72	0.174	0.078	0.078	0.025
$proc_t$	72	0.212	0.043	0.027	0.034
org_t	72	0.290	0.048	0.031	0.038
mrk_t	72	0.243	0.066	0.060	0.033
Independent variables					
Innovation types:					
$prod_{t-1}$	63	0.174	0.078	0.077	0.025
$proc_{t-1}$	63	0.213	0.042	0.025	0.034
org_{t-1}	63	0.290	0.049	0.031	0.039
mrk_{t-1}	63	0.242	0.064	0.059	0.031
Innovation activities[a]:					
$intra_{t-1}$	63	6.513	2.643	2.745	0.426
$intra_{t-2}$	54	6.488	2.686	2.792	0.398
$extra_{t-1}$	63	5.337	2.754	2.845	0.524
$extra_{t-2}$	54	5.357	2.718	2.804	0.518
$aqknow_{t-1}$	63	2.204	1.533	1.438	0.694
$aqknow_{t-2}$	54	2.261	1.504	1.412	0.675
$aqmachin_{t-1}$	63	4.938	1.456	1.457	0.453
$aqmachin_{t-2}$	54	4.952	1.488	1.486	0.463
$exrights_{t-1}$	63	2.489	1.949	1.896	0.742
$exrights_{t-2}$	54	2.446	1.893	1.838	0.724
$consult_{t-1}$	63	4.055	1.555	1.576	0.417
$consult_{t-2}$	54	4.084	1.545	1.565	0.412
$nonrd_{t-1}$	63	5.324	1.701	1.754	0.339
$nonrd_{t-2}$	54	5.315	1.705	1.752	0.360
Innovation cooperation[a]:					
$coop_{t-1}$	63	0.325	0.069	0.055	0.045

[a] All variables are log transformed

Number of entities (industries): n = 9

Number of time-points: T = 8

4.2.4 Data limitations

The two data sets used in this is paper are considered of high quality, as they both went through several steps of validation before they were published by the Danish Statistics. However, using data from the Innovation Survey in econometric analysis requires proper and careful handling as specific difficulties may arise when implementing and interpreting the data. There are several possible reasons for why such complications may occur.

Some of the variables used from the Innovation Survey may be subjective, in the sense that the respondents answer with a personal understanding and judgment. Many of the questions in the survey will be difficult to answer by respondents, especially those related to definitions and classification of innovation. In particular, many of the variables that include innovation types will be subjective in nature, for instance, it may be difficult for respondents to know the exact definition of a new or improved product or process (Mairesse & Mohnen, 2003).

Furthermore, the quality of the variables depends on the respondents' answers that are affected by an individual's knowledge and judgment. Usually, the innovation responses are based on accounting figures or internal expertise, but such information will not be as easily available to all firms. Therefore, random errors in the measurement and classification of the variables may occur. In particular, such errors may apply to the quantitative variables of investment in innovation activities, which are often of low quality. Apart from R&D spending that firms are accustomed to report, expenditures associated with new products, processes, organizational methods or marketing methods are rarely registered separately from each other (Mairesse & Mohnen, 2003).

Finally, the Innovation Survey data comes in the form of a cross-sectional data set, which makes it more challenging to cope with endogeneity issues, and to comment on the direction of causality. Many of the variables, especially those related to innovation activities, are strategic decisions that the firms must undertake. Such choices involve decisions regarding the implementation of R&D, acquisition of external knowledge or consulting services, applying for protection of intellectual property rights, or cooperation, and are often determined simultaneously. Moreover, these decisions may also depend on other unknown factors, which are difficult to detect because too few environmental variables are included. Therefore, as recommended by Mairesse and Mohnen (2003), a panel data set is constructed to carry out an appropriate analysis of the causal relationships between the variables.

4.3 Methodology

The methodology section will present two quite different theoretical frameworks to conduct econometric analysis on panel data. Traditional linear methods will be described first before the Generalized Estimating Equations (GEE) follows. As explained in the description of the variables (section 4.1), this thesis consists of several regressions, where the dependent variables possess unique characteristics. Traditional linear panel data methods will be applied to handle the continuous variable used as firm performance, whereas the GEE addresses the complicated nature of the innovation types, to be exact, product, process, organizational, and marketing innovation.

4.3.1 Traditional linear methods

This thesis analyzes the relation between innovation activities, innovation types and firm performance. Historically many relating studies were done with cross-sectional data from one specific Community Innovation Survey (CIS). However, as more innovation surveys were produced over the years, longer series of data were available to analyze, and with the development of modern econometric tools, panel data escalated rapidly in popularity. Traditional data with cross-sections varying over time is commonly referred to as panel data or longitudinal data. This is data with repeated observations over time for the same entity, thus being a combination of cross-sectional and time-series data. A panel data with k regressors is notated as:

$$(X_{1it}, X_{2it}, ..., X_{kit}, Y_{it}), \quad i = 1, ..., n, \quad t = 1, ..., T \tag{4.1}$$

where, n = number of entities (as industries), and T = number of time periods (as years)

Panel data is often split between two distinct types, namely, balanced and unbalanced. A balanced panel is known to have no observations missing, and all variables are observed for all entities (industries) and all time periods (years), otherwise noted as unbalanced. This paper deals exclusively with a perfectly balanced panel set, hence not discussing the characteristic differences between the two types.

Given the fact that panel data contains both elements of multiple entities observed over two or more points in time, the application of the regression model following the econometric theory is more complex than the separate methodologies. However, there are several reasons why the application of panel data is widely used. Baltagi (2008) summarized seven benefits:

1. The possibility to control for individual heterogeneity. The risk of obtaining biased results decreases substantially as panel data suggests that entities and individuals are heterogeneous, in contrast to time-series and cross-sectional studies.

2. Panel data enables management of more complex data. It provides more informative data, more variability, less collinearity among the variables, more degrees of freedom and more efficiency.

3. Another important advantage is that panel data enhances the ability to study the dynamics of adjustment.

4. Pure cross-sectional and time-series data are more limited than panel data, hence they are not suited to detect and measure effects that the panel data can.

5. Also, the models used for panel data let us compose and test more complicated behavioral models than in cross-sectional or time-series data.

6. Micro panel data is often more precisely measured than similar variables measured at the macro level. However, biases from aggregation may be reduced or eliminated.

7. Macro panel data, in general, have a longer time span. Besides, panel unit root tests have standard asymptotic distributions, whereas the unit root test in the time-series analysis has the problem of nonstandard distribution.

The most prominent advantage econometricians favor when using panel data, is the possibility to control for individual heterogeneity. First of all, pooled ordinary least squares estimation for panel data, and how the method fails to account for industry-specific heterogeneity is introduced. The following subsections will include a discussion of models that specifically manages individual heterogeneity, that is, random and fixed effects models.

Pooled OLS

Pooled OLS estimation is simply an Ordinary Least Squares (OLS) technique performed on panel data. It is beneficial to begin defining a panel data model linear in parameters as

$$y_{it} = \alpha_i + \beta_1 X_{it1} + ... + \beta_k X_{itk} + u_{it} \tag{4.2}$$

where α_i is the intercept given as the industry-specific effects, and u_{it} is the two-way error term. The combination of α_i and u_{it}, is usually referred to as the composite error

(Wooldridge, 2003). u_{it} consists of $\mu_i + \lambda_t + \nu_{it}$, where μ_i is the unobservable individual effect, λ_t is the the unobservable time effect, and ν_{it} accounts for the remainder stochastic disturbance term (Baltagi, 2008). In this type of linear format, a random sample from the cross section and no perfect collinearity is anticipated. Wooldridge (2003) continues that the error term, often called idiosyncratic or time-varying error, is normally assumed to be **strictly exogenous**, given by

$$E(u_{it}|X_i, \alpha_i) = 0, \tag{4.3}$$

homoskedastic, thus

$$var(u_{it}|X_i, \alpha_i) = \sigma_u^2, \tag{4.4}$$

and with no **serial correlation**:

$$cov(u_{it}, u_{ir}|X_i, \alpha_i) = 0, \tag{4.5}$$

Pooled models assume that the independent variables are exogenous, thus writing the error term as the traditional stochastic disturbance term ν_{it} rather than using the time and individual effect characterizing panel data. Equation (4.2) is then written as,

$$y_{it} = \alpha + \beta X'_{it} + \nu_{it} \tag{4.6}$$

In equation (4.6), the X_{it} does not include a constant, whereas in equation (4.2), X_i would additionally include a constant term (Cameron & Trivedi, 2010). The OLS estimation in itself is fairly uncomplicated, but it eventually should control for correlation of the error over time for each individual (within correlation) and potential correlation over individuals (between correlation). Consequently, there are mainly two reasons why pooled OLS is usually not consistent. Firstly, the fixed effects terms can be correlated with the regressors, and secondly, and the error terms are often correlated. As a result, both the assumptions regarding exogeneity and uncorrelated observations are violated (Wooldridge, 2003).

Individual-effects models

The individual-specific-effects model for the scalar regressand y_{it} is defined as,

$$y_{it} = \alpha_i + \beta X'_{it} + u_{it} \tag{4.7}$$

where X_{it} are regressors, α_i are random individual-specific effects, and u_{it} is an idiosyncratic error (Cameron & Trivedi, 2010).

There are two rather different models for the α_i, that is fixed-effects and random-effects models. However, deciding whether or not an individual-effects approach or a pooled OLS approach is appropriate is essential. A Breusch-Pagan Lagrange multiplier test decide between an individual-effects model or a pooled OLS model (Baltagi, 2008). The null hypothesis in this test is that variances across clusters are zero, which means that there is no panel effect and no significant difference across entities (industries). A significant test, at the five per cent level, leads to the discussion between the two following individual-effects models.

Fixed Effects

The fixed effects (FE) model is an approach to eliminate the fixed effects present. Fixed effects estimation come from the basic assumption that the unobserved heterogeneous components are constant over time (Baltagi, 2008). This permits the covariates to be correlated with the time-invariant part of the error term, by tolerating a limited form of endogeneity. Fixed effects estimation can be exercised in two ways, namely the Within Group estimator and the Least-Squares Dummy Variables (LSDV) model, whereas the former method is applied in this thesis.

This specific type of model permits the α_i in equation (4.7) to be correlated with the independent variables X_{it}, thus allowing the limited form of endogeneity. As the estimating of μ_i and λ_t is anticipated to be fixed, and the tradtional error term, (ν_{it}), is identical independent distributed (i.i.d.), then $u_{it} = \mu_i + \lambda_t + \nu_{it}$ represent a two-way fixed effects error component model. Since the fixed effects term, α, is constant in time, it is possible to subtract a time-average from all variables on the normal linear model to remove the fixed effect (Baltagi, 2008). Thus subtracting,

$$\bar{y}_i = \alpha_i + \beta_1 \bar{X}_{i1} + ... + \beta_k \bar{X}_{ik} + \bar{u}_i \tag{4.8}$$

on equation (4.2) gives,

$$(y_{it} - \bar{y}_i) = \beta_1 (X_{it1} - \bar{X}_{i1}) + ... + \beta_k (X_{itk} - \bar{X}_{ik}) + (u_{it} - \bar{u}_i) \tag{4.9}$$

As a result, the fixed effects term is removed and the equation can be estimated by applying a pooled OLS model. Following the assumptions specified in chapter 4.3.1, the estimator is now unbiased.

Random Effects

On the other side, the random effects (RE) model assumes that α_i in equation (4.7) is entirely random. This provides a sharper assumption considering that α_i is uncorrelated with the independent variables (Cameron & Trivedi, 2010). This relationship states that the fixed effects α_i are uncorrelated with the independent variables, and are explained by,

$$cov(X_{it}, \alpha_i) = 0 \tag{4.10}$$

However, given this relationship, the pooled OLS method examined in chapter 4.3.1 would also be unbiased. Further, the same chapter mentioned the issue with serial correlation, which would still be existing in this case. The pooled OLS will produce a biased and inefficient outcome since the variations would be overestimated. Nevertheless, a 'quasi-demeaning' transformation could be utilized to achieve conditionally uncorrelated observations and eliminate the serial correlation (Wooldridge, 2003);

$$y_{it} - \lambda\bar{y}_i = \beta_1(X_{it1} - \lambda\bar{X}_{i1}) + ... + \beta_k(X_{itk} - \lambda\bar{X}_{ik}) + (\nu - \lambda\bar{\nu}_i) \tag{4.11}$$

where,

$$\lambda = 1 - \sqrt{\frac{var(u_{it})}{var(u_{it}) + Tvar(\alpha_i)}} \tag{4.12}$$

Conclusively, this equation has now an abscense of serial correlation, and as the fixed effects model, the equation is transformed and could be estimated with an OLS approach.

Choosing between the individual-effects models

In conclusion, the random effect model, contrary to the fixed effect, tolerate independent variables that are constant in time. Although the random effect estimator is most accurate, the assumption made in equation (4.10) is extremely restrictive and not valid in many circumstances. Consequently, which individual-effects model to use should rely on whether or not this assumption holds (Wooldridge, 2003). The random effects model, as discussed above, is consistent when this presumably holds and is a superior method compared to the fixed effect method and vice versa. This specific assessment is generally referred to as the Hausman test (Hausman, 1978).

On the other hand, the Hausman test exhibit a major limitation. The test primarily needs an efficient random effect estimator, which require that α_i and u_{it} are i.i.d. This assumption is neglected if cluster-robust standard error for the random effects estimator is considerably different from the default standard errors. Very often a robust version of the Hausman test is necessary (Cameron & Trivedi, 2010).

Table 4.5: Overview: Fixed effect versus Random effects

	Fixed Effect Model	Random Effect Model
Functional form	$y_{it} = (\alpha + u_i) + \beta X'_{it} + \nu_{it}$	$y_{it} = \alpha + \beta X'_{it} + (u_i + \nu_{it})$
Assumption	-	Individual effects are not correlated with regressors
Intercepts	Varying across group and/or time	Constant
Error variances	Constant	Randomly distributed across group and/or time
Slopes	Constant	Constant
Estimation	LSDV, Within effect estimation	Generalized Least Squares (GLS)
Hypothesis test	F test	Breusch-Pagan LM test

Reproduced from Park (2011)

4.3.2 Generalized estimating equations

Generalized Estimating Equations (GEE) is a marginal model frequently used for longitudinal and clustered data analysis in medical and social science studies (Hardin & Hilbe, 2013). The model extends the Generalized Linear Models (GLMs) by Nelder and Wedderburn (1972) to fit the modeling of correlation data. Earlier, econometricians would try to fit a linear model to data that came in clusters, as panel data, and afterward adjust for the standard errors to handle the clustering. However, this post-hoc method does not affect the parameter estimates in the regression. Hence, the GEE should be utilized to account for this type of estimation bias, ultimately providing several advantageous features.

The first noteworthy feature is that the variance-covariance matrix of responses is regarded as nuisance parameters in GEE. Therefore, it becomes easier to fit, and the model is usually favored if the overall treatment effect is the main concern (McCullagh & Nelder, 1989). Second, the parameter estimates are consistent and asymptotically normally distributed under moderate regularity conditions. This is valid even when the “working” correlation structure of responses is wrongfully stated, and the variance-covariance matrix is estimated by robust standard errors. Third, GEE only requires true measures of marginal mean and variance in addition to the link function [4], thus loosening the underlying assumptions (Wang, 2014).

[4] Introduced under ”Link function and distribution” on page 55

The GEE approach is divided between two distinct options, either subject-specific (SS) or population-averaged (PA) models. The former specifies a model of the source of heterogeneity where the coefficients have an interpretation for each individual, meaning it is subject specific. The latter (PA) includes the within-panel dependence by averaging effects over all panels. This means that the marginal outcome of individuals is measured, changing the coefficients to be interpreted as the response averaged over the population. GEE-PA assumes a common correlation of the repeated subject measures. Hence, this method is frequently used when there are several individual-specific responses, but when each subjects' pattern of responses is not important. GEE-PA will be further referred to as GEE.

In conclusion, the GEE is a population-level approach based on a quasi-likelihood function and provides the population-averaged estimates of the parameters (Wedderburn, 1974). The main reason for introducing this model is due to its application to fractional response models, which is the case with several dependent variables in this research (see for instance Wu, 2012; Phelps, 2010). Fractional response models is a part of the generalized estimating equations and are naturally restricted between 0 and 1, which raises issues relating to inference and functional form (Papke & Wooldridge, 2008). GEE models also consider the within-subject correlation, which reduces the variance of the parameters and usually indicate a too large measure of significance. Three elements are needed to construct an adequate GEE model; the specification of a link function, the distribution of the dependent variable, and the correlation structure of the dependent variable (Liang & Zeger, 1986). Link function, distribution and correlation structure, as well as diagnostics test will be further explained in the following subsections.

Link function and distribution

The link function and the distribution family of the dependent variable is an important aspect of the GEE. This is easily illustrated by setting up a generalized linear model of the relationship between y_i and the covariates X_{it}:

$$g\left\{E(y_{it})\right\} = X_{it}\beta, \quad y \sim \text{ F with parameters } \theta_{it} \tag{4.13}$$

for $i = 1, ..., m$ and $t = 1, ..., n_i$, with n_i observations for each group identifier i. g($\cdot$) is called the link function, and F is the distributional family. Substituting various definitions for g($\cdot$) and F results in a wide array of models (Liang and Zeger, 1986; StataCorp, 2017). The fractional response variables originates from a dichotomous variable, thus being binary in nature. The binomial distribution is therefore appropriate, leaving the choice of link function in this specific regression analysis to be either; logit, probit, log or idenity.

Correlation structure

The working correlation matrix specifies the within-group correlation structure and is an important aspect of estimating the GEE (Hardin & Hilbe, 2013). The most simple correlation structure belongs to the independence model, which assumes no correlation within entities and excludes additional parameters in the estimating equation. The working covariance matrix is merely the identity matrix to the model. The autoregressive structure (AR) differ from the independent version as it assumes a temporal dependence within entities. The AR structure could be augmented to include several lags of the dependent variable, with the number following the AR-term denoting the number of lags used (e.g., AR1). α is the only parameter of this type of structure and shows that the level of correlation depends on the distance over different periods. Observations with few lags are presumably more correlated than those with more lags. Other possible correlation structures include exchangeable, stationary, nonstationary and unstructured, however, will not be the focus as it not applied in the analysis.

QIC

There are several criterion measures suitable to evaluate econometric models with panel data. The Akaike Information Criteria (AIC) is one of the most well-established goodness-of-fit statistics for likelihood-based model selection in general (Hardin & Hilbe, 2013). This is characterized by $AIC = -2\mathcal{L} + 2p$, specifying that $\mathcal{L}$ is the log likelihood and p is the number of parameters present. However, it is necessary to generalize the AIC into a measure for quasi-likelihood models, such as the GEE method. Thus, an extension of the AIC is the Quasi-likelihood Information Criterion (QIC), which is the quasi-likelihood under the independence model information criteria (Hardin & Hilbe, 2013). QIC is noted as,

$$QIC(R) = -2Q(g^{-1}(x\beta_R)) + 2trace(A_I^{-1}V_{MS,R}) \tag{4.14}$$

The first part of the equation on the right-hand side $(-2Q(g^{-1}(x\beta_R))$ is the value of the quasi-likelihood calculated, including the chosen correlation structure R, and the inverse of the link function (g^{-1}). The second part defines A_I as the variance matrix under the independence model, and $V_{MS,R}$ as the sandwich estimate of variance under the hypothesized correlation structure R (Hardin & Hilbe, 2013).

Chapter 5

Descriptive analysis

This chapter will present the characteristics of Danish business enterprises, based on the yearly Innovation Survey from Statistics Denmark. Moreover, it will provide a better insight into the relationship between innovation input, innovation output, and firm performance. This chapter is divided into two sections. First, an overview of the four innovation types (i.e., product, process, organizational, and marketing innovation) will be presented, including an examination of the distribution of the innovation types across industries. The second section provides an overview of the innovation activities divided by industry and years, followed by a descriptive analysis of innovative companies and cooperation. This will be discussed in terms of the distribution among industries and types of partners most prevalent.

5.1 Innovation types

An overview of innovative enterprises in Denmark is presented in figure 5.1. It shows the total innovative firms, that is, firms that have launched one or more innovations in the respective year, together with the distribution of the types of innovations introduced. Looking at the period 2014 to 2016, 44 per cent of the Danish business enterprises introduced one or more types of innovation. The proportion of innovative firms is thus unchanged compared to the previous three years and has not changed significantly during the entire period examined.

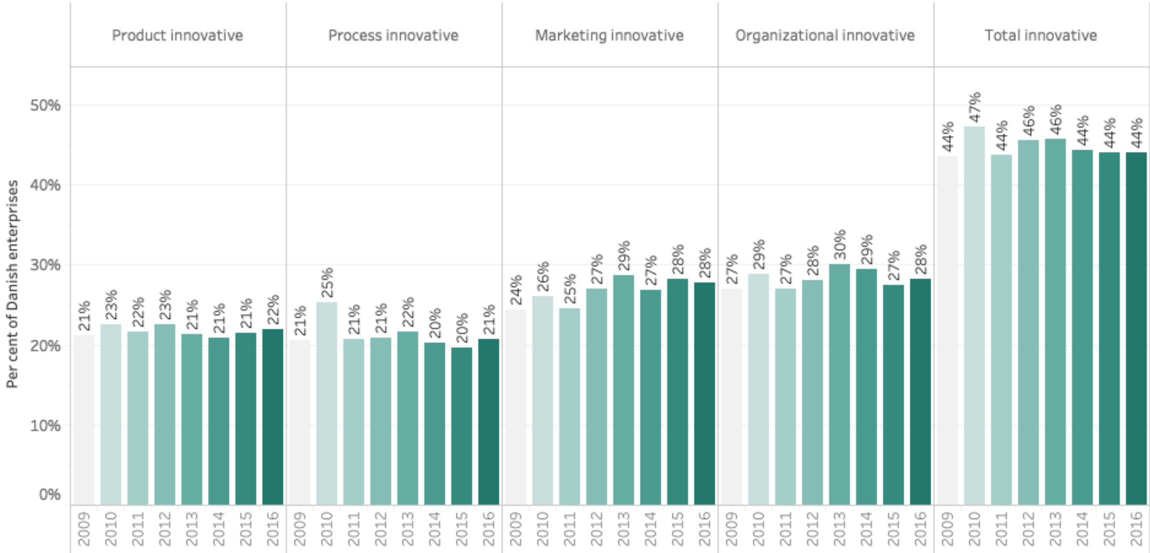

Figure 5.1: Innovative firms broken down by innovation types and year (Statistics Denmark)

Considering the four innovation types, neither have drastically changed throughout the analysis, supporting the findings in the descriptive statistics (subsection 4.2.3). About 21 per cent of the Danish enterprises have introduced new products, either in the form of new goods or new services, and firms that have introduced new processes in their business activity has been somewhat stable over time. Further, firms implementing a new organizational method has increased with about 1 per cent point from 2009 to 2016, whereas firms with a marketing innovation have increased with around 4 per cent point in the same period.

About every fifth firms in Denmark launched new or significantly improved products in the years 2009 to 2016, and with the exception of 2010, an almost equal share introduced new processes. Further, 28 per cent of the Danish firms introduced a new marketing method in 2016 that involves changes in the product's design, packaging, pricing or promotion. Finally, the same share of enterprises introduced a new organizational method in their business activities, workplace or external relations in 2016.

Looking at the innovative firms broken down by industry in figure 5.2, there is a significant variation in the different industries' innovativeness. As discussed in the descriptive statistics (subsection 4.2.3), the most significant industry differences are present among product and marketing innovative firms, whereas less variation exists for organizational and process innovative firms. However, in most industries, organizational and marketing innovations are the most widespread, whereas product and process innovative firms are less prevalent.

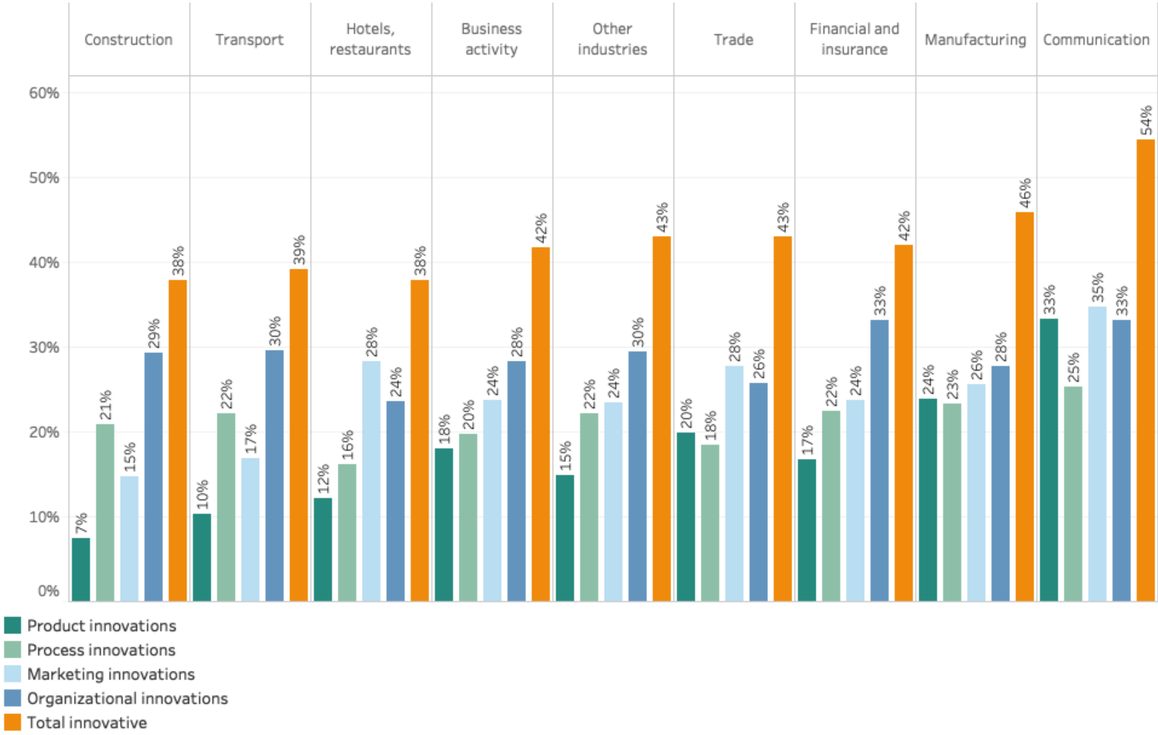

Figure 5.2: The average per cent of innovative firms broken down by innovation types and industries (Statistics Denmark)

It is clear from that *communication* is the most innovative industry on average, with 54 per cent of Danish business enterprises having launched one or more types of innovations. In addition, this industry has a high share of innovative firms within all four types of innovation. Furthermore, *manufacturing* is the second most innovative industry (46%), followed by *trade* and *other industries*, with 43 per cent. Companies that are engaged in *finance and insurance* and *business activities* both follow with 42 per cent.

For the *manufacturing* industry, the share of innovative enterprises is quite uniform across the types of innovations with about 23 to 28 per cent of the total manufacturing firms. Within the *trade* industry, marketing innovations followed by organizational innovations were the most regular innovation types introduced, with 28 and 26 per cent respectively. The introduction of new products or processes, however, were in contrary slightly lower during the period examined.

For the category *other industries*, organizational innovations occurred most frequently among the companies, with 30 per cent of the innovative firms. Market innovations and process innovations accounted for 24 and 22 per cent respectively, and on average 15 per cent of the companies introduced new products during the period.

Furthermore, 33 per cent of the innovative firms engaged in *finance and insurance* were organizational innovative, while 17 per cent introduced new or improved products.

Marketing and process innovative firms accounted for 24 and 22 per cent, respectively. Further, most of the innovative enterprises applied new organizational methods (28%) in the business activity industry. On average, 24 per cent of the firms implemented new marketing strategies during the period, 20 per cent introduced new processes, while 18 per cent launched new products or services.

The least innovative industries were *hotels and restaurants* and *construction*. In both industries, on average 62 per cent of the enterprises did not launch any innovation type. Out of the 38 per cent of innovative firms concerning *hotels and restaurants*, most of the enterprises introduced new marketing approaches (28%) and new organizational designs (24%). Product and process innovations were the least implemented innovations, with 12 and 16 per cent, respectively. Further, most of the innovative enterprises classified under *construction* were organizational innovative, whereas only 7 per cent were product innovative. Similarly, within *transport*, 61 per cent of the firms were not innovative, but out of the innovative enterprises, 30 per cent had introduced a new organizational method.

5.2 Innovation activities

Figure 5.3 presents an overview of the total expenditure related to innovation activities in Denmark. It shows the total amount invested in Danish kroner (DKK) concerning each innovation activity, in the respective year.

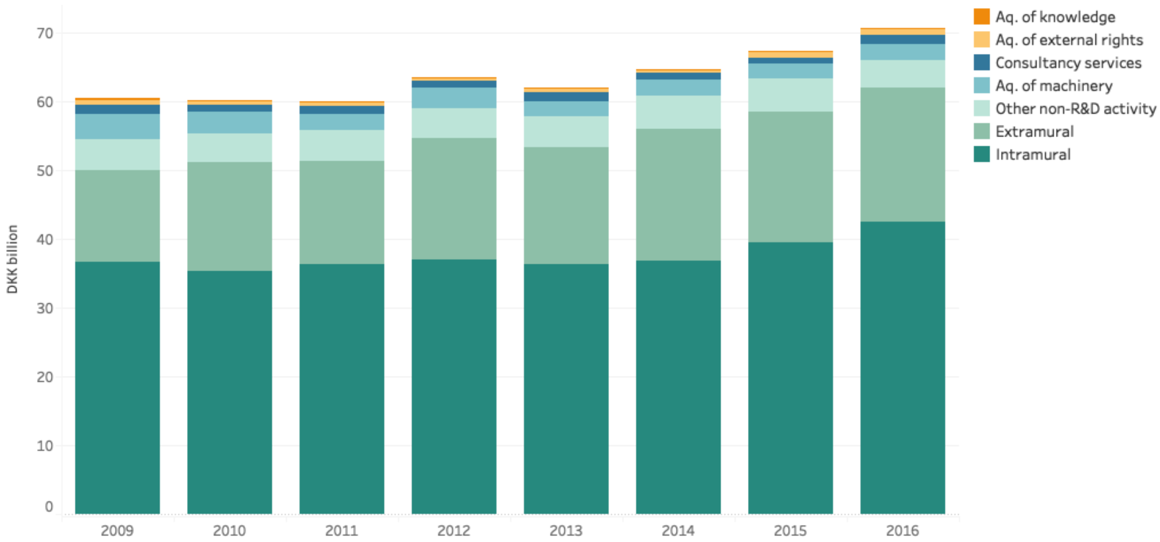

Figure 5.3: Total expenditures related to innovation activities broken down by year (Statistics Denmark)

In the period 2009 to 2016, the total expenditures related to innovation activities has steadily increased, and the distribution of the various activities has been relatively similar. This is in line with the findings in the descriptive statistics (subsection 4.2.3). In 2009, Danish firms spent 60.48 billion on innovation activities, while the total investment in 2016 was 70.69 billion, corresponding to an increase of about 17 per cent and an increase of 5 per cent compared to the previous year. Figure 5.3 further shows that in Denmark, R&D activities (i.e., *intramural* and *extramural* activities) were more costly than innovation activities non-related to R&D . Thus, the enterprises invested the most in R&D activities performed "in-house", followed by R&D activities purchased from other external organizations.

In 2016, Danish enterprises had total *intramural* expenditures equivalent to 42.42 billion, corresponding to 60 per cent of the total innovation expenditures, whereas *extramural* expenditures had a share of almost 30 per cent or 19.61 billion. Further, 3.86 billion (6%) of the total innovation expenditures was used to *other non-R&D* related activities, such as activities and preparations related to launching new products or workflows. A total of 2.42 billion (3%) was in 2016 spent on *acquisitions of machinery*, equipment, and software, partly used for the production of new products and new processes. Further, expenditures related to purchases of *consultancy services* accounted for 1.39 billion (2%) of the total innovation expenditures. Purchases of *external rights*, such as patents, brands, and designs, and *acquisitions of knowledge*, accounted for 0.91 and 0.08 billion, respectively, which together accounted for less than 1.5 per cent of the total innovation expenditures.

In order to get a more detailed overview of the investment patterns among the various industries, figure 5.4 shows the innovation-related expenditures as an average share of total sales in 2009 to 2016, broken down by industry. The figure further shows the proportion of expenditures allocated for each innovation activity.

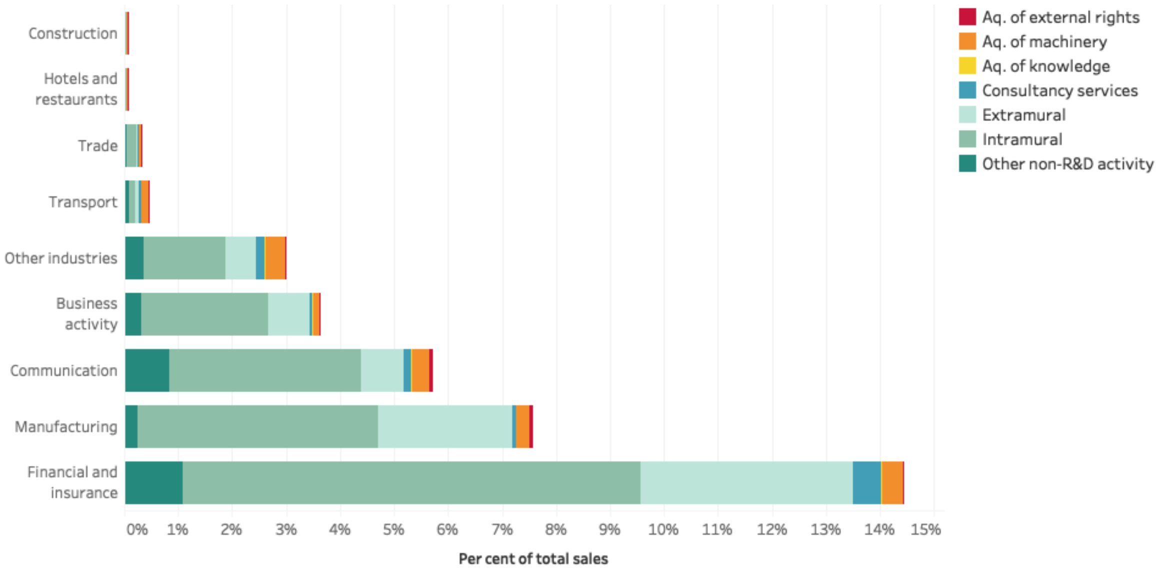

Figure 5.4: Investment in innovation activities as an average share of total sales, broken down by industry (Statistics Denmark)

The figure above shows that *finance and insurance* were the industry that spent the most on innovation to total sales. Further, the firms invested on average 14 per cent of their total sales in activities related to innovation. The *manufacturing* industry invested the second most in innovation activities, followed by *communication*, with 8 per cent and 6 per cent, respectively. The industries that spent the least on innovation activities relative to sales were *construction*, followed by *hotels and restaurants*, then *trade*, and finally *transport*, all of which invested less than 1 per cent of the respective industry's total sales.

In addition, a closer inspection of the investment patterns among the industries shows apparent industry differences in the expenditures related to the innovation activities. The *financial and insurance* industry spent relatively more than the other industries on all activities, except for the acquisition of knowledge and acquisition of external rights, which was exceeded by *other industries* and *manufacturing*. Furthermore, the most significant industry differences are present in expenditures related to intramural and extramural activities, supporting the findings in the descriptive statistics (subsection 4.2.3). Firms engaged in *financial and insurance* invested on average 8 and 4 per cent in intramural activities and extramural activities, respectively. In contrast, the *manufacturing* industry, invested on average 4 and 2 per cent of the total sales in such R&D activities.

Innovation cooperation

Figure 5.5 presents an overview of the most cooperative industries regarding innovation activities. The industries are displayed as the average share of the innovative firms that carried out innovation activities in collaboration with external partners during the period 2009 to 2016.

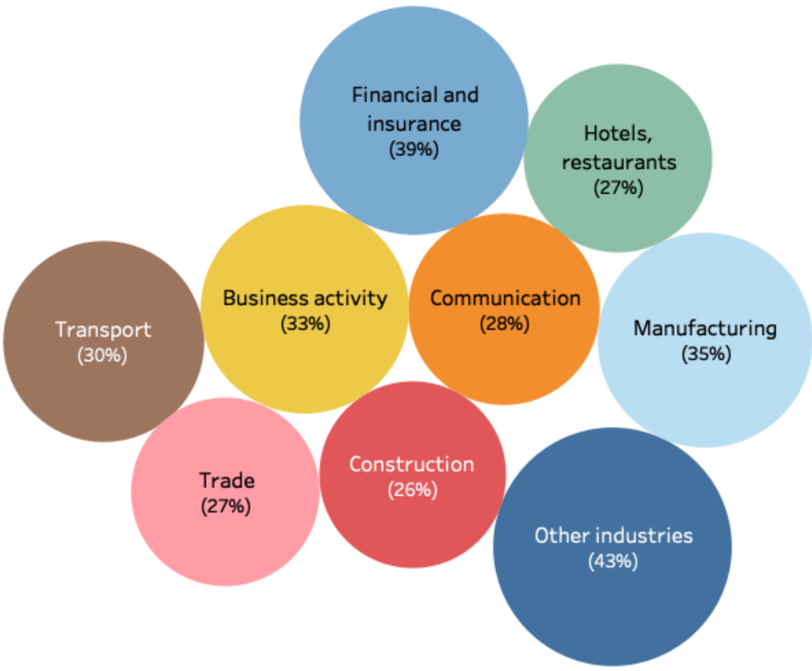

Figure 5.5: Innovative enterprises with innovation cooperation concerning innovation activities (Statistics Denmark)

Among all the innovative enterprises, 32 per cent were on average cooperating on their innovation activities during the examined period. Further, by looking at the figure, it is clear that there exist some differences between the various industries. In particular, firms in the classification *other industries* were more likely to cooperate on the innovation activities, and on average 43 per cent of the innovative firms reported that they had cooperated in the period. Also, the industries *financial and insurance* and *manufacturing* had on average a strong demand for collaboration, with 39 and 35 per cent, respectively. Among the other industries, marginal industry differences were present (between 26%-33%) in the share of innovative firms that had undertaken innovation activities in collaboration with one or more partners. Nevertheless, the innovative firms that cooperated the least were the firms engaged in *construction*, with 26 per cent of the firms, followed by those engaged in *trade* and *hotels and restaurants* (27% for both industries).

Furthermore, firms often cooperate with more than one partner, where the type of partners is usually dependent on the market and the nature of the firm. As discussed in the key concepts (subsection 3.5.3), there are differences between mature firms operating in a stable sector and firms in a more volatile environment. For instance, suppliers and customer's market signals are usually the primary partners of mature firms, as the firms are more driven by the costs and turnover of the inputs. Figure 5.6 presents an overview of the most prevalent partners among innovative firms, to further examine the characteristics of the Danish enterprises. The various partnerships are presented as an average share of the total innovative enterprises that have cooperated concerning their innovation activities.

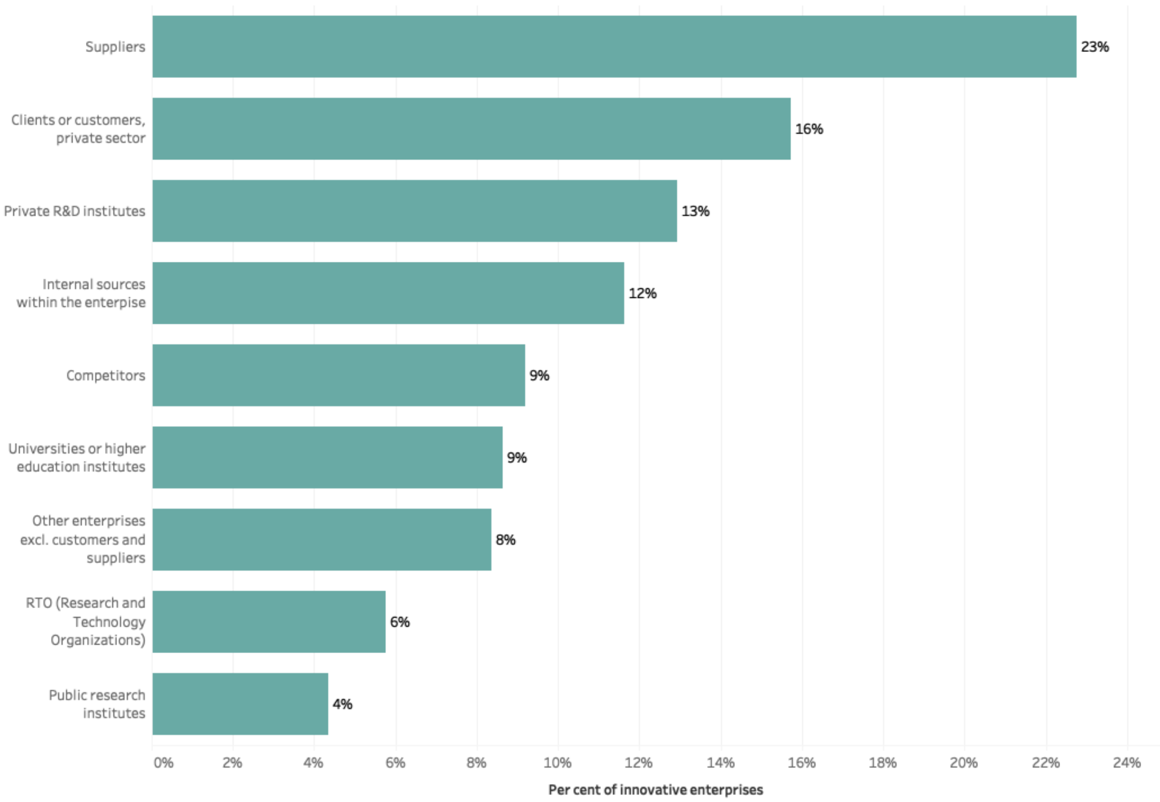

Figure 5.6: Innovative enterprises with innovation cooperation broken down by partners (Statistics Denmark)

It is clear that Danish enterprises primarily collaborated with their suppliers of equipment and software in the innovation activities, with 23 per cent of all innovative firms. Also, clients or customers from the private sector were often reported as the leading innovation partners, with correspondingly 16 per cent. This is closely followed by private R&D institutes and internal sources within the enterprise that accounted for 13 and 12 per cent of the firms' collaboration partners, respectively. In contrast, merely 4 per cent of the firms reported public research institutes as their collaboration partners, making this the least prevalent partnership among the innovative firms.

Chapter 6

Econometric analysis

The econometric analysis concerning research question two, three and four will be investigated in this chapter and will follow the sequential order stated in the problem statement (section 1.2). Thus, this chapter will make the foundations for the discussion regarding how innovation activities affect the implementation of innovation types, and, how innovation affects firm performance. All three sections are structured equally; beginning with a model specification to find the appropriate model, before the results of the specified model is presented. The econometric models will be estimated on a balanced data panel covering nine industries from 2009 to 2016.

6.1 Innovation activities and sales

This section addresses how expenditures to innovation activities affect sales for Danish enterprises. The model will first of all be constructed and justified, before the result of the second research question is presented in 6.1.2.

6.1.1 Model specification

The dependent variable, *sales*, is a continuous variable well suited to be analyzed with the traditional linear approaches in panel data. As explained in the methodology (section 4.3), the approaches is divided between three main methods; the pooled OLS, fixed effect, and random effects. It is essential to assess the appropriate model before executing the regression and interpreting the results.

A one-way effect model is composed if either the entity or the time variable is considered, that is industry and years, whereas a two-way effect model includes two dummy variables, both for time and entity (Baltagi, 2008). Both individual and time fixed effects were introduced in section 4.3 for illustrative reasons, noted μ and λ, respectively. Nevertheless, the regression in this section will be treated as a one-way model with only industries as the primary consideration, as there is no direct cause to include time dummies as regressors. The reason being that *sales* exhibit a persistent trend within each entity, presented in the descriptive statistics in chapter 5, having a low standard deviation and without significant fluctuations. Thus, no extraordinary circumstance suggests that year dummies would be beneficial to include in the model, and hence the time fixed effects, λ, will be disregarded as a predictor.

The predictor variables, more specifically, innovation activities, are lagged by two periods. All independent variable will be lagged throughout this thesis to account for time effect, concerns of reverse causality and simultaneity (Phelps, 2010). Expenditures to innovation activities are considered a first move toward implementing a product, process, organizational, or marketing innovation while introducing one of these innovation types is regarded as the first move toward firm performance. As innovation activities to innovation types are lagged by one period (introduced in section 6.2), and innovation types to firm performance are also lagged by one period (section 6.3), it is rational to assume a two periods delay from innovation expenditures to firm performance [1]. Lags among expenditures and firm performance are frequently used in innovation studies (Artz et al., 2010; Ernst, 2001; Langowitz and Graves, 1992; Kondo, 1999).

Panel data tests

Several tests are designed to choose between the methods. In general, there are some acknowledged ways to test specifically for fixed, random and OLS characteristics. Fixed effects are usually tested by a standard F-test against the residuals from the pooled OLS, and random effects are normally observed by the Lagrange multiplier (LM) test, with the latter being the most applied one (Breusch and Pagan, 1980; Baltagi, 2008; Cameron and Trivedi, 2010; Wooldridge, 2002). Significant findings in both tests give reasons to reject the null in favor of pooled OLS. The Hausman test method examines the relationship between fixed effects and random effects (Hausman, 1978). The null hypothesis entails that the individual effects are uncorrelated with any other independent variable. If rejected, then the fixed effect model is preferred to the random effect model. Finally, the Chow test examines the poolability of the regressors' slopes (Baltagi, 2008). In the case of rejecting the null hypothesis of poolability, then all industries could have their own slopes of regressors, and hence fixed, or random effects are no longer the best option. However, there is no reason to believe that the heterogeneity includes slopes changing across entities and time. Consequently, a Chow

[1] Figure **1.3** in the introduction provides a comprehensive overview of the research framework

test to analyze the poolability of the panel data is not conducted. Both the F-test and LM test assess the presence of individual-specific effects contrary to the pooled OLS, while the Hausman chooses between the two approaches if the tests are rejected. Conclusively, the approach according to Kennedy (2003) is followed, that is, applying the LM test for individual-specific effects against pooled OLS, and the Hausman to select the most appropriate model of the two.

It is beneficial to think critically before applying the tests. A natural beginning is with the OLS and looks at the potential problems in regards to the heterogeneity, both observed and unobserved. By looking at sales and innovation activities in different industries, it is natural to believe that values from one cluster to another will differ substantially. For instance, the expenditures in intramural and extramural are significantly higher in the manufacturing industry than in construction and hotels/restaurants 5.2, which leads to the assumption of heterogeneity, and thus, the presence of individual-specific effects. Continuously, if it is more likely that the individual heterogeneity is collected in the disturbance term and the industry effect is uncorrelated with the independent variables, then a random effects model is preferred. The fixed model is preferred when the heterogeneity managed by industry-specific intercepts and the individual effect could perhaps be associated with any other dependent variable. In other words, a fixed effects model should be applied if each industry has its initial level and shares the same disturbance variance with other industries, and a random effects model should be applied when each entity has its distinctive disturbance because it will more effectively sort out the heteroskedastic disturbances (Park, 2011).

The first test executed is the LM test, where pooled OLS is tested against random effects model, that is, individual-effect models, to see which one is most suited. More specific, the LM test inspects if the individual specific variance parameters are zero. If the test yields significant values and it is reasons to believe that the null is rejected, then it is promising to determine random effects in the panel data. Meaning, that this way of handling the individual effects are more capable of managing heterogeneity than the pooled OLS. Performing the LM test in Stata yielded a chi-squared of 76.16, thus rejecting the null of pooled OLS in favor of the random effects model ($p < .0001$).

The Hausman test matches fixed and random effect models with the null hypothesis that individual effect is uncorrelated with any independent variable (Park, 2011). If the findings suggest rejecting the null of no correlation, then a fixed effects method is appropriate, and vice versa for the random effects. When the null is not rejected, the estimates of the individual-effects methods should not be systematically different. The Hausman method says that "the covariance of an efficient estimator with its difference from an inefficient estimator is zero" (Greene, 2008, p. 208). If the null hypothesis were to be rejected, it is possible to determine that the individual effects, μ, are significantly correlated with one or more independent variables. This leaves the random effect model challenging to handle and the fixed effects model as a suitable choice. The robust test used in this section extends the standard Hausman test to be a heteroskedastic and cluster-robust version and restricts the covariance between the fixed effect and the regressors to zero (Arellano, 2003; Wooldridge, 2002). Thus, the choice between fixed effects and random effects estimator relies on the robust Hausman test. Applying this test in Stata produces a statistic of 54.76 and a highly significant p-value ($p < .0001$). This gives a pretty strong impression that the key assumption of random effects model is violated; hence a fixed effects estimation will be used. Both the Breusch-Pagan LM test and the Robust Hausman test is summarized in table 6.1 below.

Table 6.1: Test summary POLS vs RE vs FE

Test	H_0	Test Statistic	p-value	Test of H_0	Conclusion
Breusch-Pagan LM	Pooled OLS	76.16	0.0001<	Reject	Individual-effect
Robust Hausman	Random Effect	54.76	0.0001<	Reject	**Fixed-effects**

See Stata output in appendix A.2.1

Within vs LSDV

After deciding that the fixed effects model is appropriate, it is important to decide which type of fixed effect to use. As briefly mentioned in the methodology chapter (4.3), there are two ways of executing the fixed effect, the LSDV and "within" estimation. LSDV approach generates the same amount of dummy variables as entities. This could be somewhat problematic with many entities, and thus the within effect model would be more suitable because it utilizes mean-transformed variables without creating dummies (Park, 2011). The parameter coefficient reported by both methods is equal. However, when interpreting measures related to standard errors, R-squared, and F-statistic, one has to be cautiousness since these numbers will differ. This analysis will use the within effect for the estimation and correct the diagnostic measures if needed.

Robust Standard Error

When errors for different observations are correlated, the assumption from equation 4.5 is violated. In this case, normal and robust estimates of the variance-covariance matrix are invalid, as the errors are clustered (Cameron & Trivedi, 2010). Thus, it is sensible to assume a difference between the industries and the correlation within. This has to be appropriately adjusted so that the regression model does not over-predict the dependent variable for one given industry. Otherwise, it is likely to over-predict for individual members of that industry, giving a positive correlation within. Thus, cluster-robust standard errors are applied to the equation below to adjust for errors that are correlated within each entity and uncorrelated across entities.

To sum up, *sales* are regressed against innovation activities in a one-way fixed effects model using the "within" transformation with cluster-robust standard errors. The model is illustrated with the following equation:

$$\begin{aligned} sales_t &= (\alpha + \mu_i) + \beta_0 + \beta_1 intra_{i(t-2)} + \beta_2 extra_{i(t-2)} + \\ &\quad \beta_3 aqknow_{i(t-2)} + \beta_4 aqmachin_{i(t-2)} + \beta_5 exrights_{i(t-2)} + \\ &\quad \beta_6 consult_{i(t-2)} + \beta_7 nonrd_{i(t-2)} + \nu_{it} \end{aligned} \tag{6.1}$$

6.1.2 Regression results

This subsection will present the estimation of the regression specified in the previous subsection. Only the concluding fixed effects “within” model will be presented and commented, however, both the random and fixed effects were fairly similar, showing sufficient robustness of the data. [1] The regression output is given in table 6.2.

Table 6.2: Fixed effects "within" estimation

Dependent variable:	**Sales**	
Expenditures to Innovation Activities		
Intramural	0.0532*	(0.0172)
Extramural	-0.0298	(0.0153)
Aq. of Knowledge	-0.0226	(0.0153)
Aq. of Machinery	0.0106	(0.0275)
Aq. of External Rights	-0.0072	(0.0153)
Consultancy Services	-0.0480**	(0.0142)
Other non-R&D activity	0.0201	(0.0356)
Constant	11.80***	(0.3520)
R-squared	0.314	
F-statistic	20.89***	
Observations	54	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As the table displays, *intramural* expenditures have a significant positive effect on the predicted variable *sales*. Effectively, the estimated model suggests that all expenditures to R&D activities performed within the enterprise have a positive effect on domestic sales at the .05 significance level. This finding aligns well to the expectations according to the literature review and the Oslo Manual.

The results concerning the acquisition of *consultancy services* provides some interesting and unexpected results. The parameter coefficient turned out to be negative at

[1] See appendix A.2.1 for the estimated random effects model

the .01 level, indicating a strong negative relationship between *sales* and the acquisition of *consultancy services*. This stands in contrast to the expectation that all innovation activities are intended to increase sales and firm performance. *Consultancy services* were expected to have a positive effect on firm performance since the services are often acquired at the aim of increasing firm performance by developing and implementing product and process innovations.

Expenditures to the *acquisition of machinery* and *other non-R&D* activities indicated a positive relationship, while expenditures to *extramural*, *acquisition of knowledge* and *acquisition of external rights* surprisingly suggested a negative relationship with *sales*. However, these relations were not significant at the .05-level and are somewhat unreliable in any reasoning or discussion.

The overall regression model has a high F-statistic, thus concludes that the coefficients are jointly different from zero and that the performed model is deemed adequate ($p < .001$). The regression yielded an R-squared of about 30 per cent, however, aiming to create a model with a high explanatory power of *sales* is difficult given the fact that many factors and elements may explain the sales of a company in any given industry. The objective of this thesis is to indicate and measure the influence specific variables have on each other, rather than finding a suitable model applied for prediction or similar. Thus, the R-squared value (nor the F-statistic) will not be examined further.

This specific model is a log-log model, that is, both dependent and independent variables are log-transformed, which makes the parameter coefficients the elasticity of the dependent variable to the regressors, ceteris paribus (Baltagi, 2008). The effect on *sales* from the innovation activities are lagged by two periods, meaning that there is a delay before the effect noted by the coefficient parameters occur. Nevertheless, this thesis is concentrating on the relative effect and significance, not the specific measured effect of each coefficient. Having an aggregate perspective suggests that a specific numerical effect would not give a whole lot of reasonable interpretation and intuition. It is valuable to keep in mind the mathematical interpretation of each coefficient to fully understand the effect; however, the focus of the discussion is ultimately the overall significance one variable has on another.

6.2 Innovation activities and types of innovation

The section addresses how innovation activities and cooperation affect types of innovation pursued by Danish enterprises. The model will first of all be constructed and justified, before the result of this specific research question is presented.

6.2.1 Model specification

Compared to the model in 6.1, this model entails different characteristics and require an alternative approach than the traditional linear methods explained in the methodology chapter. The dependent variables in this section are the four specific innovation types, that is, product, process, organizational, and marketing innovations, regressed against the independent variables, being expenditures to innovation activities as the model in section 6.1.1. All the measures of the innovation types differ from the continuous variable regarding domestic sales, in the way that they are fractional responses. As explained in chapter 4, these innovation types are noted as a fraction of responses to the total amount of respondents who exercised one or more of the types during a given year. According to Papke and Wooldridge (2008), applying the GEE method to these dependent variables yields satisfactory and well-acknowledged results.

This estimated model is appropriately constructed to answer the third research question under the main problem statement. The innovation activities were again used as regressors, along with the interaction terms between cooperation and activities, to examine the different dynamics relating product, process, organizational, and marketing innovations. All regressors are lagged by one period, following the same intuition as in section 6.1.1. That is, expenditures to innovation activities are regarded as a first move toward implementing an innovation type while introducing one of these innovation types is regarded as the first move toward improved firm performance. Given the non-linear nature of these dependent variables, the construction and focus are hence different from the linear methods, especially regarding time and cluster effects. Both year and industry dummies are frequently included as control variables when applying this type of framework, to be able to capture more complex relationships. Calendar-year dummies are therefore included to control for fluctuations in the economy, along with circumstances that could occur unrelated to innovation activities in the Danish enterprises. The industry dummies are included as it possesses some of the similar attributes, and to account for industry-specific effects using the classifications from DB07.

Link Function

To fit a GEE model according to Liang and Zeger (1986), it is essential to specify a link function, the distribution of the dependent variable, and the correlation structure of the dependent variable. The first element that needs to be specified of the GEE is the link transformation function (Ballinger, 2004). This is to accurately fit the parameter coefficients of the population-averaged response of the entire sample (McCullagh & Nelder, 1989). The consensus is that the distribution of the dependent variable decides which link function to use. Having to deal with dichotomous binary response variables usually indicates a probit or logit link function, where the logit link being the standard linking function (Ballinger, 2004). Following Ballinger (2004), the logit link function is applied, where the log of the odds ratio transforms all regressors. This specified link attributes the estimating equation to map the interval from zero to one. It is worth mentioning that other link function for binary response variables, as probit for cumulative predictive analysis, provides essentially the same coefficients and results in general as logit; thus this specification is not crucial for the final conclusion.

Distribution

The next stage after the link specification involves the distribution of the predicted variables. It is necessary to specify the correct distribution in order to calculate the variance as a function of the mean response (Hardin & Hilbe, 2013). Chapter 4.3.2 introduces specifications of distributions from the exponential families, that is, normal, inverse normal, binomial, Poisson, negative binomial, and Gamma distributions. When fitting a GEE (or any generalized linear model), it is crucial to specify the distribution of the response variable. If not done correctly, the variance could be inefficiently calculated as a function of the mean and misinterpretations of the parameter coefficient might occur (McCullagh & Nelder, 1989). As all response variables in the section are binary in nature, a binomial distribution is specified.

Correlation Structure and QIC

The final step involves a specification of how the responses within subjects are correlated, and a suitable specification will in general increase the efficiency of the estimation. The QIC was used to choose between independent, autoregressive, stationary, given that exchangeable, unstructured and nonstationary would not converge with Stata's standard maximum iterations. The models with the lowest QIC is interpreted as the most reliable. As the table presents 6.3, the values from the test did not differ considerably, however, testing the QIC on the GEE model with the innovation activities favored the independent correlation structure consistently throughout the dependent variables.

Table 6.3: Summary of QIC values

	Independent	Autoregressive (1)	Stationary
Product	**56.115**	56.118	56.117
Process	**64.894**	64.914	64.909
Organizational	**75.658**	75.660	75.661
Marketing	**69.162**	69.162	69.162
Including Cooperation			
Product	**56.109**	56.109	56.109
Process	**64.876**	64.889	64.886
Organizational	**75.646**	75.646	75.646
Marketing	**69.155**	69.157	69.157

Performing the QIC function in Stata produces both QIC and QIC_u. According to Cui (2007), QIC_u could choose a different model as it is just an approximation to QIC. Consequently, QIC_u will not be appropriate to select the most suiting correlation structure, thus QIC is recommended. Only QIC is presented in table 6.3, whereas the independent correlation structure is prevailing amongst the other structures.

However, choice of correlation structure should not undoubtedly be guided by theory, as there are many correlation structures which are not suited for time-dependent correlation structures, like for instance the exchangeable structure (Ballinger, 2004). As noted above, the independent correlation structure was favored by the QIC; however, the criterion scores were marginally different from each other. Following Ballinger (2004), it is important in this case, and in general, to select and discuss the model that makes most theoretical and intuitional sense. A Woolridge test (Wooldridge, 2002) could address the within-subject correlation, but it is assumed that the responses of each industry are independent of each other, which makes the independent correlation structure a sensible choice. An autoregressive correlation structure could be specified to address the correlation within clusters. Nonetheless, both the independent and the AR-model yields reasonably similar results, thus the independent structure is deemed appropriate as a base model.

Semi-robust Standard Errors

One characteristic of the GEE estimation is that the estimated standard errors are classified as "semi-robust", rather than robust. The coefficient parameters of a GEE model utilizes the Huber-White sandwich estimator, often referred to as robust standard error estimates. This is because, in theory, the results of a GEE analysis are robust against a wrong choice of the working correlation matrix (Twisk, 2013). However, this is only when the link function is specified correctly, and the model identifies the mean accurately. The standard errors are thus labeled semi-robust. However, it is not possible to specify cluster-robust standard errors as used in the regression in section 6.1, but the results are seemingly as well specified as the clustering option.

In sum, according to the characteristics of the dependent variable, the GEE models are estimated with a logit link function, binomial distribution, independent correlation structure, and semi-robust standard errors (Ballinger, 2004; Papke and Wooldridge, 2008; Markus, 2013). The four models are specified with the following equation:

$$\text{logit}\left\{E(y_{it}^{\text{Innovation types}})\right\} = X_{it}^{\text{Innovation activities}}\beta \qquad y \sim \text{Bernoulli} \tag{6.2}$$

6.2.2 Regression results

Table 6.4 includes the results of the GEE regression models explaining innovation types with the expenditures to innovation activities along with cooperation as regressors. There are two sets of regressors present in the table below. The first set of regressors is the stand-alone expenditures to innovation activities, while the second set is the same expenditures including the effect of cooperation. This section will present the output before specific numbers are commented with respect to expectations concerning sign and significance.

Table 6.4: Generalized estimating equations: innovation types

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Product		Process		Organizational		Marketing	
Expenditures to Innovation Activities								
Intramural	0.0346		0.0119		0.0902*		0.0187	
	(0.0299)		(0.0472)		(0.0406)		(0.0377)	
Extramural	0.0230		-0.00662		0.00214		0.0774*	
	(0.0477)		(0.0265)		(0.0580)		(0.0390)	
Aq. of Knowledge	0.0820		0.0866***		0.0970**		0.0780	
	(0.0497)		(0.0257)		(0.0336)		(0.0482)	
Aq. of Machinery	-0.0280		0.0577		0.00141		0.0393	
	(0.0279)		(0.0426)		(0.0544)		(0.0840)	
Aq. of External Rights	-0.0180		-0.0175		0.00291		-0.00536	
	(0.0227)		(0.0140)		(0.0195)		(0.0203)	
Consultancy Services	-0.0913		-0.164***		-0.0479		-0.0181	
	(0.0690)		(0.0490)		(0.0657)		(0.0466)	
Other Non-R&D activity	0.140*		-0.0469		-0.00333		-0.0437	
	(0.0699)		(0.0349)		(0.0657)		(0.0839)	
Intramural*Coop		0.123		0.275***		0.260		0.0455
		(0.232)		(0.0491)		(0.172)		(0.100)
Extramural*Coop		-0.117		-0.0927		-0.0722		0.180
		(0.155)		(0.0934)		(0.154)		(0.103)
Aq. Knowledge*Coop		0.301*		0.279***		0.326***		0.228
		(0.131)		(0.0491)		(0.0859)		(0.120)
Aq. Machinery*Coop		-0.199*		0.154		-0.105		0.0804
		(0.0910)		(0.103)		(0.117)		(0.141)
Aq. External Rights*Coop		-0.0243		-0.0352		0.0132		-0.00812
		(0.0439)		(0.0414)		(0.0552)		(0.0393)
Consultancy Services*Coop		-0.360		-0.469***		-0.200		-0.114
		(0.193)		(0.118)		(0.191)		(0.180)
Other Non R&D*Coop		0.258		-0.113		-0.160		-0.203
		(0.136)		(0.157)		(0.158)		(0.179)
Constant	-2.557***	-1.556***	-0.817	-1.431***	-1.846*	-1.045***	-2.145*	-1.654***
	(0.722)	(0.253)	(0.730)	(0.297)	(0.855)	(0.148)	(0.909)	(0.312)
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald chi-sq	492.48***	145.99***	835.51***	195.36***	100.06***	13.34	583.24***	32.51***
Observations	63	63	63	63	63	63	63	63

Semi-robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Firstly, product innovation has only one significant relationship without cooperation interaction terms. Expenditures to *other non-R&D* activities have a positive effect on product innovation at the .05 significance level. This is in accordance to the expectations that these costs are directly linked to the implementation of product innovations, as well as the other types, in the form of technical preparation associated with new products, workflows, new marketing or productions processes. However, when including the effects of cooperation, the relationship is still positive but insignificant at the .05 level. That is, cooperation with external partners does not improve the implementation likelihood of product innovations significantly. *Acquisition of knowledge* and *machinery* turns significant when interacted with *cooperation*, indicating a positive and negative relationship, respectively. These findings could suggest that including cooperation in the variables increases the chances of effect toward complete product innovations.

Secondly, from the regression output, it is evident that two innovation activities have a highly significant ($p < .001$) effect on process innovation, namely *acquisition of knowledge* and *consultancy services*. Neither of the coefficients changes relative to each other nor changes the significance level while being exposed to *cooperation*, indicating a vigorous effect toward process implementations. An interesting finding is that the acquisition of *consultancy services* indicates a negative relationship with process innovation, suggesting that the acquisition of *consultancy services* is not necessarily the way to go about if the objective is to implement a process innovation. *Intramural* expenditures on the other side, changes when exposed to *cooperation*, showing a strong positive and significant effect on process innovation.

Thirdly, when it comes to organizational innovations, that is, application of new organizational design in the firms' business activities, workplace organization, and external relations, two variables are especially worth mentioning; *intramural* and the *acquisition of knowledge*. Both expenditures were significant and indicated a positive relationship to organizational innovation at the .05 and .01 significance level, respectively. Interestingly enough, when *intramural* expenditures are exposed to *cooperation*, there is no longer a significant effect on organizational innovations, indicating that external partners relating all expenditures to R&D activities within the enterprise are not necessarily beneficial to complete the attempted innovation. The *acquisition of knowledge* suggests an even stronger significant relationship (now $p < .001$) to the dependent variable when exposed to *cooperation*. However, it is worth noting that the GEE model used with the interaction term as regressors returns an insignificant Wald chi-square (p-value = 0.101, not rejecting H_0).[2]

The fourth and final dependent variable is marketing innovation. The only significant indication from the regression output is that *extramural* expenditures have a positive effect ($p < .05$) on whether or not marketing innovations are accomplished. The marketing innovations usually refer to the application of a new or significantly

[2] Wald chi-square is commonly used to test whether the coefficients are different from each other or equals zero in GEE models (Ballinger, 2004). This could cause interpretation problems when there is not enough evidence to reject H_0

changed marketing method, related to a product's layout, packaging, promotion or pricing, whereas the expenditures to extramural activities refer to the R&D purchased from external organizations. This relationship may indicate that it is advantageous to address external sources in order to complete a marketing innovation. That is, not performing the marketing innovation "in-house" or in cooperation with external partners.

The Wald chi-square used to test the significance of all parameters (also individual), appeared to be highly significant on all regressions, except the already mentioned regression number (6) in table 6.4. The p-value is used as the decisive factor compared to the critical value of .05 and is the probability of obtaining the Wald chi-square statistic if there is no joint effect of the regressors on the regressand.

The GEE models applied in this section follow the same specifications investigated in section 6.2.1, which provides a population-averaged (or marginal) model. This type of model gives an average response for the observations. More specifically, for every unit increase in a regressor of the total population, GEE expresses how much the mean response would vary (Zorn, 2001). The parameter coefficients of the estimated equations are equivalent to odds ratios from the standard logistic regression, as it utilizes the logit link function and the independent correlation structure.

6.3 Types of innovation and sales

The section addresses how the innovation types (i.e., product, process, organizational, and marketing innovation) affect sales in Danish enterprises. The applied model will first of all be constructed and justified, before the result of the fourth research question is presented.

6.3.1 Model specification

The attention in this section is the relationship between innovation types and firm performance, whereas the dependent variable *sales* is regressed against the different innovation types. In order to analyze the results accurately, it is imperative to find the most appropriate model. The applied model will be specified and justified in this subsection. However, it is worth noting that the model will be quite similar to the model specified in subsection 6.1.1 as these two possess many similarities and have the same predicted variable. The reader is thus referred to that subsection for a more comprehensive explanation and specific sources.

The model constructed at the end of this subsection is a one-way effect model, with industry clusters as the primary consideration through a fixed effects transformation.

Few factors indicate a two-way effect model in this case, as all variables have a somewhat persistent trend and no significant set of outliers is present. Thus, to avoid complications in overfitting the model, a two-way effects model is disregarded.

Following Kennedy (2003), the LM test is applied for individual-specific effects, more precisely, random effects, against pooled OLS. If the LM test is rejected, then the Robust Hausman test selects the most appropriate model between the individual-effects models (fixed and random effects). Briefly explained, if it is more likely that the individual heterogeneity is collected in the disturbance term and the industry effect is uncorrelated with the independent variables, then a random effects model is favored. Nevertheless, if the heterogeneity is managed by industry-specific intercepts and the individual effect could be associated with any other dependent variables, the fixed model is favored.

Table 6.5: Test summary POLS vs RE vs FE

Test	H_0	Test Statistic	p-value	Test of H_0	Conclusion
Breusch-Pagan LM	Pooled OLS	136.20	0.0001<	Reject	Individual-effect
Robust Hausman	Random Effect	55.88	0.0055	Reject	**Fixed-effects**

See Stata-output in appendix A.2.2

The LM test rejected the null of pooled in favor of random effects with a chi-square of 136.20 ($p < .0001$), indicating that the individual specific variance parameters are different from zero. Thus, concluding with individual specific effects. The robust Hausman yields a test statistic of 55.88, hence rejecting the null of random effects on a .01 significance level ($p = 0.0055$) in favor of fixed effects. Thus, it is reasonable to believe that the individual effects are significantly correlated with one or more independent variables, indicating that the random effect model is difficult to manage and hence that the fixed effects model are a superior choice under these circumstances. A summary of the effect tests is given in the table above (table 6.5).

As a result, the estimating model used to regress *sales* against innovation types is a one-way fixed effects estimation using the "within" transformation with cluster-robust standard errors. All innovation activities are lagged by one period to account for the time delay from implementing a product, process, organizational, or marketing innovation to the effect on firm performance. The constructed model is illustrated as:

$$\begin{aligned} sales_t &= (\alpha + \mu_i) + \beta_0 + \beta_1 prod_{i(t-1)} + \beta_2 proc_{i(t-1)} + \\ &\quad \beta_3 mrk_{i(t-1)} + \beta_4 org_{i(t-1)} + \nu_{it} \end{aligned} \tag{6.3}$$

6.3.2 Regression results

The coefficients from the estimated model specified above are given in table 6.6. As in subsection 6.2.2, only the fixed effects "within" model will be examined, as this was deemed most accurate in the discussion of the model above.[3]

Table 6.6: Fixed effects "within" estimation

Dependent variable:	**Sales**	
Introduced Innovation Types		
Product	0.0896	(0.498)
Process	-0.532	(0.444)
Organizational	0.168	(0.409)
Marketing	1.060^{**}	(0.244)
Constant	11.65^{***}	(0.105)
R-squared	0.185	
F-Statistic	8.541^{**}	
Observations	63	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

According to the estimated parameter coefficients, only marketing innovation has a significant effect on the dependent variable, *sales*. The coefficient indicates a strong positive relationship (at the .001 significance level) between the dependent and independent variable. This is a rather common finding according to peers [4], as marketing innovations are usually explicitly intended to increase sales of the company. As marketing innovation is the predictor variable and the predicted variable is revenue rather than pure profits, this finding makes a whole lot of intuitive sense.

Both product and organizational innovations provide indications of a positive relationship to *sales*, whereas process innovation suggested a negative relationship. These variables were, however, not significant at the .05 level, which is quite intriguing. First of all, product innovation was expected to have a positive and significant relationship

[3] The estimated random effects model found very similar results, showing sufficient robustness of the data. See appendix A.2.1 for Stata output.

[4] See section 4.1 for an explanation of marketing innovation and the "peers" mentioned in particular

with *sales*. Similarly, organizational innovation was also expected to have a positive and significant relationship with *sales*. The regressors were positive, although, the effects proved out to be insignificant. The finding related to process innovation was even more interesting. This type of implemented innovation was, from the previous discussion (section 4.1) expected to have a positive and significant effect. This is due to process innovations being aimed at enhancing products and services through new or improved processes, thus increasing sales. Nevertheless, the results from this specific regression, focusing on Danish enterprises, found the opposite relation. However, the coefficient on process innovation contains limited information to interpret going further, due to the insignificant relationship.

The estimated model has a rather low R-squared (18.5%) and a F-statistic on 8.54 ($p < .01$). Nonetheless, the fairly similar constructed model in section 6.1 excluded an in-depth discussion of the F-statistic and R-squared. Hence, it will also be excluded in this section as this thesis does not specifically aim to construct a regression with high explanatory power for prediction or modeling, but instead seeks to examine the overall influence and practical interpretation of the variables. The effect on *sales* from the different innovation types, particularly product, process, marketing, and organizational innovation, are lagged by one period as determined in the model specification. One unit change in the independent variables leads to a log-unit change of the coefficients' value in the dependent variable *sales*. However, the used aggregate approach makes the numerical changes from the regressors less relevant, whereas the relative change and significance between the parameter coefficients become focal.

Chapter 7

Discussion and interpretations

The overall objective of this thesis is to examine the relationship specified in the problem statement using quantitative methods. The progression of this paper went from defining concepts and analyzing the data from the Innovation Survey and Statistics Denmark, into panel data regressions, that enabled determinations and interpretations of meaningful relationships. This chapter will discuss the relationships found between innovation input, innovation output, and firm performance according to each research question, and examine possible explanations to contribute to the understanding of innovation and firm performance.

7.1 How do innovation activities and types differ between industries?

To better comprehend the dynamics and the relationship between innovation input, innovation output and firm performance for Danish enterprises, the differences between the various industries had to be further inspected. Nearly half of the enterprises were innovative in the period 2009 to 2016 and introduced one or more type of innovations. Most of these enterprises were organizational or marketing innovative, whereas every fifth launched either new or significantly improved products or processes in the period. Further, the industries *communication*, *manufacturing*, and *finance and insurance* invested the most in innovation activities relative to their sales. Indeed, these industries were also the three most innovative, indicating a relationship between innovation input and innovation output. In addition, the least innovative firms were those engaged in *hotels and restaurants* and *construction*. Not surprisingly, these firms were also the ones investing the least in innovation activities supporting the indicated relationship.

When looking at the characteristics of the Danish enterprises in light of *cooperation*, almost one out of three innovative enterprises cooperated with external partners concerning their innovation activities. The industries most likely to cooperate were *manufacturing* and *financial and insurance*, looking apart from the category *other industries* which includes companies such as arts and entertainment. As these two industries were also the most innovative, it may indicate that cooperating is vital for firms to implement innovations successfully.

7.2 How do innovation activities affect firms' sales?

The second research question addresses the relationship between input and firm performance, or more specifically how various innovation activities affect firms' sales. The findings concerning this relationship provided some expected and unexpected results.

Intramural expenditures were found to have a positive and significant effect on *sales*, indicating that R&D performed "in-house" are the only type of innovation activity suggesting an increased revenue outcome. Thus, investments in *intramural* R&D could provide firms with a competitive advantage over firms that are not investing in such activities. This finding concurs with previous research, who argue that investment in R&D is an important aspect for firms to improve their performance (e.g., Branch, 1974; Wieser, 2005). This finding may help enterprises' management influence future revenue with its decisions regarding R&D investments.

However, the results concerning the acquisition of *consultancy services* provided some interesting and quite unexpected results. The relationship between *consultancy services* and *sales* was found to be strong and negative, indicating that the purchase of consultancy services for the purpose of implementing a specific type of innovation, harms the enterprises' sales. A negative relationship is in contrary to the expectations, as consultants are often acquired to increase the firms' performance according to the Oslo Manual. More specifically, external professionals are attained to provide the firms with advice concerning the firms' performance (see section 4.1). However, as stated by Mitchell and Hamilton (1988), not all innovation activities are successful, which could explain the findings relating to consultancy services.

First of all, purchases of consultancy services are usually very costly, which could imply that the investment will be at the expense of other possible performance-generating investments. As a result, the sales from one year to the next could decrease as potential profitable investments are relinquished at the benefit of consultancy services. Another potential explanation could relate to the management shifting focus from sales to new or improved processes or organizational design that are severely time-consuming. Consequently, the ongoing innovation implementations could delay increased income by more than two years. As stated in the variables description (section 4.1), all innovation activities are implemented to increase firm performance. However, lengthy projects may result in declining sales longer than the period examined in this paper, before the relationship turns positive.

Another interesting finding is concerning the relationship between *extramural* expenditures, and *sales*. Expenditures related to external R&D activities were found to have a negative, however insignificant, effect on firms' sales. Aligned with the results found by Pandit et al. (2011), one explanation could be that this finding is an indicator that R&D activities are risky and increases the volatility of future performance. However, this result was not significant meaning that the inverse relationship is not valid. Further, as with *extramural* R&D expenditures, insignificant relationships were found between other innovation activities and domestic sales. One explanation for these insignificant relationships may be that it takes longer than two periods for the initiated innovation activities to generate revenue. Thus, as innovation activities lagged by two periods are examined, looking at a longer time perspective could result in different findings.

7.3 How do innovation activities and cooperation affect types of innovation?

The third research question relates to how innovation activities and cooperation affect different types of innovation. This relationship is essential to examine the relationship between innovation output and firm performance, as well as understanding the whole dynamics regarding innovation and firm performance in Danish enterprises.

The findings regarding **product innovation**, provided some interesting results. *Other non-R&D* activity indicates a positive relationship suggesting a positive effect toward product innovations. However, when exposed to *cooperation* the effect turned insignificant, which may be an indication of a non-beneficial relationship concerning partnerships and trade of information. Since the effect is fading when exposed to *cooperation*, one could insinuate several reasons for this. *Other non-R&D* refers to operating expenditures for innovations that are not R&D, which in turn could point to the fact that there are few suitable partners for this specific type of activity in the Danish market. That is, the innovation procedure may not be efficient enough when involving external partners. As *other non-R&D* activities involves preparations related to the introduction of new products, workflows, marketing approaches or production processes, this finding could imply that this specific type of activity is more efficiently performed without partners.

The expenditures related to the *acquisition of machinery* and the *acquisition of knowledge* suggest an insignificant relationship toward **product innovations**. Nonetheless, both activities turned significant when exposed to *cooperation*, indicating that cooperating on these activities increases the likelihood of an effect toward implementing a product innovation. However, for the *acquisition of machinery* the effect is negative, suggesting that it is not beneficial to collaborate on the activity if the objective is to implement a product innovation. This finding could relate to the fact that machinery is often time-consuming to completely mount, thus shifting the focus from other innovation activities that may increase the likelihood of implementing new products. The one year lag may also not be an appropriate period to illustrate the development from *acquisition of machinery* to the introduction of product innovations. Continuing, even though innovation activities aim at increasing innovations, some activities are proven harder to implement than others. Therefore, the *acquisition of machinery*, especially in cooperation with external partners, may cause more harm than good if a product innovation is the intended target. On the other side, the *acquisition of knowledge* suggest a positive relationship to product innovations when interacted with *cooperation*, thus indicating that an effective step toward a product innovation could be through the trade of information with external partners. Intuitively, this makes sense seeing that the most common partnerships (viewed in section 5.2) are suppliers and customers. Forming alliances and trading information is most likely beneficial to align supplier expectations and customer needs to ultimately invent a new product or service.

The main findings concerning the **process innovation** without the effects of *cooperation* relates to *acquisition of knowledge* and *consultancy services*. The joint effect of *cooperation* and *intramural* R&D expenditures indicates a significant relationship to process innovations. First of all, why *intramural* expenditures along with external partners had a negative relationship to process innovations may be due to several factors. Process innovations refer to new or significantly improved production or delivery methods that are introduced and is often referred to as a comprehensive innovation type to accomplish given its complicated nature. Therefore, the significant interaction term with *cooperation* could prove that process innovations is somewhat complex, and some sort of partnership is advantageous to most efficiently allocate the firms resources. The second argument regarding process innovation is the effect concerning the *acquisition of knowledge* and *consultancy services*, whereas both types of activities indicate a significant relationship, positive and negative, respectively. The *acquisition of knowledge* was expected to be positively related to process innovation, with and without cooperation, and displayed a robust relationship to the implementation of processes in the regression analysis.

However, the expenditures to *consultancy services* provided a more remarkable result, indicating a destructive effect when implementing new processes. The acquisition of *consultancy services*, as well as the other innovation activities from the Innovation Survey, are invested at the aim of introducing one or more of the innovation types (i.e., product, process, organizational, and marketing innovation). Thus, the negative relationship between consultants and processes contradicts with common perceptions in today's business practices, and would arguably stand as one of the most controversial findings of this thesis. Nevertheless, a couple of coherent arguments will be presented in order to attempt to substantiate the effect. First, as mentioned in the literature review in chapter 2, Baker and Freeland (1975) and Mitchell and Hamilton (1988) found that not all R&D and innovation projects are successful. Considering that many projects are started but not finished as a complete improvement or innovation, could contribute to the negative relationship between consultancy and process innovation. However, as industries are in focus and the numbers are aggregated, it is hard to conclude a valid company-specific proposal from the variables, and hence is not applicable in all companies and industries. The results could also be due to the sample used on Danish industries or the specific time period. Moreover, another important feature to note regarding the data sample is that the expenditures to *consultancy services*, in this context, is especially aimed at introducing one or more innovation types. Thus, the definitions by the Oslo Manual and the Statistics Denmark emphasized in the key concepts in chapter 3 is vital to be aware of to interpret the results and discussion of this thesis appropriately.

Both expenditures to *intramural* and *acquisition of knowledge* indicates a significant positive relationship to **organizational innovations**. The joint effect of the innovation activities and *cooperation* were deemed invalid by an insignificant test statistic; [1] hence it is not to be relied on too extensively in the conclusions. *Intramural* R&D expenditures provided an interesting result, given that it indicates a significant positive effect on organizational innovations. However, when exposed to *cooperation*, the effect turned insignificant, which is slightly unexpected according to earlier anticipations. The significant effect in general to organizational innovation is expected, in the sense that these expenditures are directed at the development and application of all innovation types (product, process, organizational, and marketing innovations). However, the insignificant interaction with *cooperation* was not as expected, but could yet be a sensible finding. Organizational innovations are not only referred to the firms' organizational

[1] See Wald chi-square from the results in section 6.2.2

structure but are frequently related to the profitability by reducing administrative or transaction costs. Thus, investing in *intramural* R&D could ultimately reduce these cost, either with the direct objective or indirectly through systematical research. If so, it could be a reasonable discovery that it does not have any significant effect to cooperate. The reason for this could be that it does not pay off to trade information with external partners, or that the cooperation requires more time to implement the organizational innovation.

Marketing innovation provided very few significant relationships regressed against the innovation activities and the cooperation interaction terms. The most striking finding relates to the *extramural* R&D expenditures which suggest a positive relationship to the dependent variable; however, turned insignificant when subject to *cooperation*. Thus, the results suggest that it is not necessarily advantageous to cooperate with an external partner in order to purchase R&D specific activities if the objective is to implement new marketing methods. Thus, *extramural* R&D has less effect on marketing innovation if executed in collaboration with external partners, as it could be more time-consuming or encounter too much bureaucracy along the way. As a result, it is better to approach the market directly than to spend time cooperating and addressing different methods. An important aspect of this finding is to look at the most common partnerships from section 5.2, which are suppliers and customers. As marketing often address suppliers and customers, it would seem inefficient in terms of marketing innovations to also cooperate, looking at an aggregate perspective.

7.4 How do types of innovation affect sales?

The fourth research question considers how innovation types affect firm performance in Danish enterprises. This relationship continues the third research question by looking at the final objective of innovation, namely firm performance. The firm performance measure, as familiar, relates to domestic sales.

The relationship between innovation types and firm performance is examined to study whether innovative firms, more specifically, firms that have launched one or more innovations in the past year, will affect the firms' sales. All four innovation types were anticipated to have a positive effect on performance as it is usually the intended target of the new implementations. Product innovation was expected from the literature review to have a strong significant effect on *sales*, through the invention of new or improved products or services. Process innovations were also expected to have a significant positive effect on *sales*, as the consequence of new or improved production or delivery methods. Findings relating to organizational and marketing innovation on domestic sales were actually according to expectations, in contrast to product and process innovations. Organizational innovations were found from the regression analysis to be positive but insignificant, due to its cost-reducing nature. This is opposed to income-generating objectives, whereas marketing innovations were thought to be positive and significant related to *sales*, as it proves out to be.

Marketing innovations were, as mentioned, expected to have a considerable positive impact on sales in the selected time-span. The reason behind this assumption relates to the general purpose of marketing, which includes raising awareness to the organization and promoting services or products. Marketing innovations aim to ultimately increase a firm's sales by satisfying customers, expanding into new markets, or introducing new products. This finding makes practical and commercial sense in an overall economy divided by industries and aligns impeccably with findings from the literature review (such as; Karabulut, 2015).

However, as all four innovation types supposedly should increase firm performance, why is only marketing advances meaningful in explaining sales for the overall economy? A suitable justification could involve the time span of the innovations. The analysis in question is primarily looking at how an implementation of a given type of innovation will affect the sales in distinctive industries by a one year delay. Thus, the perspective is rather short-term. It is reasonable to assume that the effect of marketing implementations affect firms' sales faster than any other innovation type, as it is meant to spread a message extensive and rapid across the existing and potential audience. The results could suggest that any other innovation accomplishment with the intention to increase firm performance, has a long-term effect on sales. Process innovations, where new or improved methods are implemented, may need a longer time perspective to have a significant positive effect. The effect of product, process, and organizational innovation are not deemed irrelevant; however, it is discovered to be insignificant with the perspective applied in this thesis.

Chapter 8

Conclusion and further discussion

8.1 Conclusion

Innovative firms have proven to be more productive and more adjustable to change, and innovation is, therefore, a vital tool for enterprises to gain competitive advantage and to survive. According to OECD (2010), this may result in firms improving their performance and creating value for their stakeholders. However, there are divided opinions on the innovation-performance relationship that may be due to inadequate understanding of factors influencing this relationship. Therefore, this study has been conducted to acquire a better perception of the complex dynamics of innovation and to contribute new insight to the Danish market. The relationship between innovation input, innovation outputs, and firm performance is examined using data from the Innovation Survey provided by Statistics Denmark.

The **first research question** explores the differences in innovation activities and innovation types between Danish industries. The descriptive analysis in chapter 5 scrutinized the characteristics of innovation activities and types and found several fundamental distinctions between the nine industries. The findings suggest that the industries investing the most in innovation activities were, in fact, the most innovative industries, indicating a positive relationship between innovation activities and innovation types in general. Furthermore, the findings suggest that cooperation could be an important factor for firms to implement innovations, as the most innovative firms were also the industries with the strongest demand for external partners. These findings gave valuable insight into the topic of innovation and created a necessary foundation to interpret and discuss the results of the regressions.

The **second research question** addresses the types of innovation activities that affect the firms' sales after two periods of the initiated activity. The findings suggest that only *intramural* R&D activities has a positive effect on *sales*. These activities involve all creative work systematically initiated to increase the knowledge base, which could provide firms with a competitive edge. However, the results also suggest that acquiring *consultancy services* harmed firms' revenue. Explicitly, the expenditures to consultancy services with the aim of increasing innovations has a negative effect on revenue after two years.

The **third research question** looks at innovation activities and cooperation concerning the effect on product, process, organizational and marketing innovations, and found some remarkable results. The findings suggest that *other non-R&D* activities are essential for introducing new products, and should not be in collaboration with external organizations. This means that activities such as industrial design involving technical specifications, testing and evaluation, and setup and engineering, has a positive effect on the implementation of product innovations. However, it was found that *acquisition of machinery* in collaboration with others, will decrease the chances of implementing new products and hence is an inefficient tool if the objective is to improve product or services.

The findings further suggest that some specific innovation activities affect process innovations, that is, influence new or significantly improved production or delivery methods. First, the *acquisition of knowledge*, such as computer services, technical or scientific services, was found to increase the likeliness of implementing process innovations regardless of cooperation. Secondly, *intramural* R&D activities were found to increase the likeliness of successfully implementing process innovations; however, only in collaboration with others. An intriguing finding is that acquiring *consultancy services* was found to have a negative impact. That is, based on the result, firms aiming to be process innovative does not enrich the intended target by investing in consultancy services, regardless of the cooperation arrangement.

Two innovation activities are affecting the implementation of organizational innovations. Both conducting *intramural* R&D activities and *acquiring knowledge* were found to have a positive effect on the implementation of organizational design. Although, only the latter had an effect when examined together with cooperation. Meaning, *intramural* R&D expenditures done in collaboration with external partners will not increase the chances of implementing new or improved organizational procedures.

Firms aiming to introduce new or significantly changed marketing approaches, should according to the findings purchase *extramural* R&D activities. Carrying out this activity in collaboration with others, however, was found to be insignificant. Thus, *extramural* activities should not be performed in cooperation, as time-consuming collaborations may result in competitors gaining the upper hand.

Finally, the **fourth research question** concerns the importance of product, process, organizational, and marketing innovations and the subsequent effect on firms' sales. Marketing innovations were the only type found to be positively related to sales. Given the short-term perspective of one year, the findings suggest that innovations related to marketing are vital to a firm's revenue. Hence, new or improved marketing methods may improve the firms' liquidity promptly, and as a result, outmaneuver competitors and provide shareholder value.

Considerations of all research questions, along with its findings, provides an intriguing insight into the topic of innovation for the Danish market and contributes to the existing scholarship. The relative and individual associations between innovation input, innovation output, and firm performance were analyzed and discussed, whereas the findings provide information about strategic decisions enterprises' management can adopt to possibly stimulate future growth and endure fierce competition.

8.2 Limitations and further discussion

Even though this thesis has provided valuable insight into an exciting and current topic, several aspects could limit the universal relevance of the findings. First and foremost, it is important to note that this paper analyses the macro-data made available from the Statistics Denmark, which limits the opportunity to instruct company-specific advisement as the numbers are aggregated into industries, regions, and size class. However, the data used is still valuable to analyze based on an aggregated perspective, keeping in mind that modifications in the interpretations are required and the overall prediction accuracy decreases. In other words, it is imperative to keep in mind that many excluded factors could influence different variables altogether.

Another noteworthy limitation could be related to the chosen firm performance measure, domestic sales, used to examine the outcome of innovation implementations and innovation activities. This measurement is a pure revenue variable that excludes all the effects of changes in the firms' costs. As some of the innovation processes aim to increase performance by reducing costs, this is not captured by the chosen performance measure and using another variable that includes costs could, therefore, give different findings. However, as discussed in the literature review, scholars have found evidence of a positive relationship between innovation and firm performance regardless of the performance measure, and the use of the variable sales is therefore deemed appropriate. On the subject of sales, it is crucial to be aware of the fact that sales could also be related to non-innovators. As explained in the key concept (chapter 3), this paper looks primarily at companies that acknowledged to implement at least one innovation type the same year, thus leaving out non-innovators. Nonetheless, as the trends in each industry are emphasized, the effect of excluding non-innovators could be negligible, yet it is critical information to recognize.

Finally, this thesis is looking at a period of one or two years, giving minor flexibility in terms of distinctive effects that may be more or less time-consuming than others. Therefore, the time-perspective could represent the fact that many of the variables do not have a significant effect in the chosen period, and subsequently, could yield different results in another outlook. Although interesting dynamics could be unveiled by asserting a long-term perspective, the survey data does not stretch far enough back to make robust inferences.

Appendix A

A.1 Do-file

The DO-file is provided to give a brief insight into the regressions analysis and diagnostic test applied in this thesis. Only the most relevant commands are included, that is, data manipulations and the creation of lagged values, interaction-terms, and dummies are omitted.

```
import excel "S:\H2018\Master\final_data.xlsx", sheet("percent") firstrow
egen id=group(industry)
xtset id year

* Lists of independent variables
global xlist Lintra Lextra Laqknow Laqmachin Lexrights Lconsult Lnonrd
global xlistL2 L2intra L2extra L2aqknow L2aqmachin L2exrights L2consult L2nonrd
global xlistcoop intracoop extracoop aqknowcoop aqmachincoop exrightscoop consultcoop nonrdcoop
global xlist3 L.prod L.proc L.org L.mrk


*************************************
*** Regression Research Question 2 ***
*************************************

* Breusch.Pagan LM test for random effects vs OLS

quietly xtreg sales $xlistL2, re vce(cluster id)
xttest0

* Hausman test for fixed versus random effects model

quietly xtreg sales $xlistL2, re vce(cluster id)
xtoverid

* Random effect model

xtreg sales $xlistL2, re vce(cluster id)

* Fixed effect model

xtreg sales $xlistL2, fe vce(cluster id)


*************************************
*** Regression Research Question 3 ***
*************************************

** Generalized Estimating Equations
```

```
* Innovation types and innovation activities

xtgee prod $xlist i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee proc $xlist i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee org $xlist i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee mrk $xlist i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)

* GEE Including Cooperation

xtgee prod $xlistcoop i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee proc $xlistcoop i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee org $xlistcoop i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)
xtgee mrk $xlistcoop i.id i.year, family(binomial 1) link(logit) corr(ind) vce(robust)


*************************************
*** Regression Research Question 4 ***
*************************************

* Breusch.Pagan LM test for random effects vs OLS

quietly xtreg sales $xlist3, re vce(cluster id)
xttest0

* Hausman test for fixed versus random effects model

quietly xtreg sales $xlist3, re vce(cluster id)
xtoverid

* Random effect model

xtreg sales $xlist3, re vce(cluster id)

* Fixed effect model

xtreg sales $xlist3, fe vce(cluster id)
```

A.2 Stata output

A.2.1 Regression analysis 1: Innovation Activities and Sales

```
Random-effects GLS regression                   Number of obs      =        54
Group variable: id                              Number of groups   =         9

R-sq:                                           Obs per group:
     within  = 0.3140                                          min =         6
     between = 0.0001                                          avg =       6.0
     overall = 0.0011                                          max =         6

                                                Wald chi2(7)       =    398.64
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                     (Std. Err. adjusted for 9 clusters in id)
```

sales	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
L2intra	.0541842	.0171065	3.17	0.002	.0206561	.0877123
L2extra	-.0287876	.0147826	-1.95	0.051	-.0577609	.0001858
L2aqknow	-.0223046	.0150962	-1.48	0.140	-.0518925	.0072834
L2aqmachin	.0118524	.026789	0.44	0.658	-.0406531	.0643578
L2exrights	-.0065506	.0150903	-0.43	0.664	-.036127	.0230259
L2consult	-.0468744	.0139052	-3.37	0.001	-.074128	-.0196207
L2nonrd	.0225904	.035115	0.64	0.520	-.0462337	.0914144
_cons	11.76086	.4296146	27.38	0.000	10.91883	12.60289
sigma_u	1.6755539					
sigma_e	.06255467					
rho	.99860813	(fraction of variance due to u_i)				

Figure A.1: Random effects model

```
Breusch and Pagan Lagrangian multiplier test for random effects

        sales[id,t] = Xb + u[id] + e[id,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                   sales |   1.142204       1.06874
                       e |   .0039131       .0625547
                       u |   2.807481       1.675554

        Test:   Var(u) = 0
                             chibar2(01) =    76.16
                          Prob > chibar2 =   0.0000
```

Figure A.2: Breusch-Pagan lagrangian multiplier

```
Test of overidentifying restrictions: fixed vs random effects
Cross-section time-series model: xtreg re  robust cluster(id)
Sargan-Hansen statistic  54.759  Chi-sq(7)    P-value = 0.0000
```

Figure A.3: Robust Hausman test

A.2.2 Regression analysis 3: Innovation Types and Sales

```
Random-effects GLS regression                   Number of obs      =        63
Group variable: id                              Number of groups   =         9

R-sq:                                           Obs per group:
     within  = 0.1839                                         min =          7
     between = 0.0321                                         avg =        7.0
     overall = 0.0307                                         max =          7

                                                Wald chi2(4)       =     32.69
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                                    (Std. Err. adjusted for 9 clusters in id)
------------------------------------------------------------------------------
             |               Robust
       sales |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        prod |
         L1. |   .2272271     .518823     0.44   0.661    -.7896472    1.244101
             |
        proc |
         L1. |  -.5568844    .4513606    -1.23   0.217    -1.441535    .3277662
             |
         org |
         L1. |    .097901    .4227999     0.23   0.817    -.7307717    .9265736
             |
         mrk |
         L1. |   1.094198    .2629021     4.16   0.000     .5789197    1.609477
             |
       _cons |   11.64806    .4025474    28.94   0.000     10.85908    12.43704
-------------+----------------------------------------------------------------
     sigma_u |  .61924943
     sigma_e |  .07773111
         rho |  .98448796   (fraction of variance due to u_i)
```

Figure A.4: Random effects model

```
Breusch and Pagan Lagrangian multiplier test for random effects

        sales[id,t] = Xb + u[id] + e[id,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                   sales |    1.14456       1.069841
                       e |   .0060421       .0777311
                       u |   .3834699       .6192494

        Test:   Var(u) = 0
                             chibar2(01) =   136.20
                          Prob > chibar2 =   0.0000
```

Figure A.5: Breusch-Pagan lagrangian multiplier

```
Test of overidentifying restrictions: fixed vs random effects
Cross-section time-series model: xtreg re  robust cluster(id)
Sargan-Hansen statistic  55.876  Chi-sq(4)    P-value = 0.0000
```

Figure A.6: Robust Hausman test

Bibliography

Arellano, M. (2003). *Panel data econometrics*. Oxford university press.

Artz, K. W., Norman, P. M., Hatfield, D. E., & Cardinal, L. B. (2010). A longitudinal study of the impact of r&d, patents, and product innovation on firm performance. *Journal of product innovation management*, *27*(5), 725–740.

Baker, N. & Freeland, J. (1975). Recent advances in r&d benefit measurement and project selection methods. *Management science*, *21*(10), 1164–1175.

Baldwin, J. R., Business, S. C., & Group, L. M. A. (1994). *Strategies for success: A profile of growing small and medium-sized enterprises (gsmes) in canada*. Statistics Canada, Business and Labour Market Analysis Division.

Ballinger, G. A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational research methods*, *7*(2), 127–150.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Bettis, R. A. & Hitt, M. A. (1995). The new competitive landscape. *Strategic management journal*, *16*(S1), 7–19.

Branch, B. (1974). Research and development activity and profitability: A distributed lag analysis. *Journal of Political Economy*, *82*(5), 999–1011.

Breusch, T. S. & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The review of economic studies*, *47*(1), 239–253.

Cameron, A. C. & Trivedi, P. K. (2010). *Microeconometrics using stata*. Stata press College Station, TX.

Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage publications.

Changyong, F., Hongyue, W., Naiji, L., Tian, C., Hua, H., Ying, L., et al. (2014). Log-transformation and its implications for data analysis. *Shanghai archives of psychiatry*, *26*(2), 105.

Cohen, W. M. & Levinthal, D. A. (1989). Innovation and learning: The two faces of r & d. *The economic journal*, *99*(397), 569–596.

Crépon, B., Duguet, E., & Mairessec, J. (1998). Research, innovation and productivi [ty: An econometric analysis at the firm level. *Economics of Innovation and new Technology*, *7*(2), 115–158.

Cui, J. (2007). Qic program and model selection in gee analyses. *The Stata Journal*, *7*(2), 209–220.

Danish Ministry of Economic Business Affairs. (2008). *Innovation in denmark – how danish enterprises convert new ideas and knowledge into value.*

Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. *Journal of economic literature*, 1120–1171.

DST. (2016a). Danmarks statistik; danmarks statistiks formidlingspolitik. Retrieved April 9, 2019, from https://www.dst.dk/da/OmDS/strategi-og-kvalitet/formidlingspolitik

DST. (2016b). Danmarks statistik; innovation i erhvervslivet. Retrieved April 9, 2019, from https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/innovation-i-erhvervslivet

DST. (2016c). Danmarks statistik; statistisk behandling. Retrieved April 9, 2019, from https://www.dst.dk/da/Statistik/dokumentation/statistikdokumentation/innovation-i-erhvervslivet/statistisk-behandling

DST. (2016d). Danmarks statistik; strategi og kvalitet. Retrieved April 9, 2019, from https://www.dst.dk/da/OmDS/strategi-og-kvalitet

DST. (2016e). Danmarks statistik; vores mission, vision og værdier. Retrieved April 9, 2019, from https://www.dst.dk/da/OmDS/strategi-og-kvalitet/vaerdier

DST. (2018). Danmarks statistik; innovation i erhvervslivet 2016. Retrieved February 10, 2019, from https://www.dst.dk/da/Statistik/nyt/NytHtml?cid=25287

DST. (2019). Documentation of statistics: Purchases and sales by enterprises. Retrieved April 27, 2019, from https://www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/purchases-and-sales-by-enterprises

Ernst, H. (2001). Patent applications and subsequent changes of performance: Evidence from time-series cross-section analyses on the firm level. *Research Policy*, *30*(1), 143–157.

Fagerberg, J. (2004). Innovation: A guide to the literature. Georgia Institute of Technology.

Fagerberg, J., Mowery, D. C., Nelson, R. R., et al. (2005). *The oxford handbook of innovation.* Oxford university press.

Flint, E. (2012). *Investigating the relationship between labour market status and minor psychiatric morbidity: Longitudinal and spatial analysis of the british house-*

hold panel survey, 1992-2008. (Doctoral dissertation, UCL (University College London)).

Freeman, C. (1991). Networks of innovators: A synthesis of research issues. *Research policy*, *20*(5), 499–514.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth*, *1*(1), 92–250.

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic management journal*, *21*(3), 203–215.

Gunday, G., Ulusoy, G., Kilic, K., & Alpkan, L. (2011). Effects of innovation types on firm performance. *International Journal of production economics*, *133*(2), 662–676.

Gurmu, S. & Pérez-Sebastián, F. (2008). Patents, r&d and lag effects: Evidence from flexible methods for count panel data on manufacturing firms. *Empirical Economics*, *35*(3), 507–526.

Hardin, J. & Hilbe, J. (2013). *Generalized estimating equations, second edition*. New York: CRC Press.

Harland, C. M. (1996). Supply chain management: Relationships, chains and networks. *British Journal of management*, *7*, S63–S80.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251–1271.

Heidenreich, M. (2009). Innovation patterns and location of european low-and medium-technology industries. *Research Policy*, *38*(3), 483–494.

Helfat, C. E. & Peteraf, M. A. (2003). The dynamic resource-based view: Capability lifecycles. *Strategic management journal*, *24*(10), 997–1010.

Hirschey, M. & Weygandt, J. J. (1985). Amortization policy for advertising and research and development expenditures. *Journal of Accounting Research*, 326–335.

Ho, Y. K., Tjahjapranata, M., & Yap, C. M. (2006). Size, leverage, concentration, and r&d investment in generating growth opportunities. *The Journal of Business*, *79*(2), 851–876.

Hsu, F.-J., Chen, M.-Y., Chen, Y.-C., Wang, W.-C., et al. (2013). An empirical study on the relationship between r&d and financial performance. *Journal of Applied Finance & Banking*, *3*(5), 107–119.

Jaruzelski, B., Dehoff, K., & Bordia, R. (2005). Strategy+business.

Jha, A. K. & Bose, I. (2016). Innovation in it firms: An investigation of intramural and extramural r&d activities and their impact. *Information & Management*, *53*(4), 409–421.

Jiménez-Jiménez, D. & Sanz-Valle, R. (2011). Innovation, organizational learning, and performance. *Journal of business research*, *64*(4), 408–417.

Kafouros, M. I., Buckley, P. J., Sharp, J. A., & Wang, C. (2008). The role of internationalization in explaining innovation performance. *Technovation*, *28*(1-2), 63–74.

Karabulut, A. T. (2015). Effects of innovation types on performance of manufacturing firms in turkey. *Procedia-Social and Behavioral Sciences*, *195*, 1355–1364.

Kennedy, P. (2003). *A guide to econometrics*. MIT press.

Klomp, L. & Van Leeuwen, G. (2001). Linking innovation and firm performance: A new approach. *International Journal of the Economics of Business*, *8*(3), 343–364.

Kondo, M. (1999). R&d dynamics of creating patents in the japanese industry. *Research Policy*, *28*(6), 587–600.

Kotler, P., Armstrong, G., & Harris, L. C. (2016). *Principles of marketing european edition 7th edn*. Pearson Education UK.

Langowitz, N. S. & Graves, S. B. (1992). Innovative productivity in pharmaceutical firms. *Research-Technology Management*, *35*(2), 39–41.

Li, J., Chen, D., & Shapiro, D. M. (2010). Product innovations in emerging economies: The role of foreign knowledge access channels and internal efforts in chinese firms. *Management and Organization Review*, *6*(2), 243–266.

Liang, K.-Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22.

Lööf, H. & Heshmati, A. (2006). On the relationship between innovation and performance: A sensitivity analysis. *Economics of Innovation and New Technology*, *15*(4-5), 317–344.

Lööf, H., Heshmati, A., Asplund, R., & Nåås, S.-O. (2001). *Innovation and performance in manufacturing industries: A comparison of the nordic countries*. SSE/EFI working paper series in economics and finance.

Mairesse, J. & Mohnen, P. (2003). R&d and productivity: A reexamination in light of the innovation surveys. In *Druid summer conference* (pp. 12–14). Citeseer.

Markus, A. (2013). *External knowledge sourcing and firm innovation: Essays on the micro-foundations of firms' search for innovation*. Copenhagen Business School [Phd].

McCullagh, P. & Nelder, J. A. [John A]. (1989). Generalized linear models. chapman & hall london.

Mitchell, G. R. & Hamilton, W. F. (1988). Managing r&d as a strategic option. *Research-Technology Management*, *31*(3), 15–22.

Mohnen, P., Mairesse, J., & Dagenais, M. (2006). Innovativity: A comparison across seven european countries. *Economics of Innovation and New Technology*, *15*(4-5), 391–413.

Nelder, J. A. [John Ashworth] & Wedderburn, R. W. (1972). Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, *135*(3), 370–384.

OECD. (2002). *Frascati manual 2002: Proposed standard practice for surveys on research and experimental development*. OECD Publishing, Paris. doi:https://doi.org/10.1787/9789264199040-en

OECD. (2004). Oecd glossary of statistical terms. Retrieved February 10, 2019, from https://stats.oecd.org/glossary/detail.asp?ID=3494

OECD. (2010). *The oecd innovation strategy: Getting a head start on tomorrow*. OECD Publishing, Paris. doi:https://doi.org/10.1787/9789264083479-en

OECD. (2015). *Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development*. OECD Publishing, Paris. doi:https://doi.org/10.1787/9789264239012-en

OECD/Eurostat. (2005). *Oslo manual: Guidelines for collecting and interpreting innovation data* (3rd ed.). OECD Publishing, Paris. doi:https://doi.org/10.1787/9789264013100-en

OECD/Eurostat. (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD Publishing, Paris. doi:https://doi.org/10.1787/9789264304604-en

Ortt, J. R. & van der Duin, P. A. (2008). The evolution of innovation management towards contextual innovation. *European journal of innovation management*, *11*(4), 522–538.

Pandit, S., Wasley, C. E., & Zach, T. (2011). The effect of research and development (r&d) inputs and outputs on the relation between the uncertainty of future operating performance and r&d expenditures. *Journal of Accounting, Auditing & Finance*, *26*(1), 121–144.

Papke, L. E. & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*(1-2), 121–133.

Parisi, M. L., Schiantarelli, F., & Sembenelli, A. (2006). Productivity, innovation and r&d: Micro evidence for italy. *European Economic Review*, *50*(8), 2037–2061.

Park, H. M. (2011). Practical guides to panel data modeling: A step by step analysis using stata. *Public Management and Policy Analysis Program, Graduate School of International Relations, International University of Japan*, 1–52.

Phelps, C. C. (2010). A longitudinal study of the influence of alliance network structure and composition on firm exploratory innovation. *Academy of management journal*, *53*(4), 890–913.

Porter, M. (1990). E.(1990) the competitive advantage of nations. *Harvard Business Review*, *68*(2), 73–93.

Schumpeter, J. A. (1934). *The theory of economic development. an inquiry into profits, capital, credit, interest, and the business cycle*. Harvard University Press.

Sharif, N., Baark, E., & Lau, A. K. (2012). Innovation activities, sources of innovation and r&d cooperation: Evidence from firms in hong kong and guangdong province, china. *International Journal of Technology Management*, *59*(3/4), 203–234.

Shaukat, S., Nawaz, M. S., Naz, S., et al. (2013). Effects of innovation types on firm performance: An empirical study on pakistan's manufacturing sector. *Pakistan Journal of Commerce and Social Sciences (PJCSS)*, *7*(2), 243–262.

StataCorp. (2017). *Stata statistical software: Release 15*. College Station, TX: StataCorp LLC.

Twisk, J. W. (2013). *Applied longitudinal data analysis for epidemiology: A practical guide*. cambridge university press.

Wang, M. (2014). Generalized estimating equations in longitudinal data analysis: A review and recent developments. *Advances in Statistics*, *2014*.

Wedderburn, R. W. (1974). Quasi-likelihood functions, generalized linear models, and the gauss—newton method. *Biometrika*, *61*(3), 439–447.

Wieser, R. (2005). Research and development productivity and spillovers: Empirical evidence at the firm level. *Journal of Economic Surveys*, *19*(4), 587–621.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data mit press*. Cambridge, MA.

Wooldridge, J. M. (2003). *Introductory econometrics: A modern approach*. South-Western CENGAGE Learning.

Wu, G. A. (2012). The effect of going public on innovative productivity and exploratory search. *Organization Science*, *23*(4), 928–950.

Zorn, C. J. (2001). Generalized estimating equation models for correlated data: A review with applications. *American Journal of Political Science*, 470–490.