The relationship between R&D expenditures, patents granted and firm performance in European high-tech industries.

Radboud Universiteit

Master's Thesis

Gijs Elings, S4499387
Master Economics: Accounting and Control
Supervisor: Thomas Niederkofler

Abstract

This study assesses the relationship between R&D expenditures, patents granted and firm performance. Research objects are high-tech European industries as identified by Eurostat which are; the pharmaceutical, electronics-telecommunication, computer and office machines, aerospace and scientific instruments industries. The data over these industries are collected over the UK, France, Germany and the Netherlands. The study is built up in a two-step analysis. First, a simplified DEA analysis with no modification is conducted with input variable *R&D expenditures* and output variable *patents granted*. The calculated R&D is calculated for each industry in order to measure if there are significant differences. The second step of the analysis are panel data regressions of R&D efficiency and patents granted as independent variable on firm performance, which is operationalized as *operating revenue*. The results of the study are that there is a significant difference between R&D efficiency rates between the five industries and that patents granted is a better indicator for firm performance than R&D efficiency. However, the relationship remains weak and strongly industry dependent. This study contributes to the existing literature by strengthening the importance of industry dependency in R&D effects research.

Keywords: Firm performance, patents granted, R&D efficiency, R&D expenditures, data envelopment analysis, panel data, panel data regressions, fixed effects.

Table of contents

List of tables

1 Introduction

In the second half the 20th century, many economies transitioned from predominantly manufacturing industries to service industries. The subsequent transition to knowledge industries is characterized by new types of firms and services that grew due to rapid technological developments and inventions like the modern computer (Reenen, 2005; Powell & Snellman, 2004). These firms exploit scientific developments and invest in their own development projects to introduce new products and services. Also, some firms join forces to combine research and development (R&D) capabilities in order to finance or develop new products cooperatively (Cloodt et al., 2010). To protect these lengthy and expensive research and development trials, patents are employed by predominantly high-tech firms (Panagopoulos & Park, 2018). Patents are intangible assets that give a firm the right to manufacture a certain product a specified way in order to prohibit competitors from producing an imitation of these products.

The practice of patenting has been linked to economic growth and firm performance. An example is that it improves production and revenue rates (Maradana et al. 2017, 2019; Pradhan, 2020). Hasan & Tucci (2010) find that countries with many firms that have high quality patents in portfolio experience an increase in economic growth. However, the effect of patents on economic growth is often an indirect result as it is difficult to measure the direct effect of patenting or other innovation improving tools on economic growth (Park & Ginarte,1997). In addition, Albert & Png (2013) find that R&D is mostly beneficial to developed nations and spur economic growth as in less developed countries the institutional protection from patent infringement is weaker. Where the institutional environment is weak, R&D rates and patenting practices deteriorate as patents are less enforceable and require that strong legal protection. (Seitz & Watzinger, 2017; Slivko & Theilen, 2014). Also, the study by Acs & Sanders (2012) suggests that there is a ''sweet spot'' to how countries institutionalize their patent law and regulations. Strict and elaborate patent laws decrease the amount of competition between firms in the market, which leads to reduced growth. The link between economic growth and patenting thus is proven by some studies, but is also still contested in some regard (Hall, 2007). Patenting may decrease technological advancements as some research is accumulative in nature. When the newest innovation is patented, it will possibly lead to a deterioration in innovation in the sector (Albert & Png, 2013) and limit competition (Arora et al. ,2008). Patents are thus tools to foster and protect innovation. By employing patents, firms are incentivized to increase R&D investments and activities (Lindman & Söderholm, 2016). However, the influence of patents on innovation rates defers between different industries (Allred & Park, 2007) and countries should be aware of how they try to foster innovation and patent output, as R&D incentives may also decrease the quality of R&D outputs (Chen & Zhang, 2019) as firms are incentivized to put as much patents out as possible. Patents are thus widely used in the current economics and business and lead to chances but may also prohibit economic and technological developments. In addition, the influence of patents on economic or firm growth remains under scrutiny as the influence of patenting is difficult to measure reliably.

There thus are some studies that stress the link between R&D input, patent output and the subsequent improvement in firm performance or countries performance. However, the degree to which R&D processes are effective and how these are measured are dependent on the models, inputs and outputs that are used in this calculation. With this study, I wish to add to the existing literature by assessing the relationship between R&D expenditures, patents granted and firm performance by employing a new approach to the subject. With a two-step analyses, I aim to analyse how well R&D expenditures are turned into patents granted and the subsequent influence of patents granted and R&D efficiency on firm performance.

There are a number of novelties in this study compared to similar studies. First, this approach hasn't been conducted before in the existing literature, where the chain is analysed from R&D input up to firm performance with patents granted as step between. Second, in this study there are two hypothesis that aim to explain whether patents granted or R&D efficiency is the better predictor for firm performance. Third, this study assesses the difference in R&D efficiency in an uncommonly researched domain; European high-tech industries that are based in the UK, Germany, France and the Netherlands. The study aims to add to the current debate regarding R&D efficiency and patents granted and its influence on firm performance. The relationship between patenting and firm performance is under scrutiny as the influences of patents are difficult to measure. In addition, there are indications in the literature (I refer to chapter 2) that R&D efficiency is dependent on industry specific factors and that further research has to be conducted in order to define the similarities and differences. As will become clear in the literature review of this study, the relationship between R&D expenditures, patents granted and firm performance is more complicated than one may think. The research question of this study is therefore; ''What is the relationship between R&D expenditures, patents granted and firm performance in European high-tech industries?"

The scope of the thesis consists of the UK, France, The Netherlands and Germany and their presence in the five dominant high-tech industries in Europe. The high-tech industries are selected on the basis of an Eurostat (2017) . The scientific relevance of the paper is that I try to achieve insight by analysing the relationship between R&D investments, patents granted and firm performance. I do this by assessing the European high-tech industry with a two-step analyses which is a new approach to the existing knowledge gap. The practical relevance of the paper is that management is advised on the efficiency of R&D expenditures and their subsequent influence on firm performance. For the examination of the efficiency of R&D expenditures on patents granted, an simplified DEA model is used, which measures how input variables are transformed into output variables. The relationship between patents granted and R&D efficiency on firm performance is examined using an panel data fixed effects regression.

The main expectations of the outcomes of the study is that there are significant differences in R&D efficiency between different industries, suggesting that R&D expenditures are more efficiently turned into patents granted in one industry than in the other. Second, there is an expectation that on the basis of industry differences the relationship between R&D efficiency and patents granted on firm performance is different for every industry. This indicates that the relationship isn't as linear as one may expect and that there thus are underlying factors that explain these differences between industries.

The findings of the study indicates as expected that there are differences in R&D efficiency among industries and that the notion that R&D efficiency linearly leads to an improvement in firm performance is false. In addition, the study shows that among industries the relationship between R&D efficiency and patents granted on firm performance is different. This strengthens the initial discussion that the process is much more complicated and context dependent than one may initially expect. The main conclusion that derives from the study is that the relationship between R&D expenditures, patents granted and firm performance is dependent on underlying factors like possible industry characteristics. The subsequent additional conclusion that is drawn is that comparative studies that aim to find similarities among industries and the influence of R&D efficiency in general are flawed, as there thus is evidence that there is a strong link between underlying characteristics and the performativity of R&D in these industries.

The study is structured as follows. first the theoretical background and literature review is displayed in the next chapter. In this chapter, the relevant literature is studied and prior research is assessed on the basis of its use for this research. In this literature review, 3 hypotheses are formulated that are to be tested and that should help answer the research question. In the following chapter the study design is presented and there is an elaboration on the used methodology. Subsequently, the research is conducted and analyses are presented and argued. Next, the discussion that elaborates on the flaws of the study and possible directions for future research and the conclusion that summarizes the study and answers the research question. Finally, an overview of the amount of missing values and known values for R&D expenditures are assessed with an robustness test.

2 Literature review

In this section of the study, the relevant literature is discussed in an literature review which goes over articles regarding the research subject. From the literature review derives the hypothesis that will be tested in this study.

2.1 Patents in practice

Patents are used as a legal protection against infringement from competitors. When a product is patented, competitors aren’t allowed to replicate the novelty without consent of the patent holder. However, patents are also used for a wide variety of other purposes.

Patents are used as signalling devices (Useche, 2014, Vo, 2019). Signalling is the use of a certain object that represents value or assurance (i.e. an annual report) to convey a message to stakeholders outside of the firm. Patents thus may be used to improve a company’s image or draw attention to the firm. Ciftci & Zhou (2016) find in their research that in markets where intellectual protection is strong, investors are likely to prefer the disclosure of patent output over the disclosure of R&D expenses. Vismara (2013) finds that new firms can use patents to signal ‘’technological maturity’’, which is a more persuasive factor for investors than for example firm size. Bessler & Bittelmeyer (2008) find that start-ups with patents indicate long and short term performance succes. Holgerrson & Granstrand (2017) find that patenting is a common tool to improve corporate image, but it is deemed less important than for instance product protection. The paper also underlines the notion that patents are widely used to increase stock value and to try to increase investment chances. An example of a similar conclusion comes from the paper by Noel & Schankerman (2013). With a study in the software industry they show that patenting can be used as a tool to increase stock value. They find that a firm that has a number of patents in portfolio can measure a ‘’patent premium’’ in their stock value.

Apart from signalling value to possible investors, patents are used as tools for organizational strategies. A strategic benefit of patenting is the ability to lock out competitors of your market segment by patenting knowledge or processes. The ability of locking out competitors of the market is elaborated on by Guellec et al. (2012), who finds that there are strategies employed by firms to use pre-emptive patenting to lock competitors out of certain activities. For example, procedures are patented that aren’t on the edge of the technological frontier, but are just patented so that others can’t patent them and thus lock out possible competitors. Holgerrson & Granstrand (2017) emphasize the importance of this, as they find in their paper that the ability to lock out competitors is regarded as just as important as protecting newly developed products and services or preventing infringement of competitor patents.

However, there are also reasons why patenting is considered a flawed strategy for firms or is a problematic practice. When a patent is accepted by a patent office, the patent description is publicly available. Cohen et al. (2000) find that firms tend to choose not to patent a product because they fear that company or R&D intelligence is compromised by the patent descriptions, which makes it easier for competitors to catch up. They also find that the cost to put out patents may not weigh up to the benefits as patenting costs may be high. In addition, some fields have to deal with ‘’patent thickets’’, which are overlapping procedures or inventions that are patented and lead to difficult innovation procedures as often patents need to be licensed in order not to infringe on them. (Cockburn et al. 2010). Cockburn finds that these thickets decrease the innovative performance of companies.

Lastly, patents are used in different industries for different purposes and face different challenges (Allred & Park, 2007). In the software industry, patents are used to strengthen the protection from competitor infringement as there inherently is copyright on computer code, but this is fairly easy to bypass (Chabchoub & Niosi, (2005). Also, as a substitute for patents, open innovation and open source platforms emerge in the software industry. Harison & Cowan (2004) mention that open source strategies have an indirect positive effect on firm revenues. However, a firm has to offer transparency regarding their codes or product plans. This deliberate choice not to engage in patenting thus may be beneficial and in some cases a more preferable alternative. In the pharmaceutical industry, patents are used because R&D costs are very high but these often lead to monopoly positions for certain drugs and other products (Grinols & Henderson, 2007). Moral implications arise as manufacturers are more concerned with profit than public health demands (Barton & Emanuel, 2005). The telecommunication industry is characterized by a fast evolving industry where innovation is very diverse and the importance of specific products and their share of the market guides innovation (Noh et al., 2016).

2.2 R&D efficiency and patents granted

R&D efficiency is regarded to be the measure of how well input is turned by the R&D process into output. However, the relevant input and output variables and methods that are used to correctly measure R&D efficiency defer from study to study. Wu et al. (2019) determines R&D efficiency as how well different input indicators are eventually turned into patent output. The authors argue that R&D efficiency can be measured using different tools, of which Stochastic Frontier Analysis (SFA), Data Envelopment Analysis (DEA) and specialized customized tools are the most common in the industry. The basic premise of SFA is that it is an econometric model that measures how effective productivity is designed within a firm or a collection of firms and aims to assess whether production can be improved. DEA is one of the most commonly used tools in the field and is a formula that calculates how effective input is turned into output. Input and output can consist of multiple variables and the importance of variables can be stressed as the method allows one to put extra on them in the calculation. One of the reasons why it is commonly used in R&D efficiency studies is that it can be customized to better suit the object of analysis or study design. For instance, it is possible to account for constant or variable output and to conduct an additional stage of calculation to account for additional effects. Lastly, there are a number of different designs that are used to assess R&D efficiency in specific situations. For instance, the paper by Thomas et al. (2011) measures efficiency by analysing differences in R&D expenditures and patents ratio's over time for different states in the United States. Their notion is that to follow the effects of R&D expenditures, one should assess developments over time rather than on one point in time.

The literature thus shows that there are multiple ways to calculate R&D efficiency. In addition, there are per study different variables that are used as input and output variables. Wu et al. (2019) states in its literature review that R&D expenditures is one of the most agreed upon variables in the relevant literature to play an important role in R&D efficiency as it is the primary input for the R&D process, which is compared to a production process. Other studies regarding R&D efficiency that stress the importance of R&D expenditures as one of the main input variables for research are Sharma and Thomas (2008), Rousseau and Rousseau (1998) and Rousseau and Rousseau (1997). Other input variables that are common are for instance number of scientists (Wu et al. 2019; Wang, 2007; Wang & Huang, 2007) and a countries GDP (Rousseau and Rousseau, 1997, 1998).

According to Wu et al. (2019), patenting is used as an R&D efficiency output variable as it is unique in reflecting the technological and business capability of a firm or country. The authors state that especially in high-tech industries, patents are used as indicators of for a firms technological skills and R&D strategies that are employed. In addition, they find that patents is the best measurable output to determine R&D efficiency because it says the most about the possible benefits that a firm may experience from its R&D development. The paper by Thomas et al. (2011) stresses that patents are the most reliable indicator of innovation or R&D success and that it is a reliable measure as patent output per region or country is well documented. For their research, they uses the variable patents granted. Which is operationalized as patents that are accepted by a patents office and offers active protection from infringement by competitors. In the paper, it is argued that patents granted are useful as it is the only patent specification that states that it is in effect and thus offer active protection. In addition, one may thus suggest that patents granted is the only patent specification which may possibly lead to economic benefits as it are patents that offer that protective status. Patents granted remain for that reason an often used variable in R&D efficiency research (Johansson et al. 2015; Sharma & Thomas, 2008; Wang & Huang, 2007; Wang, 2007).

R&D efficiency research is divers (Karadayi & Ekinci, 2019) and there is dispute about which factors influence R&D efficiency rates and whether there is a linear connection between input and output. The study by Lee & Park (2011) state that differences between R&D efficiency rates are created because of country characteristics. For instance, They find that countries have different policies regarding R&D efficiency motivation. This may lead to differences in R&D efficiency rates between industries and countries. The study by Cullman et al. (2012) argues that differences in R&D efficiency may be influenced by their finding that industries with low entry barriers experience higher R&D efficiency than industries with high entry barriers, as start-ups are more common to emerge. In addition, the study states that it is expected that R&D differences occur due to country specific characteristics as different countries operate on different levels of the technological frontier. An example of this is that they find that Germany is one of the countries that is on the edge of technological research and it is thus implied that other countries Grant et al. (2019) argues in their study that the most important indicator for R&D efficiency for a firm is their previous investments in R&D processes. In addition, the study suggests that there is no linear connection between R&D expenditures and patent output, as it is likely that R&D productivity is also dependent on other variables than only R&D expenditures. Lastly, the paper by Grant et al. (2019) argues that in some industries due to industry specific characteristics, there may be differences in how linear the relation is between R&D input and output. The paper by Hashimoto and Haneda (2008) stresses this notion and suggests that rather than a linear relationship between input and output there may be in some industries an U-curve, where R&D efficiency doesn't keep following a linear increase when input increases. In their study, they find support for this claim in the pharmaceutical industry.

One can thus expect the relationship between R&D efficiency input and output to be more complicated than input follows output. I wish to test the hypothesis that there are differences in R&D efficiency levels in the largest high-tech industries In the European Union, meaning that R&D efficiency is different per research object. The contribution that is aimed by testing this hypothesis is to assess what difference remains when country specific influences are minimized, in order to assess what the stretch is of underlying factors. Testing the hypothesis means that it becomes more clear whether industries should be compared or that it would be more beneficial to only conduct individual industry studies as comparison is too difficult. In addition, because the study data is gathered from the Netherlands, France, Germany and the UK the study adds to the Cullman et al. (2012) study by minimizing country effects as one can thus determine whether differences in R&D efficiency occur because of differences between industries. An additional incentive to study this hypothesis is that according to Grant et al. (2019) R&D efficiency between industry studies are conducted in limited amounts.

Hypotheses 1: There is a significant difference in R&D efficiency between European high-tech industries.

2.3 Innovation and firm performance

The subsequent influence of R&D efficiency and patenting on firm performance has been studied but also remains under scrutiny as the relationship is difficult to measure.

Most of the research focusses on a definition of innovation rather than the specific role of patents. Frietsch et al. (2014) measured the economic importance of patents by measuring the influence of patent applications on export figures, for which they find a strong relation. Maradana et al. (2017, 2019) finds that patents are associated with increases in firm production rates and revenue. Bloom & Reenen (2002) Also find that patents have an effect on production rates. in addition, they suggest that the relationship may be indirect as patenting may over time increase a firm's stock price. This indirect effect is a common theme in studies (i.e. Chen et al., 2019) that assess the relationship between R&D efficiency and firm performance as the effect of the independent variables is often difficult to measure as it is lagged. This means that logically, the results of patents and other innovations can only be perceived over a period of time. An example of this is the study by Czarnitski & Kraft (2010). The study aims to bypass the indirect effects of patents on firm performance by not assessing the relationship in a particular year, but by assessing a firms complete patent portfolio that is described as the ''patent stock''. The authors calculate the present value of the patent portfolio to determine whether it has had a strong effect on firm performance over the years, which they find it has. Jiménez- Jiménez & Sanz-Valle (2011) find a positive relation between firm performance and innovation and they argue that the majority of available literature finds a positive relationship between innovation efforts and firm performance. Agostini et al. (2015) notes that the relationship between patenting and firm performance is different as they find that not patents in themselves lead to an increase in firm importance. They point towards the indirect benefits of patenting practices, like the protection of R&D investments. They argue that the believed direct benefit of patenting thus is misunderstood and that the influence of patenting is thus a much more subtle one.

There thus are indications that patenting may have a beneficial influence on firm performance, but the relationship is under scrutiny as effects are difficult to measure. Previously mentioned Maradana et al. (2017 and 2019) find an influence of R&D efforts on production and revenue rates, but it is an indirect effect. On one side, patents in themselves have an indirect influence on some factors of the organization (i.e. R&D success may improve future investments in R&D processes) and in other cases it has a direct influence on firm performance as newly patented technology may improve internal organizational processes, reduce productional costs or lead to the introduction of a new product. Oliveira et al. (2018) mention that innovation may lead to financial performance, but for instance new products don't necessarily lead to improvements in a firms financial performance. The paper as well provides a literature review that elaborates on the discussion whether innovation leads to improvements in firm profitability or not. In their literature review they mention a study by Simpson et al. that argues that the relation between innovation efforts and firm performance may be negative due to the risk and high costs that are associated with R&D projects. Mahajan et al. (2018) find that the perceived benefits of R&D efficiency may be smaller than perceived, as R&D efficiency seems to be related to economies of scale. If R&D efficiency lead to new products, the profitability of these products (and thus the influence of R&D efficiency on firm performance) is dependent on the economies of scale factor, which is different for each industry.

Also, there are a number of mitigating factors related to patent output and firm performance. The paper by Jiménez- Jiménez and Sanz-Valle (2011) find that there is a positive relation between innovation and firm performance but this relation is influenced by the industry in which the firm operates, firm size and level of firm maturity. In addition, the study suggests in their literature review, determining whether innovation has an effect on firm performance is complex. Measuring the relation between firm performance and innovation or patent output is difficult as many factors come into play. In addition, the literature review shows that there are multiple ways to define and assess the relation between patents, innovation, firm profitability and firm performance. However, the notion of industry importance is contested as Thornhill (2006) states that firms that engage in innovation activities are very likely to experience improvements in firm performance across all industries.

To contribute to the existing literature, I wish to further assess the relationship between R&D efficiency, patents granted and firm performance by assessing whether the products of innovation or the efficiency of the innovation process itself is an important indicator for firm performance. As argued above, the existing literature is undecisive in what the exact role of innovation is in relation to firm performance. Oliveira et al. (2018) thus argues that new products and thus patents granted don't necessarily improve firm performance. On the other hand, Maradana et al. (2017 and 2019) state that the effect is likely not that strong as patents predominantly affect firm performance in an indirect way. This notion is supported by Ambrammal & Sharma (2016). In addition, the paper by Ghapar et al. (2014) finds an relationship between patenting activity and firm performance, but the effect is very small and the data gathered is not strongly convincing. Grant et al. (2019) finds a strong relationship between R&D investments and firm performance. However, Jiménez-Jiménez and Sanz-Valle (2011), Thornhill (2006) and Frietsch et al (2014) find a strong relationship between innovation and firm performance. There thus is in the field discussion about the role of innovation in firm performance. What this study aims to make insightful is the relationship between R&D expenditures, patents granted and firm performance. However, the relevant literature doesn't seem to differentiate whether R&D efficiency, the process that defines how well input is turned into output, or patent output is the dominant determiner of firm performance. This is a relevant difference as when is found that R&D efficiency is a good determiner of firm performance, it means that the process itself is what determines firm performance benefits regarding R&D efforts. If patents granted is the better determiner, it means that not the process but the outcome of the process is the better determiner of R&D benefits for firm performance. This seems trivial, but for investors and stakeholders it might help in analysing R&D efforts and determine whether one can expect future firm benefits. Also, there is no comparative study conducted that has assessed this difference before on the same research subjects. In addition testing both hypothesis are instrumental for answering the research question.

Hypotheses 2: There is a significant relationship between R&D efficiency and firm performance.

Hypotheses 3: There is a significant relationship between patents granted and firm performance.

2.4 Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA)

Chapter 2.4 goes briefly over two of the most common R&D efficiency analysis; Stochastic Frontier Analysis (SFA), Data Envelopment Analysis (DEA).

The Stochastic Frontier Analysis is employed in several efficiency studies (Wang, 2007; Wang & Wong, 2012; Hu et al. , 2011; Perelman 1995) and is operationalized by Wang (2007) as a tool that is used to measure how efficient a production process is in terms of production numbers with regards to input factors or how efficient the production is in reducing associated costs. In addition, he operationalizes DEA analysis as the general analysis which can be used to study processes where inputs variables are turned into output variables and measure which ratio is applied. The difference in the two thus lies in the application, goal of the analysis and how knowledge creation is conceptualized. On the methodological level, the SFA tries to build a production function, which takes the effects and interactions between all input and output variables into account and explains how efficient the production process of how inputs are turned into output is. The paper by Wang (2007) for instance is edged on determining how efficient R&D capital and manpower is turned into patents and publications. The conceptual scope that is applied is the believe that R&D efficiency is a production process where a maximum and minimum level of efficiency is assumed. The goal of the study is to determine where on the efficiency line the current production level is and how this may be improved or how this is influenced by external factors outside of the production of knowledge itself. Wang and Wong (2012) employ SFA to determine technical and production differences between countries and also aims to assess what the score is of their panel data on the efficiency line and determine how efficiency may be improved. Perelman (1995) stresses the importance of technological efficiency in using the analysis and Hu et al. (2011) emphasizes the production concept by also taking manpower into account as input variable.

The second widely used tool of analysis in R&D efficiency is Data Envelopment Analysis (DEA). Wu et al. 2019) operationalizes DEA as a model where a formula is used to measure the ratio between input indicators and output indicators. In relation to SFA research, DEA analysis approaches R&D efficiency processes as a process rather than an production process which may be efficient or inefficient. In general, a basic DEA model sets measured and weighted input variables against the measured and weighted output indicators. The reasons why DEA is often used, is because DEA is a comprehensive method where a large amount of input and output variables can be combined and it is highly customizable to better suit particular study designs. The model thus can be altered by adding or lowering weight to input or output variables or modify the formula as is. Examples of customized DEA models are BCC and CCR. CCR, which stands for Charnes, Cooper & Rhodes, which is a DEA modification that assumes that there is a constant return to scale. This means that increases and decreases in the amount of input variables will have an proportionate effect on the output variable. Another model that is used is BCC, which stands for Banker, Charnes and Cooper. This version of DEA assumes that the return to scale is variable, which means that the output variable can increase or decrease more than the increase or decrease in the input variables. By combining these different models or applying customized versions of the DEA model, one may account for vulnerabilities in the methodology or for specific variables that are particular for a certain industry in order to measure the desired data more accurately. One of the main weaknesses of DEA according to Wu et al. (2019) is that DEA relies more on the input and output variables that are chosen by the user themselves (when compared to SFA) and should therefore be supported with relevant literature.

The two commonly used methods to conduct R&D efficiency studies have been assessed and for this study it is clear that DEA better fits the needs of this study. The study aims to assess R&D efficiency as the efficiency ratio with which input (R&D expenditures) are turned into output (patents granted) rather than to assess the production function of a firm or industry and determine how R&D efficiency can be improved. In addition, the high level of modification of the DEA model can be used to better fit the overall goals of this study.

3 Methodology

In this chapter, the methodology and study design for this study are explained and assessed. First I present a general short overview of the study that I wish to conduct. Subsequently the variables are operationalized and a brief summary of how the data is gathered is summarized. Following data collection, sample tables are presented. In addition, the main research tools of this thesis, the simplified DEA analysis, independent two-sample t-test and panel data regressions are explained.

3.1 General study design

The goal of this paper is to gain insight in the relationship between R&D expenditures, patents granted and firm performance. In regard of Chua's (1986) distinction of scientific discourses, this master thesis falls in the positivistic research category. Positivistic research states that reality is quantifiable, measurable, can be experience objectively (reality is independent of the perceptions of the researcher) and there is a strong reliability on empirical research. This type of discourse believes that the connection between variables is stable and can be replicated. There thus is a stable connection between cause and effect. Also, positivistic research is characterized by the formulation of measurable hypotheses.

As research objects for this study, the largest European high-tech industries according to Eurostat (2017) are used. The industries that are the most prevalent in the European Union are the Pharmaceutical, electronics and telecommunication, computers and office machines, Aerospace and development of Scientific instruments industry. The countries from which data is gathered are The Netherlands, Germany, The United Kingdom and France. These countries were selected as there are relatively similar GDP differences (Eurostat 2019) technological capabilities and R&D efforts (Eurostat 2017). In this way, I try to minimize the accountability of external or country factors that have influence on R&D efficiency or firm performance in comparison to patents granted.

Data is collected from Orbis intellectual property. Orbis IP is preferred over regular Orbis as it offers patent information that is needed to determine how many patents were granted per firm per year. Without the database, the study could thus not be carried out. In addition, Orbis IP contains less financial information about firms than the regular Orbis database, but it does offer more R&D expenditures information than regular Orbis and provides the necessary information for this study. Firm patent and financial data is gathered over the period 2010-2018 in order to enable panel data analysis and account for the possible lagged effects of patents granted.

The data is analysed in two steps. First, a simplified DEA analysis is conducted to find the efficiency rate of R&D investments for the five industries in relation to patents. subsequently, independent two-sample t-tests are conducted to assess if there is a significant difference between industries. Secondly, I use an panel data regression to try to find a formula which explains the influence of patents granted on firm performance (hypothesis 3) and R&D efficiency on firm performance (hypothesis 2). These tests are performed using STATA and Excel.

3.2 Operationalization

The input and output variables of the R&D efficiency determination are *R&D expenditures* as input variable and *patents granted* as output variables. R&D expenditures are for this research operationalized as all costs that are related to R&D activities and that are necessary for conducting R&D operations. R&D expenditures and costs should be made in the period that is used in the research, which is 2010- 2018. However, it should be clear in what year which costs are incurred. The variable *patents granted* is chosen as a firm puts out multiple patents per year, but not all patents are granted as they lack a degree of novelty, are not properly administrated or were denied for a number of reasons by the patent office to which the patent was submitted. However, only patents that are granted contribute to an expected increase in revenue over time as they offer active protection from infringement. After R&D efficiency calculation, the relation between patents granted and firm performance is examined. *Operating revenue* is selected as firm performance indicator because revenue comes closest to increases and decreases in sales figures which are closely connected to for instance the introduction of new products and services. Net profit is a flawed measure as profit is also determined by organizational costs that are unrelated to the R&D process and the percentage of costs per company differ strongly.

As is argued by Wu et al. (2011), it is of great importance that the variables that are chosen for DEA analysis and these kinds of research are supported by enough literature as the selection of the wrong variables strongly suggests the studies outcome. *R&D expenditures* is a common (Wu et al. (2019; Thomas et al. 2011; Wang, 2007; Grant, 2019) variable In R&D efficiency studies, but rarely chosen as the only input determiner (Wu et al., 2019). R&D expenditures are selected as input variable for this study mainly because the interest of this study lies in the connection between R&D expenditures and patent output. Patents granted is used as output variable and is used in some studies (Wu et al., 2019; Johansson et al. 2015; Sharma & Thomas, 2008; Wang & Huang, 2007; Wang, 2007), which state that patents granted is one of the main indicators of innovation performance and as the indicator that may have an influence on firm performance. Firm performance is operationalized as operating revenue (Maradana et al. 2017,2019). Operating revenue is chosen as dependent variable as it comes closest to operations output(Ernst, 2001) and isn't influenced by additional firm costs which may influence the relationship between patent output and an firm performance indicator like profit margin.

The research objects of the thesis are four European countries and five industries per country. The countries that I select are Germany, The United Kingdom, France and the Netherlands. These countries are selected as they are comparable to each other in macroeconomic figures like i.e. GDP, part of the same socio-economic space (European free trade zone) and they are all developed countries. The industries that are examined are all high-tech industries, as these are the most knowledge intensive and dependent on their R&D processes. The high-tech industries that are selected are the pharmaceutical industry, electronics-telecommunication, computers-office machines, aerospace and scientific instruments developers. These industries have been selected as they are, according to the high-tech industry report from Eurostat (2017)[1], the most valuable high-tech industries in terms of export in the European economic space.

3.3 Data collection and preparation

Data is collected via Orbis Intellectual Property. The intellectual property database contains a very detailed account of patenting information and additional information about current owners, patent details, patent impact and initial applicants. In addition, it is possible to extract basic firm financials like ROE and operating revenue from it. To answer the research question and hypotheses of this paper, from Orbis IP, the R&D expenditures, patents granted in the period 2010-2018 and operating revenue is collected. In addition, for possible additional research, the output variables ROE, ROCE, profit margin, EBIT and EBITDA are collected. Data thus is gathered for 4 countries and 5 industries per country over the period 2010-2018. In addition, only financials have been extracted from firms that Orbis found that have patents which are granted in the period 2010-2018. One of the difficulties of using this database is that Orbis intellectual property exports patent data and firm financials separately in two different excel files, which for this research have to be reconciled. However, after reconciliation (as discussed in chapter 4), it is found that there is a discrepancy between the firms that are exported by financials and patent data. Because of this, observations were lost. Chapter four discusses in more depth how these two exports were reconciled. The second difficulty using Orbis is that it only recognizes three industry classification systems (i.e. NACE rev 2.). While the industry classification of the high tech industries is presented in Eurostat in an format that isn't supported by Orbis. Because of this, the industry classifications had to be translated from system to system. For elaboration on this, I refer to the discussion of this study. Three descriptive tables are presented below about the data that is gathered through Orbis. Table 1 presents the distribution of firms that were extracted per industry, per country. Table 2 shows how many R&D expenditure observations collected per year per industry. Table 3 shows how much operating revenue observations is collected per year and per industry. Countries are shortened by country code. NL = Netherlands, GE= Germany, FR= France, United Kingdom = UK. Industries are shortened by name. Pharmaceutical = pharma. electronics and

[1] Eurostat is an institution that is linked to the European Union and collects relevant statistics regarding the EU commissioned by the European commission.

telecommunications= elec-tel. computers and office machines = comp-off. Aerospace= aer. Scientific instruments are SI. These names are used throughout the study for tables.

Table 1: Distribution of firms across industries

Industries	Countries				
	GE	NL	FR	UK	Total
Pharma	856	166	330	521	1873
Elec-tel	1585	147	405	812	2949
Comp-off	1023	336	456	1246	3061
Aer	1410	129	329	617	2485
SI	1014	101	236	971	2322

Table 2. Total amount R&D expenditures known values per year per industry.

Industry	Number of total firms	2010	2011	2012	2013	2014	2015	2016	2017	2018
Pharma	1873	149	161	165	172	163	156	190	205	199
Elec-tel	2949	193	208	211	202	172	134	188	209	201
Comp-off	3061	101	100	109	113	89	73	107	116	116
Aer	2485	140	136	142	146	116	97	137	139	151
SI	2322	194	214	217	214	184	246	311	223	224

Table 3. Total amount of operating revenue (as indicator of Firm performance) known values per year per industry.

Industry	Number of total firms	2010	2011	2012	2013	2014	2015	2016	2017	2018
Pharma	1873	567	610	638	734	729	726	713	698	663
Elec-tel	2949	964	1005	1044	1162	1160	1126	1128	1103	1045
Comp-off	3061	608	638	666	748	696	685	661	632	579
Aer	2485	779	822	866	1018	1013	979	947	897	846
SI	2322	705	725	751	850	869	830	831	814	768

The data that is available in Orbis strongly differs from account to account. As table 2 presents, the amount of known R&D expenditures per industry per year is low. This is due to the fact that Orbis IP discloses that many firms choose to not disclose their R&D expenditure information. Another possible explanation comes from the paper by Enache & Srivastava (2018) who argue that R&D expenditures data may also be disclosed in the SG&A account. However, to recover this data from the accounts requires an amount of time which goes beyond the scope of this study. Therefore, this is a suggestion for future research in order to improve the number of R&D expenditures observations.

3.4 Data Envelopment Analysis (DEA)

DEA analysis is used in this study to calculate the R&D efficiency rate of the high-tech industries. Data envelopment analyses uses weighted input variables and measures how efficient these variables are turned into weighted output variables (Leitner, 2005). The application of DEA for R&D efficiency measurement is besides stochastic frontier (Wang, 2007) analyses common. (Wu et al.,2019 ;Wang et al.,2020 ;Lee et al., 2011) and Karadayi & Ekinci (2019) use DEA analyses for R&D efficiency calculation and they state that the analyses is widely used for a variety of purposes and determining R&D efficiency in a variety of ways.

As discussed in chapter 2.4 of this study, DEA analysis is preferred as this study aims to measure the level of R&D efficiency rather than formulating an production function formula (SFA approach). In addition, DEA analysis is highly customizable. The DEA analyses compares the input and output rates with a predetermined benchmark, which can be a certain value that is backed by literature or to compare it to other firms in the selection. In addition, there are 2 ways regular DEA analyses can be modified. CCR and BCC (Wu et al. 2019). CCR modification requires that there is a constant return to scale, which means that output can only be increased with the amount of input. BCC doesn't require a constant return to scale, and thus can be used in research where it is likely that there is a difference between the amount of input and amount that is put out.

There are studies that approach DEA analyses in a similar way but with different or multiple input and output variables. Wu et al. (2019) studies the R&D efficiency of the semiconductor industry in Taiwan and uses 6 different input variables (i.e. total assets, R&D expenditures) and 3 output variables like patents. The study examines 42 firms (operationalized as decision making units (DMU's)) and assesses the R&D efficiency rate for each of them. The study shows the different outputs DEA research offers like overall efficiency figures, efficiency figures compared to other firms in the selection and a quartile analyses to assess differences between groups in the selection. Sharma & Thomas (2008) conducted an study where variable and constant returns were both assumed (combining CCR and BCC) to analyses the R&D efficiency for a selection of countries in predominantly Europe and Asia. Rousseau and Rousseau (1998) study patents and citations and compare multiple input and output measures to each other to determine the most efficient country in terms of R&D efficiency in Europe. The paper by Cullman et al. (2012) further identifies studies that use DEA analysis or stochastic frontier analysis to assess R&D efficiency on industries and countries. These studies show that R&D efficiency or other efficiency studies with different input and output variables and different research objects are all appropriate for DEA analysis. The simple DEA analysis formula can be written as follows;

$$efficiency\ per\ DMU = \frac{output\ variable\ x * assigned\ weight\ variable\ x}{input\ variable\ y * assigned\ weight\ variable\ y}$$

The model above can be expanded by increasing the amount of input and output variables on either side. For this master's thesis, the relation between R&D expenditures and patents granted is conducted using an DEA model without modification. Usually, the DEA model calculates the ratio of R&D efficiency by comparing weighted input and weighted output variables. However, as there is only one input variable and one output variable in this R&D efficiency calculation, the DEA formula can be simplified to an efficiency function as displayed below.

$$efficiency = \frac{output}{input}$$

In this study, R&D efficiency for a an firm (decision making unit) can thus be calculated for period T as patents granted in period T divided by R&D expenditures in period T. Adding CCR or BCC modification isn't possible as these models rely on multiple variables and assigning weights to these variables.

$$R\&D\ efficiency\ per\ DMU = \frac{Patents\ granted\ in\ period\ T}{R\&D\ expenditures\ in\ period\ T}$$

3.5 Independent two-sample t-tests and panel data regressions

Three hypothesis are formulated in the literature review of this study. To assess the hypothesis, fitting tools for analysis and tests have to be selected. The first hypothesis, which aims to test whether there is a significant difference in R&D efficiency rate among industries is answered by conducting independent two-sample t-tests. The t-tests that are conducted should be independent as there is no evidence that the high-tech industries are influenced by each other and one industry thus isn't expected to have influence over the R&D efficiency rate of the other industries. Because the R&D efficiency means thus are unrelated, an independent two sample t-test is conducted. Secondly, the independent two-sample t-test achieves insight in the differences in means between the two selected industries. As is shown in chapter 4, by pairwise comparison of industries, it is possible to build a pairwise comparison matrix which shows between which industries the R&D efficiency means are different and which are similar. by gaining insight in cross-industry similarities and differences, it is possible to answer whether there are significant differences in the means between industries. To conduct the two sample t-test, the means of the R&D efficiency rates per industry per year are calculated. Then, per industry, the R&D efficiency means per year are assembled. This is done because the observations for R&D efficiency differs from year to year due to missing values or when input value is 0 and output value is 0. When the averages per year are taken and assembled per industry, there are 9 observations (for the years 2010-2018) that are reliable. An assessment was made how the two-sample t-tests should be conducted and with which data, but I found that this was the most reliable and accurate way to provide data for cross-industry independent two-sample t-tests.

The second part of the thesis is concerned with hypothesis 2 and hypothesis 3. These hypothesis test whether R&D efficiency or patents granted has an influence on firm performance in order to achieve insight in the relationship between research and development and firm performance. Both hypothesis are tested by using panel data regressions. Panel data regressions are drawn from panel data, which is data that follows a selection of firms over multiple years. Panel data thus collects values per entity over time in order to make developments over time insightful. As is discussed in the literature review, panel data is often used in studies that assess the influence of patents or R&D efficiency on firm performance as it is able to evaluate the lagged effect of innovation on firm performance. This study collects data over the same entities in five high-tech industries in four countries over the period 2010-2018, which is used as its panel dataset. Over this dataset, linear panel data regressions are conducted to assess the influence of R&D efficiency on firm performance (hypothesis 2) and the influence of patents granted on firm performance (hypothesis 3). In addition, the panel data regressions are conducted with the fixed effects assumption. The regression is used to find the best fitting (formula) description of the relationship between R&D efficiency and patents granted with firm performance and assess whether there is a significant value indicating an valid relationship.

To conduct the panel data regression and assure the validity of the regressions, the data is tested for BLUE assumptions. Best linear unbiased estimator (BLUE) (Casson & Franzco, 2014) assumptions about the data that is used in the dataset which should attain a number of standards. Because this is not an OLS regression, only three assumptions are extensively tested as they may possibly contaminate panel data. The dataset is tested for autocorrelation, heteroskedasticity and multicollinearity. Autocorrelation is tested by conducting an Woolridge test (Drukker, 2003; Woolridge, 2002). The Woolridge test indicates that autocorrelation is present when the test calculates an Prob > F value that is significant (<0.05). When autocorrelation is found, it is subsequently treated by differentiating the dependent and independent variables. After differentiating to the first degree, the Woolridge test is conducted again to assess if autocorrelation score is improved. Heteroskedasticity is examined by conducting an Wald test (Baum, 2001), which is a test that can be used to test for heteroskedasticity in fixed effects regressions and does accounts for datasets like panel regression data. If the Wald test calculates an Prob > chi2 value that is significant (>0.01), heteroskedasticity is assumed. To correct for heteroskedasticity, an robust fixed effects panel data regression is advised. Finally, multicollinearity is tested by calculating the VIF value (Shieh, 2010). Multicollinearity is debunked with VIF lower than 1. Which is checked by doing the regression, then the formula: $1/(1-R^2)$ to determine the Variance inflation factor (VIF). If the value is above 5, it is an indicator of multicollinearity. If the TOL value (1/VIF) is lower than 0.2 or 0.1, multicollinearity is expected. In this study however, none of the datasets that were used in the regression showed multicollinearity.

Important to disclose is that for the panel data regression conducted on the relationship between R&D efficiency and firm performance and the panel data regression conducted on the relationship between patents granted and firm performance two different datasets are used. The dataset that is used for the R&D efficiency panel data regression is presented in table 10 and the dataset that is used for the patents granted panel data regression is disclosed in table 12. Two different datasets were used because there is a limited amount of R&D efficiency data available to test hypothesis 2. However, the number of operating revenue and patents granted observations that is available to test hypothesis 3 is larger which leads to a more reliable regression. Because of the large differences in the availability figures of patents granted and R&D efficiency, it was chosen to use two datasets. One which is used for R&D efficiency regression which contains all R&D efficiency data with corresponding firm performance figures and one dataset which contains all patents granted figures with corresponding firm performance figures.

4 Data analyses and results

This chapter of the study consists of an elaboration of the different analysis tools that are employed and which data is gathered by using them. Furthermore, the results of the analyses are presented and the hypotheses that are formulated in chapter 2 are answered.

4.1 DEA analyses and R&D efficiency

The first analyses conducted is the simplified DEA analysis to determine the R&D efficiency of the five industries. Initially, two different datasets are extracted from Orbis IP. The first dataset consists of firm data of different industries per country. For instance, one of the extracted datasets consists of the pharmaceutical firms in France that have an accepted publication bit (the patent is granted). Financials that are collected are R&D expenses, operating revenue, EBIT, EBITDA, Return on Equity and ROCE over the period 2010-2018. The second type of datasets extracted consists of all individual patents that have a grant date between 2010 and 2018. Because the amount of patents granted per firm per year isn’t an variable that can be extracted by Orbis IP. The variable patents granted per firm per year thus has to be constructed manually. However, this leads to a problem. The datasets can be reconciled, but only partly as the database generates slightly different firms for the financials export and the patent data export. This means that the number of observations extracted does differ from the amount of observations that can be used in the study. The number of additional missing values accounts for 10 to 15% of the data per industry, per country. This problem is also created as firms book their R&D expenditure costs on one subsidiary, but register the patents on another subsidiary. Because of this, it is very difficult and time consuming to correct reconciliation errors. For future research, it is therefore advised to try to find another patenting database which may have the patent granted per year variable already available.

By following the analysis and data preparation steps above, the patents per year per firm were determined and reconciliated in the firm financials file in order to conduct the DEA. The analysis measures how well input (R&D expenditures) are turned into output (Patents granted) and thus aims to answer the first hypothesis, which is concerned with how R&D efficient European high-tech industries are and if there is any significant difference in between industries. In order to answer the hypothesis, DEA was conducted. As is discussed in chapter 3.4, DEA analysis in this study is conducted in a simple form with no specific modification. A problem that was encountered was that the program (STATA) couldn’t account for situations where there was input, but no output or where there was output but no input. Therefore, DEA analysis was carried out manually by using Excel. To handle the firms that had no or maximum efficiency (input, but no output and vice versa), these were assigned an efficiency of ‘’0’’ when their input didn’t generate any output or ‘’1’’ when firms generated patents without investing in research and development.

By conducting DEA analysis, the R&D efficiency rate is measured per decision making unit. The efficiency rate states how efficient input is turned into output as a rate between 0 and 1, where 0 is not efficient and 1 is fully efficient. This measure is the main measure that is used when conducting R&D efficiency analysis as it indicates an rate of productivity, which can thus be compared to other industries. When calculating R&D efficiency , the cases that were fully efficient and not efficient were kept in the selection as they should be included as they are valid data and should thus be taken into account and because they tell something about how efficiency works in the sector (i.e. an industry with high mean efficiency is likely to have more firms that have no input, but do experience output). These efficiency figures are used to answer hypothesis 1.

4.1.1 Difference in industry R&D efficiency means

To answer the first hypothesis, independent two sample t-tests are conducted over the means of the R&D efficiency rates per year, assembled per industry in order to compare industry averages. The two sample t-test is used as it measures the differences in means between two variables that don’t have a perceived effect on each other. Confidence interval that is used is 95%, so the results that have a T-value below 0,05 are deemed significant and it may be assumed that there is a significant difference. For the t-test results I refer to table 4,5,6,7,8 and significant values are summarized in table 9.

Table 4: Two sample t-test results for comparison with pharmaceutical industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
P value	-	0.0535	0.0001	0.0000	0.0001
T value	-	2.0840	5.3341	5.6538	5.4232
Mean	0.0982185	0.0684851	0.0378572	0.0338577	0.0411626
Standard error	0.0092444	0.0108672	0.0065268	0.0066429	0.0050226
Standard deviation	0.0277331	0.036016	0.0195803	0.0199287	0.0150679
Observations per industry	9	9	9	9	9
Degrees of freedom	16	16	16	16	16

Table 5: Two sample t-test results for comparison with electronics-telecommunication industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
P value	0.0535	-	0.0280	0.0152	0.0365
T value	2.0840	-	2.4161	2.7187	2.2822
Mean	0.0982185	0.0684851	0.0378572	0.0338577	0.0411626
Standard error	0.0092444	0.0108672	0.0065268	0.0066429	0.0050226
Standard deviation	0.0277331	0.0326016	0.0195803	0.0199287	0.0150679
Observations per industry	9	9	9	9	9
Degrees of freedom	16	16	16	16	16

Table 6: Two sample t-test results for comparison with computers and office machines industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
P value	0.0001	0.0280	-	0.6733	0.6935
T value	5.3341	2.4161	-	0.4295	-0.4014
Mean	0.0982185	0.0684851	0.0378572	0.0338577	0.0411626
Standard error	0.0092444	0.0108672	0.0065268	0.0066429	0.0050226
Standard deviation	0.0277331	0.0326016	0.0195803	0.0199287	0.0150679
Observations per industry	9	9	9	9	9
Degrees of freedom	16	16	16	16	16

Table 7: Two sample t-test results for comparison with aerospace industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
P value	0.0000	0.0152	0.6733	-	0.3934
T value	5.6538	2.7187	0.4295	-	-0.8772
Mean	0.0982185	0.0684851	0.0378572	0.0338577	0.0411626
Standard error	0.0092444	0.0108672	0.0065268	0.0066429	0.0050226
Standard deviation	0.0277331	0.0326016	0.0195803	0.0199287	0.0150679
Observations per industry	9	9	9	9	9
Degrees of freedom	16	16	16	16	16

Table 8: Two sample t-test results for comparison with scientific instruments industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
P value	0.0001	0.0365	0.6935	0.3934	-
T value	5.4232	2.2822	-0.4014	-0.8772	-
Mean	0.0982185	0.0684851	0.0378572	0.0338577	0.0411626
Standard error	0.0092444	0.0108672	0.0065268	0.0066429	0.0050226
Standard deviation	0.0277331	0.0326016	0.0195803	0.0199287	0.0150679
Observations per industry	9	9	9	9	9
Degrees of freedom	16	16	16	16	16

Table 9: Pairwise comparison of independent two sample t-test p value measuring efficiency.

	Pharma	Elec-tel	Comp-off	Aer	SI
Pharm	—	0.0535	0.0001	0.0000	0.0001
Elec-tel	0.0535	—	0.0280	0.0152	0.0365
Comp-off	0.0001	0.0280	—	0.6733	0.6935
Aer	0.0000	0.0152	0.6733	—	0.3934
SI	0.0001	0.0365	0.6935	0.3934	—

The results of the two sample t-tests helps understand some aspects of the industries and their differences and similarities in order to answer hypothesis 1. As is disclosed in tables 4-8, there are significant differences found between the high-tech industries that are tested in this study. Table 9 shows that there are significant differences between the pharmaceutical industry and the computer-office machines (0.0001), aerospace (0.0000) and scientific instruments (0.0001) industries. Also, significant differences are found between the electronics-telecommunications industry and the computer-office machines (0.0280), aerospace (0.0152) and scientific instruments (0.3934) industries. These findings suggest that the hypothesis is partly true, as the two-sample t-test shows that there seems to be two groups that are significantly related with each other or not. The pharmaceutical industry and electronics industry don't show any significant difference between each other, but both do have a significantly different efficiency rate than the computers-aerospace-scientific instruments group. The industries in the latter group show no significant differences among each other. The results thus imply that there is a difference between both groups. Possible directions for future research may be that the two groups have a different market structure, where pharmaceutical and electronics usually require high costs of entry and is largely in the hands of a few very large firms. The scientific instruments and computer industries may seem to have lower costs of entry and thus have a different rate of efficiency. Different efficiency figures may also be due to the fact that there are more firms that are able to produce patents without investments, which may be an indicator that entry barriers thus are low. Why aerospace belongs to this group as well is unknown and may be subject to future research. Another emerging conclusion comes from the average efficiency figures per year. The conclusion is that the pharmaceutical is the most efficient industry on average. However, this doesn't correlate with previous research (González & Gascón, 2004) that states that due to high R&D fail chances as other firms may patent a new drug before other firms and that subsequently investments become worthless it is expected that R&D efficiency is low.

Hypothesis 1: H0: there isn't a significant difference between European high-tech industries (Aerospace, Pharmaceutical, Electronics, Computers and Scientific instruments) efficiency rate.
H1: There Is a significant difference between European high-tech industries efficiency rate.

To answer the hypothesis, there are significant differences found in the means between industries, which indicates that the R&D efficiency level between industries significantly differs. For future research, it may be interesting to assess the differences between these groups in more detail. As there is some difference between industries, it may be stated that there is a significant difference in efficiency rates among the industries. However, these are between groups, therefore the H0 hypothesis is wrong and the H1 hypothesis is accepted.

4.2 Panel data regressions

4.2.1 Influence of R&D efficiency on firm performance

The second hypothesis is linked to the third hypothesis in the sense that the subsequent relationship of R&D efficiency with firm performance is assessed. However, with the second hypothesis, the importance of R&D efficiency as a function of firm performance is explained. An panel data regression was conducted to describe the relationship between R&D efficiency and firm performance, which is operationalized as operating revenue. in this panel data regression, firm performance is selected as dependent variable and R&D efficiency as independent variable. To determine to account for random or fixed effects in panel data, a Hausman test was conducted over all industries. The test showed a significant score indicating that fixed effects should be employed. Important to point out is that for this analysis, the fully efficient firms were removed as they posed outliers compared to the other efficiency data. Where applicable the data is treated per industry for the relevant assumptions to assure model fit and validity. First, table 10 follows with the descriptive analytics for the dataset. For the results of the regression, I refer to table 11.

Table 10: Descriptive table Firm performance and efficiency

Distribution of observations						Efficiency		Firm performance in Operating Revenue	
Industries	Total Group obs.*	GE	NL	FR	UK	Mean	SD*	Mean	SD*
Pharma	261	43	7	48	163	0.00000833	0.0001621	2030000000	764000000
Elec-tel	292	64	9	29	190	0.00000273	0.0000182	2290000000	9360000000
Comp-off	184	17	3	26	138	0. 00000750	0.000086	320000000	848000000
Aer	209	36	3	8	162	0.0000105	0.0001779	1640000000	7300000000
SI	342	55	5	27	255	0.0000117	0.0002781	624000000	1910000000

Total group observations are 1268 groups. Of which 215 firms originate from Germany, 27 firms originate from the Netherlands, 138 originate from France and 908 from the United Kingdom.

**Standard deviation

Table 11: Regression of R&D efficiency on firm performance by industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
Firm performance (Operating revenue)	0.311* 1.02	0.747* (0.32)	0.395* (0.85)	0.014* 2.49	0.274* (1.10)
Coëfficient	57500000000	-48500000000	-238000000	1960000000	-28000000
Standard error	56600000000	150000000000	279000000	786000000	34700000000
Fixed effects for years	Yes	Yes	Yes	Yes	Yes
Fixed effects for firms	Yes	Yes	Yes	Yes	Yes
R2 within	0.0000	0.000	0.000	0.0000	0.0000
R2 between	0.0003	0.0018	0.0000	0.0002	0.0001
R2 overall	0.0002	0.0006	0.0000	0.0000	0.0000
Number of observations	1273	1453	507	792	1185
Number of groups**	234	267	129	169	250
F value	(1233) 1.03	(1,266) 0.10	(1128) 0.73	(1168) 6.20	(1249) 1.20
Prob > F	0.3109	0.7473	0.3949	0.0137	0.2740

*First, the P value is disclaimed, second the T value.
** The amount of total groups differs from the summary table due to panel regression which doesn't take firms into account with only one known value in the timespan 2010-2018.

Table 11 shows the results of the panel data regression. There are no significant (p value < 0.05) values found for the pharmaceutical (0.311), electronics-telecommunication (0.747), computers and office machines (0.395) and scientific instruments (0.274) industry. However, for the aerospace industry, an significant relation was found (0.014). These values indicate that for the five industries, only in one of the industries a valid relationship between R&D efficiency and firm performance. This finding indicates that on average, one may not state that there is a relationship between R&D efficiency and firm performance, but that it seems that the relationship is a sense dependent on the industry. The R values indicate that on average one may expect that the possible influence of R&D efficiency on firm performance to be very small. For the aerospace industry, the R value (0.0000) is low but the result is still significant. The high T (2.49) and F value (6.20) solidifies that the significant result is valid, meaning that there is a significant effect but that the explanatory power of R&D efficiency on firm performance is very low. It is also possible that the variance in the aerospace industry dataset is too high. What can be concluded from the analysis on the level of results and analysis figures is that the overall relationship between R&D efficiency and firm performance cannot be established except for the aerospace industry, which may be only a weak significant link. In addition, all outcomes show low explanatory power (R2) and a relatively high prob > F Value, indicating regression strength.

This means that on average, in the European high-tech industries, R&D efficiency is not a valid predictor to determine future firm performance. On the one hand, this finding is interesting as the results suggest that how efficient R&D is conducted isn't necessarily a guarantee for success. It does thus stress that what is important may be the output that is generated rather than how efficient this process is. This may seem illogical as increasing R&D efficiency should lead to more output and potentially thus also more patent output, but may possibly be influenced by the firms that invest a lot in R&D, but don't generate any patents. In addition, there should be attention for the role of the fully efficient firms (efficiency ''1''). These firms don't invest in R&D, but are able to generate patents. The amount of fully efficient firms differs strongly from industry to industry. The scientific instrument industry comprises of the most fully efficient firms, which is on average 75 observations higher than in other industries. This in itself may be interesting for future research regarding R&D efficiency research. Why an significant effect is found in the aerospace sector, but not in other industries may be subject to future research.

Second, the finding suggest that the best fit would be on average a negative slope, where an improvement in R&D efficiency would lead to deterioration in firm performance. The result of the study seems paradoxical as an increase in R&D efficiency should logically lead to an increase in patent output and thus an increase in firm performance. A logical interpretation may be that the regression on firm performance and R&D efficiency is so weak that the coefficients don't say anything valid about the relationship in the first place. It may be that the model has difficulties to draw a best fitting line between all variables, meaning that the coefficients aren't relevant as there is in four out of five cases a very weak p value score. One may thus question whether these coefficients indicate something relevant or is just a misfit from the model. However, for the aerospace sector, a significant effect was found and the coefficient indicates an positive slope, meaning that in the aerospace sector, a higher R&D efficiency rate is correlated with firm performance.

What one may take from this analyses is that R&D efficiency is on average not a predictor for firm performance. However, the results stress that it may also be in some cases industry specific, meaning that in some cases there may be a strong relationship due to underlying industry factors. These findings call for future research. One direction for future research is whether R&D efficiency may be a more valid indicator of firm performance in different other industries. Another direction for research may be the role of efficiency ''1'' firms and how these patents are created as there is a hint (scientific instruments has twice the amount of efficiency ''1''cases than other industries) that it may be industry specific. Finally, a last direction for future research is the possible negative relationship between R&D efficiency and firm performance in the electronics sector.

Hypothesis 2: H0 There is no relationship between R&D efficiency and firm performance
H1: There is a relationship between R&D efficiency and firm performance

To answer hypothesis 2, on average one may say that there is no significant relationship between R&D efficiency and firm performance. Therefore H1: is rejected and H0 is kept. However, it is important to disclaim that there may be a relationship, but that this is likely only applicable to some industries.

4.2.2 Influence of patents granted on firm performance

Subsequently to the panel regression regarding R&D efficiency, an assessment was made of the relationship between patents granted and firm performance. The objective of hypothesis 3 is to gain insight in whether patents granted is a better predictor of firm performance than R&D efficiency and if there is a valid relationship between both. For this analyses, the analyses steps from the hypothesis 3 analyses are repeated. For this test, a Hausman test was conducted which indicated that fixed effects should be used. First, Table 12 containing dataset information is presented. Second, the results from the panel data regression of patents granted on firm performance are shown in table 13. For this analysis, data validity assumptions are tested and corrected where applicable.

Table 12: Descriptive table firm performance and patents granted

	Distribution of observations*					Patents granted		Firm performance in Operating Revenue	
Industry	Total Group obs.*	GE	NL	FR	UK	Mean	SD**	Mean	SD**
Pharma	1623	772	123	320	408	10.14103	56.70988	552000000	3660000000

Elec-tel	2712	1510	131	375	696	12.43277	143.9804	543000000	4130000000
Comp-off	2537	934	204	395	1004	1.849255	10.3294	128000000	662000000
Aer	2319	1353	114	317	535	6.733636	36.9955	508000000	3690000000
SI	2141	958	94	222	867	7.675471	38.0773	200000000	951000000

Total group observations are 11332 groups. Of which 5527 firms originate from Germany, 666 firms originate from the Netherlands, 1629 originate from France and 3510 from the United Kingdom.
*comprises of group observations, meaning the amount of firms that have multiple known values for timespan. 2010-2018
**Standard deviation

Table 13: Regression of patents granted on firm performance by industry

	Industries				
	Pharma	Elec-tel	Comp-off	Aer	SI
Firm performance (Operating revenue)	0.250* (1.15)	0.829* 0.22	0.022* 2.29	0.360* 0.92	0.004* (2.88)
Coefficient	4921222	618597.1	3147950	1140154	-574388.50
Standard error	4275032	2858701	1373622	1244364	199758.4
Fixed effects for years	Yes	Yes	Yes	Yes	Yes
Fixed effects for firms	Yes	Yes	Yes	Yes	Yes
R2 within	0.0288	0.0005	0.0213	0.0006	0.0070
R2 between	0.3035	0.0945	0.0012	0.0002	0.0382
R2 overall	0.2814	0.0908	0.0150	0.0005	0.0028
Number of observations	6006	9606	4426	6350	5661
Number of groups**	1034	1752	1011	1330	1119
F value	(1,1033) 1.33	(1.1751) 0.05	(1,1010) 5.25	(1,1329) 0.84	(1,1118) 8.27
Prob > F	0.2499	0.8287	0.0221	0.3597	0.0041

*First, the P value is disclaimed, second the T value.
** The amount of total groups differs from the summary table due to panel regression which doesn’t take firms into account with only one known value in the timespan 2010-2018.

As can be perceived in table 13, of the five industries, only two industries show a strong relationship between patents granted and firm performance. In the computers and office machines (0.022) and scientific instruments (0.004) sector a significant effect was found indicating that the number of patents granted per firm has an influence on a firms operating revenue (performance). For the pharmaceutical (0.250), electronics and telecommunications (0.829) and aerospace sector (0.360), no significant relationship was found. The R2 values of the significant values show that patents granted have some explanatory power over firm performance, but the influence remains small. In addition, the F and T values of both significant industries indicate strong support that the calculated p value is correct (strong evidence that the significance level is correct). Overall, an positive coefficient indicating a positive relationship between patents granted and firm performance can be perceived, which means that per extra patent granted, firm performance increases with one times the coefficient. For the scientific instruments industry, it was found that there is a negative coefficient, meaning that there is a significantly negative slope relationship between patents granted and firm performance. This finding may be subject for future research.

Additional findings are that the coefficients of the pharmaceutical, electronics and computer industry are fairly similar, while the aerospace coefficient is lower. This may indicate that the possible benefits from patents granted is different across industries. This may be due to underlying industry factors. For instance, as is suggested by González & Gascón (2004), the pharmaceutical industry should have the largest benefits from patents granted, as in that industry R&D trials are specific in the sense that when the product is patented, often one cannot build upon the knowledge done. It is a market where the winner takes all. Also, It is possible that the positive relationships are linked to the average R&D investments required per year to generate an patent (I refer to hypothesis 1), as the average required investment amounts also were fairly similar in these industries. This would also explain why patents granted in the scientific instruments industry may only generate modest benefits. Lastly, the difference in the scientific instruments industry is hard to explain, but It may be possible that it is due to industry specific characteristics. For instance, there is literature that states that some firms do not patent their innovations as they find that the cost of protection does not weigh up to the benefit of not exposing one's secrets (Cohen et al. 2000). In an industry where there is likely constant innovation that is based on previous innovations, it may be beneficial to choose not to patent innovations, which would lead to the negative connection. However, it is a finding that may call for further investigation or an industry study that is focused on the relationship between patents and firm performance in the scientific instruments sector. Another interesting direction for future research is that the relationship between patents granted and the computers industry is significant and that patenting does benefit firm performance in that sector. However, as discussed in the introduction of this study, there are indications that some software and computer firms choose to engage in open source innovation as substitute for patenting strategies. An suggestion for future research may thus be to research whether open source is a better alternative to patenting as it thus is proved that patenting can also benefit firm performance. To answer the hypothesis;

Hypothesis 3: H0: There is no relationship between patents granted and firm performance.
H1: There is a relationship between patents granted and firm performance

One may conclude that on the basis of the results that are disclosed above, that the relationship is true with the disclaimer that the assessment remains industry dependent as the computer and office machines and scientific instruments show a significant relationship while the pharmaceutical, electronics and aerospace industries didn't show any significant relationship.. This means that H0 is wrong and H1 is accepted with thus the disclaimer that it remains industry dependent in the sense of significance P value and direction of the Coefficient (positive/negative relationship).

Lastly, to compare the hypothesis 2 regression with the hypothesis 3 regression, it is found that there is more evidence that there is a significant relation between patents granted and firm performance than between R&D efficiency and firm performance. In addition, it can be perceived that the explanatory power of patents granted on firm performance is higher than that of R&D efficiency. In this regard, one may conclude that patents granted seems to be the better indicator of firm performance. However, more importantly, the comparison of both regressions shows that there is no universal linear connection between R&D and firm performance. In addition, when also looking at the first hypothesis, one may state that the link between R&D input and output, and subsequently firm performance isn't linear, industry dependent (likely because of underlying factors) and more complex than one may initially perceive it to be.

5 Discussion

In the discussion, the possible implications of the research, recommendations that arise from the results and possible directions for future studies are discussed.

5.1 Research object selection and examination

During this study, there were a number of implications that arose that might affect results and should be taken into account when assessing the results or conducting further research.

First, the selection of the research objects is based on an Eurostat (2017) report, which stated that the main high-tech industries in the European Union were the pharmaceutical, electronics, computer and office engineering, aerospace and scientific instruments development industries. There are two implications regarding these research objects; Translation of inter-organizational industry classifications and industry distribution in the EU. First, Eurostat (2017) reports their high-tech industry report in a certain industry classification language (SITC 4). The language comprises of a list of all types of industries and which activity belongs to which industry. However, because Orbis IP only recognizes a limited number of industry classification languages, a translation had to be made between SITC 4 and the chosen industry classification language in Orbis IP (NACE rev 2). Eurostat provides in RAMON correspondence tables, which are accessible online. The first implication of this is that this process is time consuming. A translation had to be made from NACE rev 2 to NACE rev 1.1, NACE rev 1.1 to ISIC rev 3, ISIC rev 3 to SITC rev 3 and subsequently SITC rev 3 to SITC rev 4. Second, translation means that some variables overlap between industries. Because of this, the electronics-telecommunication and computers and office machines industries were not fully translated as they had too much overlap. This overlap was solved by manually linking which variables belonged to which industry. The last implication of this is that it is likely that errors would occur, like the overlap problem. For future research, it is recommended to use a blend of sources to determine which types of activities belong to which firm instead of only relying on one source. In addition to the translation, the reconciliation of the financials and patents export lead to some loss of observations as there is either patent data or financial data available. If one of the two is obsolete, the observation has to be deleted as it can’t be used for R&D efficiency analysis. The second implication of the selection of the high tech industries by using the Eurostat report, is that the aerospace industry is marked as an important high-tech industry in the EU. However, the aerospace industry, which consists of the manufacturing and development of air, space and helicopter crafts and undercarriages is so specific that it is very large or well developed in some countries, and a lot less in other countries. After the collection of data, it seemed that for this particular industry the distribution was high when comparing the industry between countries. The distribution of Aerospace activities is dispersed and mostly done by small groups of very large companies. Because of this, it is difficult to get a representative sample of the aerospace industry. This implication and the implication that is discussed in the next paragraph made me decide to cancel the cross-country analysis part of this study as it was not possible to get a representative sample for each country.

5.2 Research and development data availability

Another implication of this study is that it points out the importance of choosing a viable database that is geared towards the research question. Initially, Orbis and Orbis intellectual property were used as databases in order to combine the best of both worlds. Regular Orbis contains many firm financials and additional information while Orbis Intellectual property is focused on patents and factors that are involved with them. It became clear during the process that Orbis intellectual property and regular Orbis were different in their selection of firms and classifications, even if the same industries were selected. In another subsequent study, it is advised to perform a single industry study or a selection of related industries rather than a cross comparison of different industries in different countries, as some databases may contain more information about a particular industry, financial information or additional firm information. In addition, as is shown in Orbis, much R&D information is restricted by firms and thus not disclosed in databases. This also stresses the importance of careful selection of information sources in order to make a valid and accurate analysis of industry R&D’s. the paper by Enache & Srivastava (2018) states that these problems with R&D expenditures figures also occur as sometimes R&D costs are booked on SG&A. They offer an methodology to filter these values out. However, it is beyond the scope of this study to conduct this elaborate analysis. However, if this study is to be replicated, one may choose to apply this method.

As argued under subchapter 5.1. and this subchapter, one of the main implications of this study and the whole process is that R&D related data availability and collection is very difficult. First, this study collected it’s data from regular Orbis and Orbis IP, where both had to be combined to reconcile the financials with the patents granted figures. However, after reconciliation, the amount of known R&D expenditures figures remained low. Then, a comparison was carried out with the gathered data and when data would only be gathered via Orbis IP, which showed that solely using Orbis IP would generate some more values. When switching to solely Orbis IP and then reconciling financials and patent figures, it showed that the total amount of R&D expenditure availability would likely be too low to draw any valid conclusions from. Subsequently, the cross-country comparison had to be dropped and just focus on industry differences. The implication of this study thus was that R&D expenditures and corresponding patent data was hard to come by and often restricted or incomplete. For future research, it is advised to take a larger amount of industries and countries in order to reduce the chance that R&D expenditures known values are too low. Also, as is one of the conclusions that comes forth from chapter 4, cross-country comparison for multiple industries may not lead to new findings as there is a strong hint that industry dependency is important. This means that it may be much more beneficial for future research to focus on single or comparison industry studies rather than cross-country studies.

5.3 Tools of analysis and methodological implications

Some implications for this study is focused on the methodology and analysis that were conducted. as is disclosed in chapter 4 of this study, the amount of autocorrelation and heteroscedasticity was high in the industry datasets. Heteroscedasticity may be improved in the future by assessing differences in industries by sorting firm sizes into different categories. Heteroscedasticity is now tackled by drawing an robust fixed effects regression, but may thus be prevented and possibly lead to better results as the amount of outliers per category is minimized. Another implication that is in line with panel data and assumptions correction is that the total amount of observations becomes lower as panel data regression only incorporates firms with values for multiple years, which means that some observations fall. Also, it is likely that the robust regression decreases the number of observations.

5.4 Recommendations

There are a number of recommendations that derive from this study in order to improve similar studies or point towards potential interesting future studies. An improvement for future studies may be that databases or sources of financial information should be dedicated to the industries involved rather than using a general database like Orbis. This may optimize the probability that R&D expenditures information isn't restricted. In addition, the method recommended by Enache & Srivastava (2018) may improve the number of R&D expenditure figures available by extracting them from the SG&A account. Another recommendation is that one makes an assessment of industry representations in all countries in order to make a fitting research design and data extraction plan. During the thesis, it was found that in some countries R&D figures were less available in for instance the Netherlands than in Germany. This may be due to legislative, competitive or administrative reasons, but differences can be noted between information availability between countries. When one designs a study, it should be taken into account whether enough information can be gathered about these industries in the sense that they can be compared with other industries. In the thesis, it was found that aerospace industries were dispersed and little developed in for instance the Netherlands and more developed in France. In this case, it may be recommended that specific industries that are difficult to compare with others or highly specialized should be assessed in a different study that focusses on that industry as it seems that cross comparison leads to invalid results. Also, there are suggestions for future research that derive from the study. A recommendation is concerned with the further assessment of the R&D efficiency differences between on the one hand the pharmaceutical and electronics industry, on the other hand the computer, aerospace and scientific instruments industry. In addition, one may asses firms that have patent output but no R&D development budgets, which on average may play a significant part in explaining how R&D efficiency works and further elaborates on the relationship between R&D investments and patent output. Subsequently, the negative relationship of the scientific instruments industry between patents granted and firm performance may be interesting to assess. Finally, as is concluded at the end of chapter 4 and is further discussed in the conclusion, it may seem that the relationship between R&D efficiency, patents granted and firm performance is strongly dependent on the industry that is observed and likely its underlying characteristics. Therefore it may be more beneficial to carry out limited comparison or singular industry studies when assessing the relationship between R&D efficiency, patents granted or firm performance.

6 Conclusion

This study aimed to achieve insight in the European high-tech industries sectors efficiency rate, how patents granted are funded and their subsequent influence on firm performance. By testing three hypothesis, the issue at hand has been assessed and clarified, and results are presented in chapter four. In this conclusive chapter, the results of the study are summarized, remarks are made about the validity and accuracy of the study and the research question is answered. In addition, the main recommendations for future studies are summarized.

The first hypothesis aimed to assess the relationship between research and development expenditures and patents granted. The first hypothesis has been assessed by conducting a simplified DEA analyses per European high-tech industry to analyses the R&D efficiency of the industries. The efficiency rates on average over the years 2010-2018 have been compared by using an two-sample t-test and subsequent cross-comparison. The results show that there is a difference in efficiency rate between two groups. The pharmaceutical and electronics industry showed significant differences with the aerospace, scientific instruments and computer industries. There thus is a significant difference between the different European high-tech industries efficiency, but important to note is that there thus seems to be possible underlying characteristics that cause this difference. More important for the central research question is that all industries are different and that R&D efficiency is different In some industries when compared to others. This shows that it does matter in which industry one should invest when they expect that R&D expenditures will lead to patents granted.

The second half of this study was focused on the relationship between innovation and firm performance. I have tested whether patents granted is a good predictor for improvements in firm performance. In addition, this relationship is also assessed by using the R&D efficiency rate that was calculated for hypothesis 1 as independent variable. By conducting panel data research, an panel data regression was employed to assess what the relationship is between both dependent variables and firm performance, which is operationalized as operating revenue. The second hypothesis assesses the relationship between R&D efficiency and firm performance. The hypothesis aimed to gain insight whether the degree of R&D efficiency is a good predictor of firm performance. To answer the hypothesis, a panel data regression was carried out and showed that R&D efficiency only had a significant influence in the aerospace industry. In addition, the R2 value, which argues how much of variation is explained by the independent variable, was very low. One may thus conclude that on average there is no relationship between R&D efficiency rate and firm performance. However, as one significant relation was found in one of the industries, one should disclaim that while H1 is rejected, and the null hypothesis kept, that it is possible that there is a relationship between R&D efficiency and patents granted but that that relationship is strongly dependent on the industry and likely underlying characteristics.

The third hypothesis aimed to create insight in the relationship between patents granted and firm performance to assess how strong the relationship is. The relevance of the hypothesis is to measure whether patents granted are a valid indicator of a to be expected higher firm performance. In addition, one can measure if R&D efficiency or patents granted is a better indicator for firm performance. An panel data regression has been conducted and the results showed that in two of the five industries there was a significant relationship between patents granted and firm performance. The hypothesis underpins the claim that was made about the first hypothesis, meaning that there might be some difference in the importance of patenting in different industries, which is an interesting finding for future research. In addition, the negative relationship between patenting and firm performance in the scientific instruments sector may be a ground for future research as it is the only negative coefficient that was found in the H3 results.

The research question that was operationalized in the introduction of this study was '' What is the relationship between R&D expenditures, patents granted and firm performance in European high-tech industries?". By assessing the results from the three hypothesis, the answer may be simplified to: it depends. There are grounds in this thesis on which one may declare that R&D efficiency has in one case a significant effect on firm performance. Patents granted in turn has in two out of five cases a significant effect on firm performance. There thus is a case to be made that depending on the industry and independent variable, there is a possible relationship between R&D expenditures, that that in some degree is converted to patents granted (however, this process may be different in different industries and even firms) and that patents granted in some cases has an influence on firm performance. However, this research also stresses the notion that this relationship is strongly dependent on which industry one assesses. I would thus on the basis of this research conclude that in very little cases there is a positive relationship of R&D efficiency and patents granted on firm performance. However, this study thus shows that there are 3 implications to comparative studies like this one;

1. General R&D efficiency are likely to find that there are industry specific factors at play, which means that general studies like this study that compares industries are likely to only point out that conclusion. In other words, it is difficult to measure anything else of meaning other than that there are underlying reasons why industries are different from each other and therefore no real meaningful conclusions can be drawn from research like this. The results of the study states that there is on average some positive effects on average, but it also implies that for any meaningful conclusion, one should focus on separate or limited comparative industry studies to assess what really influences R&D efficiency in that industry.

2. It is difficult to properly account for cases where there are no input, but there is output and the other way around. When you keep them in your study, you will experience difficulties in traditional R&D analysis tools like DEA as these aren't designed to account for this. However, it is important to keep these variables in your selection when possible as these groups are too large to withdraw completely from the analysis. One could make the choice to exclude them from the study, but the outcome of the study would be flawed as it doesn't tell the full story of R&D efficiency in the sector, therefore one could not rely on them for investment decisions or as a calculation tool. However, if you keep them in for your analysis, you will find that when comparing a variable to R&D efficiency the results will be flawed.

3. There may be a cultural or legislative variable that isn't accounted for that does explain the amount of firms that don't have R&D costs but do generate patents. If the conclusion of this thesis is to be tested, one may choose to replicate this study in other parts of the world to determine whether there is an increase or decrease in ''0'' and ''1'' efficiency values when compared to western Europe. If there is a difference with the results in other regions, it would stress the importance of specific industry studies (most desirable if the study is conducted in one country to exclude cross-country effects) for further R&D efficiency studies.

There are some implications regarding this thesis. One of the main implications is the selection of industries. For the selection of industries, I focused on an report by Eurostat (2017) which described what the most important high-tech industries in Europe were. However, the translation table which was needed to determine industry classifications (the domain which describes what industries do and don't have to be included) had to be translated and may therefore lead to translation issues. Also, in retrospect, the distribution of high-tech activities should be assessed before the investigation is conducted. This thesis does include the 3 largest economies in Europe, but for instance the aerospace industry is represented in only a few countries, but where it is present, it has a very strong and influential presence. Such an industry could have been better assessed in a separate study. Another implication of the study is that it wasn't initially designed to go about the situations where there is input, but no output and the other way around. In a future study, one may find a way to make a better design choice in which these situations can be taken into account to properly determine R&D efficiency. Lastly, an implication of the study is data collection and gathering R&D expenditures related data. The database that was used only offered a restricted selection of data, which is according to Enache & Srivastava (2018) not an uncommon finding. They suggest that additional R&D expenditure data may be extracted from the SG&A account, but that the process takes a lot of time and is difficult to do. For the scope of this study, extracting the data is the proposed manner would be too time intensive, but it would be recommended for future research.

This thesis is concluded by pointing towards possible future research that may help to elaborate on the findings in this thesis and explain some of its points of interest. One possible direction for future research is to investigate the differences between groups in this study, for instance between the pharmaceutical and electronics industries and the computer, aerospace and scientific instruments industries. Another possible direction for future research is to assess why certain firms don't have R&D input but have patents granted output. If the data is assessed, it may also be industry specific. By investigating the reason for this, R&D efficiency studies can account for them.

7 References

- Acs, Z. Sanders, M. (2012). Patents, knowledge spillovers, and entrepreneurship. Small Business Economics, 39, 801-817.

- Agostini, L. Caviggioli, F. Flippini, R. Nosella, A. (2015). Does patenting influence SME sales performance? A quantity and quality analysis of patents in Northern Italy, European Journal of Innovation Management, 18(2), 238-257.

- Albert, H. Png, I. (2013). Patent rights and economic growth: evidence from cross-country panels of manufacturing industries, Oxford Economic Papers, 65(3), 675-698.

- Allred, B. Park, W. (2007). The influence of patent protection on firm innovation investment in manufacturing industries, Journal of International Management, 13(2), 91-109.

- Ambrammal, S. Sharma, R. (2016). Impact of patenting on firms performance: an empirical investigation based on manufacturing firms in India. *Economics of Innovation and New Technology, 25(1),* 14-32.

- Arora, A. Ceccagnoli, M. Cohen, W. (2008). R&D and the patent premium, International Journal of Industrial Organization. 26(5), 1153-1179.

- Barton, J. Emanuel E. (2005). The patents-based pharmaceutical development process: rationale, problems, and potential reforms, JAMA, 294(16), 2075-2082.

- Baum, C. (2001). Residual diagnostics for cross-section time series regression models. *The Stata Journal, 1,* 101-104

- Bessler,W. Bittelmeyer, C. (2008). Patents and the performance of technology firms: Evidence from initial public offerings in germany.

- Bloom, N. van Reenen, J. (2002). Patents, real options and firm performance, The Economic Journal, 112(478), C97-C116).

- Casson, R. Franzco, P. (2014). Understanding and checking the assumptions of linear regression: a primer for medical researchers. *Clinical & Experimental Ophthalmology, 42(6),* 590-596.

- Chabchoub N. Niosi, J. (2005), Explaining the propensity to patent computer software, Technovation, 25 (9), 971-978

- Chen, T-C. Guo, D-Q. Chen, H-M. Wei, T-T. (2019), Effects of R&D intensity on firm performance in Taiwan’s semiconductor industry, Economic Research, 32, 2377-2392.

- Chen, Z. Zhang, J. (2019). Types of patents and driving forces behind the patent growth in China. Economic Modelling, 80, 294 – 302.

- Chua, W. (1986). Radical Developments in Accounting Thought. The Accounting review ,61(4) , 601-632.

- Ciftci, M. Zhou, N. (2016). Capitalizing R&D Expenses versus disclosing intangible information, Review of Quantitative Finance and Accounting. 46, 661-689.

- Cloodt, M. Hagedoorn, J. Roijakkers, N. (2010), Inter-firm R&D networks in the global software industry: an overview of major trends and patterns, Business History, 52(1), 120-149.

- Cockburn, L. Macgarvie, M. Müller, E. (2010). Patent thickets, licensing and innovative performance, Industrial and Corporate Change, 19 (3), 899-925.

- Cohen, W. Nelson, R. Walsh, J. (2000) Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not), NBER Working paper no.7552. National bureau of Economic research.

- Cullman, A. Schmidt- Ehmcke, J. and Zloczysti, P. (2012). R&D efficiency and barriers to entry: a two stage semi parametric DEA approach. Oxford Economic Papers, 61(1), 176-196.

- Czarnitski, D. Kraft, K. (2010). On the profitability of innovative assets, Applied Economics, 42(15), 1941-1953.

- 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.

- Drukker, D. (2003). Testing for Serial Correlation in Linear Panel-data Models. *The Stata Journal, 3(2),*168-177

- Enache, L. Srivastava, A. (2018) Should Intangible Investments Be Reported Separately or Commingled with Operating Expenses? New Evidence, Management Science, 64(7), 2973-3468

- Eurostat (2017). High-tech statistics – economic data (statistics explained). https://ec.europa.eu/eurostat/statistics-explained/pdfscache/46748.pdf. Visited 7-4-2020.

- Eurostat (2017). Aggregation of products by SITC Rev. 4. – Annex 5 – High tech aggregation by SITC Rev. 4. https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an5.pdf

- Eurostat (2019). GDP numbers 2019 https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table?lang=en

- Eurostat (2020). Glossary: High-tech classification of manufacturing industries (statistics explained). https://ec.europa.eu/eurostat/statistics-explained/pdfscache/6384.pdf. Visited 06-04-2020.

- Eurostat (2020). Statistics Division- Classifications on economic statistic https://unstats.un.org/unsd/classifications/Econ#Correspondences (correspondence tables)

- Eurostat (2020). Index of correspondence tables -Translation tables EU https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=EN&IntCurrentPage=15

- Eurostat (2020). Correspondence tables - SITC rev.3 to ISIC rev.3 3https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_LINK&StrNomRelCode=SITC%20REV.%203%20-%20ISIC%20REV.%203&StrLanguageCode=EN&StrOrder=1&CboSourceNomElt=&CboTargetNomElt=&IntCurrentPage=76

- Eurostat (2020). Correspondence tables - ISIC rev 3 to NACE rev 1.1. https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_LINK&StrNomRelCode=ISIC%20REV.%203%20-%20NACE%20REV.%201&StrLanguageCode=EN&StrOrder=1&CboSourceNomElt=&CboTargetNomElt=&IntCurrentPage=5

- Eurostat (2020). Correspondence tables - NACE rev 1.1. to NACE rev 2 https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_LINK&StrNomRelCode=NACE%20REV.%201.1%20-%20NACE%20REV.%202&StrLanguageCode=EN&StrOrder=1&CboSourceNomElt=&CboTargetNomElt=&IntCurrentPage=6

- Eurostat (2020). Aggregation of products by SITC Rev.3 https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an4.pdf

- Eurostat (2020). Aggregations of manufacturing based on NACE Rev.2 https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf

- Eurostat (2020). Aggregations of manufacturing based on NACE Rev.1.1 https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an2.pdf

- Eurostat (2020). High-tech industry and knowledge-intensive services (htec) https://ec.europa.eu/eurostat/cache/metadata/en/htec_esms.htm

- Eurostat (2020). GDP numbers 2020 https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=nama_10_gdp&lang=en

- Frietsch, R. Neuhäusler, P. Jung, T. van Looy, B. (2014), patent indicators for macroeconomic growth – the value of patents estimated by export volume, Technovation, 34(9), 546-558.

- Ghapar, F. Brooks, R. Smyth, R. (2014). The impact of patenting activity on the financial performance of Malaysian firms, Journal of the Asia Pacific Economy, 19 (3), 445-463.

- González, E. Gascón, F. (2004). Sources of productivity growth in the Spanish pharmaceutical industry. (1994-2000), Research Policy, 33(5), 735-745.

- Grinols, E. Henderson, J. (2007). Replace pharmaceuticals patents now. PharmaEconomics, 25, 355-363.

- Guellec, D. Martinez, C. Zuniga, P. (2012). Pre-emptive patenting: securing market exclusion and freedom of operation, Economics of Innovation and New Technology, 21, 1-29.

- Grant, K. Matousek, R. Meyer, M. Tzeremes, N. (2019). Research and development spending and technical efficiency: evidence from biotechnology and pharmaceutical sector, International Journal of Production Research, 1-15.

- Hall, B. (2007). Patents and patent policy, Oxford Review of Economic Policy, 23 (4), 568-587.

- Harison, E. Cowan R. (2004). On substitution of intellectual property and free disclosure: an analysis of R&D strategies in software technologies, Economics of Innovation and New Technology. 13 (5) 477-487.

- Hasan, I. Tucci, C. (2010). The innovation – economic growth nexus: Global evidence, Research policy, 39(10), 1264 – 1276.

- Hashimoto, A. Haneda, S. (2008). Measuring the change in R&D efficiency of the Japanese pharmaceutical industry. Research Policy, 37(10), 1829-1836.

- Holgersson, M. Granstrand, O. (2017). Patenting motives, technology strategies, and open innovation, Management Decision 55(6), 1265-1284.

- Hu, J-L. Yang, C-H. Chen, C-P. (2014). R&D efficiency and the national innovation system: an international comparison using the distance function approach. Bulletin of Economic Research, 66(1), 55-71.

- Huang, C-H. Hou, T. (2019). Innovation, research and development, and firm profitability in Taiwan: Causality and determinant,. International Review of Economics & Finance, 59, 385-394.

- Jiménez-Jiménez, D. Sanz-Valle, R. (2011). Innovation, Organizational Learning, and performance, Journal of Business Research, 64(4), 408-417.

- Johansson, B. Lööf, H. Savin, M. (2015). European R&D efficiency, Economics of Innovation and New Technology, 24 (1-2), 140 – 158.

- Karadayi, M. Ekinci, Y. (2019). Evaluating R&D performance of EU countries using categorical DEA, Technology Analysis & Strategic Management, 31(2), 227-238.

- Lee, H-Y. Park, Y-T. (2011). An international comparison of R&D efficiency: DEA approach. Asian Journal of Technology Innovation, 13(2), 207-222.

- Leitner, K-H. Schaffhauser-Linzatti, M. Stowasser, R. Wagner, K. (2005) '' Data Envelopment Analysis as method for evaluating intellectual capital''. Journal of intellectual capital, 6(4), 528-543.

- Lindman, A. P, Söderholm (2016). Wind Energy and green economy in Europe: Measuring policy-induced innovation using patent data, Applied Energy, 179, 1351-1359.

- Mahajan, V. Nauriyal, D. Singh, P. (2018). Efficiency and Its Determinants: Panel Data Evidence from the Indian Pharmaceutical Industry. Margin: the Journal of Applied Economic Research. 12(1), 19-40.

- Maradana, R. Pradhan, R. Dash, S. Gaurav, K. Jayakumar, M. Chatterjee, D. (2017). Does innovation promote economic growth? Evidence from European Countries. Journal of Innovation and Entrepreneurship, 6 (1), 1- 23.

- Maradana, R. Pradhan, R. Dash, S. Zaki, D. Gaurav, K. Jayakumar, Sarangi, A. (2019). Innovation and economic growth in European area countries: the Granger causality approach, IIMB Management Review. 31(3), 268-282.

- Noel, M. Schankerman, M. (2013). Strategic Patenting and Software Innovation, Journal of Industrial Economics, 61 (3), 481-520.

- Noh, H. Song, Y. Lee, S. (2016). Identifying emerging core technologies for the future: Case study of patents published by leading telecommunication organizations, Telecommunications Policy, 40 (10-11), 956-970.

- Oliveira, J. Basso, L. Kimura, H. Sobreiro, V. (2018). Innovation and financial performance of companies doing business in Brazil, International Journal of Innovation Studies, 2(4), 153-164.

- Özbuğday,F. Findik, D. Özcan, K. Başçi,, S. (2020). Resource efficiency investments and firm performance: Evidence from European SME's, Journal of Cleaner Production, 252, (to be pressed).

- Panagopoulos, A. Park, I-U. (2018). Patents as negotiation patents: patenting versus secrecy for start-ups. The Economics Journal, 128 (615), 2876 – 2894.

- Park, W. Ginarte, J. (1997). Intellectual property rights and economic growth, Contemporary Economic Policy, 15(3), 51-61.

- Perelman, S. (1995). R&D, Technological Progress and Efficiency Change in Industrial Activities. Review of Income and Wealth, 41(3), 349-366.

- Powell, W. Snellman, K. (2004). The Knowledge Economy, Annual Review of Sociology, 30, 199-220.

- Pradhan, R. Arvin, M. Nair, M. Bennett, S. (2020). The Dynamics among entrepreneurship, innovation, and economic growth in the Eurozone countries. Journal of Policy Modeling, (to be pressed).

- Reenen, J. (2005). The New Economy: Reality and Policy, Fiscal Studies, 22(3), 307-336.

- Rousseau S. Rousseau, R. (1997). Data analysis as a tool for constructing scientometric indicators, Scientometrics, 40, 45-60.

- Rousseau, S. Rousseau, R. (1998) The scientific wealth of European nations: taking effectiveness into account, Scientometrics, 42, 75-87.

- Seitz, M. Watzinger, M. (2017). Contract enforcement and R&D investment, Research Policy, 46(1), 182-195.

- Sharma, S. Thomas, V. (2008) '' Inter-country R&D efficiency analysis: an application of data envelopment analysis'', Scientometrics, 76, 483-501.

- Shieh, G. (2010). On the misconception of multicollinearity in Detection of Moderating effects: Multicollinearity is not always detrimental. *Multivariate Behavioural Research*, *45(3),* 483-507.

- Slivko, O. Theilen, B. (2014). Innovation or imitation? The effect of spillovers and competitive pressure on firms R&D strategy choice, Journal of Economics, 112(3), 253-282.

- Thomas, V. Sharma, S. Sudhir J. (2011). Using patents and publications to assess R&D efficiency in the states of the USA. World patent information, 33(1), 4-10.

- Thornhill, S. (2006). Knowledge, innovation and firm performance in high- and low-technology regimes, Journal of Business Venturing, 21(5), 687-703.

- Useche, D. (2014). Are Patents signals For the IPO market? An EU-US comparison for the software industry.

- Vismara, S. (2013). Patents, R&D investments and post-IPO strategies, Review of managerial Science, 8(3), 419-435.

- Vo, D. (2019), Patents and early-stage financing: matching versus signaling, Journal of small business management, 57(4), 1252-1279.

- Wang, E. (2007), R&D efficiency and economic performance: A cross-country analysis using the stochastic frontier approach. Journal of policy modeling, 29(2), 345-360.

- Wang, E. Huang, W. (2007). Relative efficiency of R&D activities: a cross-country study accounting for environmental factors in the DEA approach, Research Policy, 36, 260-273.

- Wang, Y. Pan, J-F. Yi, B-W. Yang, G-L. (2020). Assessing the technological innovation efficiency of China’s high tech industries with a two-stage network DEA approach. Socio-Economic Planning Sciences, (to be pressed).

- Wang, M. Wong, M. (2012). International R&D Transfer and Technical Efficiency: Evidence from Panel Study Using Stochastic Frontier Analysis. World Development, 40(10), 1982-1998.

- Woolridge, J. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT press.

- Wu, H. Chen, I. Chen, J. Chien, C. (2019). The R&D efficiency of the Taiwanese semiconductor industry, Measurement, 137, 203-213

8 Appendices

8.1 Robustness test

Robustness test industries. R&D numbers are the numbers including corresponding available patent numbers								
Germany	Industry	Firms in industry that received patents granted 2010-2018	R&D2010	% of values known	R&D 2011	% of values known	R&D 2012	% of values kno
	Pharmaceutical	856	37	4%	37	4%	35	
	Electronics	1585	48	3%	51	3%	50	
	computer	1023	12	1%	12	1%	12	
	Aerospace	1410	32	2%	32	2%	30	
	Scientific instruments	1014	43	4%	47	5%	47	
Netherlands	Industry	Firms in industry that received patents granted 2010-2018	R&D2010	% of values known	R&D 2011	% of values known	R&D 2012	% of values kno
	Pharmaceutical	166	7	4%	6	4%	6	
	Electronics	147	7	5%	8	5%	8	
	computer	336	2	1%	3	1%	3	
	Aerospace	129	2	2%	2	2%	2	
	Scientific instruments	101	6	6%	6	6%	6	
France	Industry	Firms in industry that received patents granted 2010-2018	R&D2010	% of values known	R&D 2011	% of values known	R&D 2012	% of values kno
	Pharmaceutical	330	25	8%	28	8%	29	
	Electronics	405	22	5%	21	5%	25	
	computer	456	24	5%	22	5%	19	
	Aerospace	329	8	2%	9	3%	9	
	Scientific instruments	236	24	10%	23	10%	24	
UK	Industry	Firms in industry that received patents granted 2010-2018	R&D2010	% of values known	R&D 2011	% of values known	R&D 2012	% of values kno
	Pharmaceutical	521	80	15%	90	17%	95	
	Electronics	812	116	14%	128	16%	128	
	computer	1246	63	5%	63	5%	75	
	Aerospace	617	98	16%	93	15%	101	
	Scientific instruments	971	121	12%	138	14%	140	

Robustness test (continued)

R&D2014	% of values known	R&D2015	% of values known	R&D2016	% of values known	R&D2017	% of values known	R&D2018	% of values known
36	4%	36	4%	34	4%	37	4%	36	4%
49	3%	45	3%	45	3%	44	3%	42	3%
12	1%	10	1%	12	1%	12	1%	12	1%
30	2%	31	2%	26	2%	23	2%	24	2%
46	5%	43	4%	42	4%	38	4%	32	3%
R&D2014	% of values known	R&D2015	% of values known	R&D2016	% of values known	R&D2017	% of values known	R&D2018	% of values known
7	4%	7	4%	7	4%	7	4%	6	4%
9	6%	9	6%	10	7%	9	6%	10	7%
3	1%	3	1%	3	1%	3	1%	3	1%
2	2%	2	2%	3	2%	3	2%	3	2%
6	6%	6	6%	6	6%	7	7%	6	6%
R&D2014	% of values known	R&D2015	% of values known	R&D2016	% of values known	R&D2017	% of values known	R&D2018	% of values known
39	12%	42	13%	46	14%	46	14%	47	14%
26	6%	28	7%	28	7%	27	7%	27	7%
20	4%	21	5%	20	4%	22	5%	22	5%
9	3%	9	3%	9	3%	10	3%	9	3%
24	10%	136	58%	132	56%	25	11%	25	11%
R&D2014	% of values known	R&D2015	% of values known	R&D2016	% of values known	R&D2017	% of values known	R&D2018	% of values known
81	16%	71	14%	103	20%	115	22%	110	21%
88	11%	52	6%	105	13%	129	16%	122	15%
54	4%	39	3%	72	6%	79	6%	79	6%
75	12%	55	9%	99	16%	103	17%	115	19%
108	11%	61	6%	131	13%	153	16%	161	17%