MASTER THESIS

The Impact Of R&D Investments On Firm Performance For European Listed Firms

Name: Gijs Reysoo
Student number: S2417138
E-mail: g.reysoo@student.utwente.nl
Study: MSc Business Administration
Faculty: Behavioural, Management and Social Sciences
Track: Financial Management
1st Supervisor: Dr. X. Huang
2nd Supervisor: Prof. Dr. M. R. Kabir
Date: 07-07-2021

UNIVERSITY OF TWENTE.

Abstract

This paper investigates the impact of R&D investments on firm performance. There is still much ambiguity about this impact due to varying results among existing literature. Ordinary least squares (OLS) regression is applied to a sample of 472 European listed companies over the period 2011-2019. This revealed a clear inverted U-shaped relationship between R&D investment and firm performance. A non-linear relationship, which indicates that it is profitable to invest in R&D up to approximately 1.60%. Additionally, it is also found that there is a negative relationship between R&D and high R&D firms (R&D intensity >1%) and a positive relationship between R&D and low R&D firms (RD intensity <1%). It is also observed that a larger firm size enhances firm performance while a highly leveraged company is strongly restricted in its R&D opportunities.

Keywords: *R&D investment, firm performance, European firms, lagged effect, OLS regression*

Table of Content

1. Introduction

For a business, large or small, it is essential to continuously stand out from the competition. Especially, in competitive business environments that are changing rapidly. A company can distinguish itself from its competitors in several ways. For example, by providing the highest quality, the lowest price, or by offering the best service. Innovation and applying the latest technologies can also be an important tool to stay ahead of the competition. Technological innovation is seen as one of the most important forces that drive the economic growth of a company (Guo et al., 2018). This is often reflected in the R&D investments of businesses. Investing in the Research & Development (R&D) of the products or services is a crucial way to gain and maintain a competitive advantage (Bettis & Hitt, 1995). Organizations invest in R&D to create new products, improve their existing products or services, or optimize their processes. Additionally, the knowledge that is gained through R&D also leads to more collaborations with other enterprises, institutions, and universities (Chen et al., 2016). These investments will provide the competitive advantage that is needed to succeed in the ever-changing market (Lee, 2020). Innovation and R&D investments are seen as crucial factors to achieve long-term success (Lin et al., 2006). Teece (2007) describes the value of innovation as follows: "Success requires the creation of new products and processes and the implementation of new organizational forms and business models" (p. 1346).

According to Huang & Liu (2005), science- and technology-advanced countries invest heavily in R&D projects. The OECD supports this statement with actual data. The total R&D expenditures of OECD countries, mostly seen as the most developed countries in the world, have increased significantly over the past 40 years. From 378.000 million dollars in 1981 to over 1.371.000 million dollars of total R&D expenditures in 2018. The emerging Asian countries such as Singapore, China, and Japan have also shown clear increases since the 1990s (OECD, 2020). This demonstrates that the importance of R&D investment is recognized all over the world. Companies overwhelmingly understand the value of innovation.

The increased interest in R&D is also reflected in the existing literature. It is expected that R&D investments always lead to better firm performance, increased market value, or higher sales. However, it should also be considered that R&D investments are expensive. It can take years before these investments are recouped (Hartmann et al., 2006). Investing in new knowledge is likely to be beneficial in the long run but can be damaging in the short term (Vithessonthi & Racela, 2016). The existing literature, therefore, presents a variety of results and supporting arguments. In general, a positive relationship is found between R&D investments and firm performance. According to Lee (2020), R&D investments have a positive time-lag effect on the market value of a company. This implies that after a certain period, the costs that are made are covered by the benefits of the investment. In addition, research by Falk (2012) found that sales grow because of the competitive advantage that is gained through R&D investments. This competitive advantage can be created by using R&D investments to develop rare, valuable, and heterogeneous resources. As a result, R&D investments will be cost-effective and lucrative (Jaisinghani, 2016). However, there are many other results that also depend on the circumstances of the study. For example, Chen et al. (2019) found a negative relationship between R&D investments and current business performance, but a positive relationship between R&D investments and future firm performance. Therefore, implementing a lag can be a crucial factor. Vithessonthi & Racela (2016) concluded a negative effect of R&D investments on high R&D firms and a positive effect on low R&D firms. A conclusion that suggests that investing too much money in R&D is not profitable for a company. This is supported by Yeh et al. (2010), who state that R&D is a costly activity that does not guarantee its potential earnings. Increasing R&D is not always advisable because there must be a cut-off point somewhere, after which it does not generate a proportional return (Hartmann et al., 2006). Yeh et al. (2010) found a non-linear relationship, which included a threshold. This threshold is a turning point to where it is on average profitable to invest in R&D. This inverted U-shaped relationship is more frequently found during studies regarding R&D (Booltink & Saka-Helmhout, 2018; Guo et al., 2018). All these varying results together made this an interesting and relevant research topic to examine in more detail.

By using both the linear term of R&D intensity and the squared term of R&D intensity, it is possible to make a very precise estimate of the curve of the relationship between R&D investments and firm performance. Most existing studies do not use the squared term and only examine whether there is a positive or negative relationship. In addition, a comprehensive sample consisting of 13 different European countries and a wide range of industries is selected. This full sample is also divided into a low R&D-intensive subsample and a high R&D-intensive subsample to create even more insights. The use of both lagged and non-lagged dependent variables will also contribute to this. Existing literature has frequently shown that the use of a lag can change the results since it takes time for an R&D investment to be reflected in the firm performance. All samples are examined by using an OLS regression. The same OLS regression is applied to different robustness tests with other independent and dependent variables. The time period of the collected data is from 2011 to 2019. Altogether, this very comprehensive study of the relationship between R&D investments and firm performance will provide much insight and will extend the existing literature. Altogether, this leads to the following research question:

"To what extent do R&D investments influence the firm performance of European listed companies?"

This report is structured in the following order. The next chapter elaborates on the main theories regarding R&D and firm performance. It also examines the existing literature and analyzes these studies. Finally, based on these theories and literature, hypotheses are formulated which can be found at the end of this chapter. In chapter 3, the most prominent methods and variables in the current literature are explored and discussed. This is followed by an explanation of the method and variables that are applied in this paper. Chapter 4 will explain briefly about the collection and use of the data. Chapter 5 contains the main results including explanatory notes. In the final chapter, the findings are summarized in the conclusion and debated in the discussion.

2. Literature Review

2.1 Introduction to R&D

There are internal and external incentives for companies to keep innovating and investing in R&D. The external motivation is related to competitiveness and customer demand. To continue delivering value to the customer, it is crucial to fulfilling the need of the customer. Companies must keep innovating to achieve this. Innovation activities can be considered as an essential task to stay competitive and profitable (Vithessonthi & Racela, 2016). Competitive advantages can be gained when this happens faster and better than the competitors. Investing in R&D can also optimize the processes. This can result in a faster response to customer demand, allowing companies to gain a competitive advantage. In addition, a process that is set up more efficiently reduces costs (Lee, 2020). There are also tax incentives to invest in R&D, which can be seen as an internal motivation to invest in R&D. Governments stimulate investments in R&D by granting subsidies or implementing an appropriate tax policy. Companies that invest in R&D can deduct more from their taxable income and therefore pay less tax (Chen & Li, 2018). Research by Dumont (2013) shows a positive link between this government policy and R&D investments among Belgian companies. Bloom et al. (2002) studied the same relationship but under several large OECD countries such as the United States, the United Kingdom, and Germany. Their conclusion is the same: when tax benefits reduce the costs, the total investment in R&D increases. However, it is necessary to consider that each country has its own conditions and tax incentives. Consequently, the effectiveness of the tax policy varies from country to country (Li & Du, 2016).

The total amount of R&D expenditures and the choice to invest in the differentiation of the products or optimization of the processes also depends on the strategy of the company. Guo et al. (2018) argue that R&D spending is higher in companies with a product differentiation strategy. Because of the continuing need for innovation, companies must keep up with competitors and need to keep investing in R&D. It is considered as product-R&D, which requires continuous adaptation to the changing market demand. Firms with process-R&D aim to maintain quality and often adopt a cost leadership strategy (Liao & Cheung, 2002). Companies with a cost leadership strategy focus on the lowest price. This can be achieved by working as efficiently as possible and by using R&D to optimize the processes. This strategy results in lower total R&D investments. This is supported by the findings that R&D is positively related to performance in the case of a product differentiation strategy. When testing the same relationship but for a cost leadership strategy, this changes to an inverted U-shape relationship (Guo et al., 2018). Chung & Choi (2017) confirm this through an investigation among Korean firms. The effect of R&D on the growth of companies with a product differentiation strategy is more robust than on companies with a cost leadership strategy. These results suggest that R&D plays a more significant role for companies with a product differentiation strategy.

However, R&D investments also have a different side and carry risks. The uncertainty of investing in R&D is higher than investing in tangible assets (Gharbi et al., 2014). There are several reasons for this. Materializing an R&D investment in revenue-generating sources is often time-consuming. Technology changes rapidly and companies could already be too late when a project is finished (Beladi et al., 2021). It should also be taken into account that a new or upgraded product does not always become a successful one. There are large differences among companies in terms of the probability of commercialization and the probability of commercial success (Mansfield & Wagner, 1975). Depending on the size of the investment, the failure or disappointment of an investment can be costly for a business. This can be a problem for more financially constrained companies because there is a clear positive relationship between R&D intensity and distress risk (Zhang, 2015). Companies that invest more in R&D face a higher risk of defaulting on their financial obligations. Therefore, companies should always consider whether and how much they want to invest in R&D projects.

Another important point is the increasing information asymmetry between managers of R&D-intensive companies and their outsiders. Gharbi et al. (2014) suggest that this is due to companies disclosing little or no information about their projects. R&D creates new 'uniqueness' and this is their competitive advantage that they do not want to share with outsiders. Also, the value of R&D projects is not always shared on the balance sheet, so it remains guessing. A higher information asymmetry leads to difficulties in funding R&D projects in the future. Investors are reluctant to invest in projects if they know little or nothing about them (Aboody & Lev, 2000). The last point, why organizations need to think carefully about investing in R&D concerns return volatility. According to Chan et al. (2001), there is a clear relationship between R&D intensity and return volatility. The lack of accounting information, not placing it on the balance sheet, is again cited as an argument for this finding. Gharbi et al. (2014) came to the same conclusion and indicated that a firm must reduce its information asymmetry.

2.2 Theories that influence R&D

2.2.1 Resource-based theory

One of the most common theories related to R&D intensity and firm performance is the resource-based theory. Usually, it is called the resource-based view (RBV) and refers to the pivotal work of Barney (1991). Resources are important strengths or weaknesses of enterprises, depending on how they are deployed. Mostly, the focus is on the product-market perspective, what product is in demand or shortage on the market, and the resources are adjusted accordingly. The resource-based view is viewed from the perspective of the resources, and an optimal product-market activity can be created based on the available resources (Wernerfelt, 1984).

The work of Barney (1991) explores how and when resources can provide a sustained competitive advantage. A competitive advantage is sustained when it is nearly unfeasible for the competitors to duplicate a strategy. Resources can contribute to this process, but they should have four attributes. The resources are supposed to be valuable, rare, imperfectly imitable, and not substitutable. In this way, resources are as heterogeneous and immobile as possible, and a sustained competitive advantage can be achieved (Barney, 1991). Heterogeneity is the foundation and necessary for competitive advantage (Peteraf, 1993). Resources that are easy to buy or replicate by competitors will not lead to an economic benefit (Grant, 1991).

R&D investments are also a type of resource. This establishes the connection between R&D investments and the resource-based theory. R&D investments are qualified as intangible resources (Diefenbach, 2006). Diefenbach (2006) defines an intangible resource as “everything of immaterial existence used or potentially usable for whatever purpose that is renewable after use and decreases, remains or increases in quantity and/or quality while being used” (p. 410). Nowadays, intangible assets are seen as an essential source of growth and differentiation (Montresor & Vezzani, 2016). Intangible resources are more likely to generate a competitive advantage than tangible resources because they are rarer and socially complex, making them harder to duplicate and change (Hitt et al., 2001; Galbreath, 2005). As mentioned earlier, rarer resources contribute to a competitive advantage (Grant, 1991). R&D investments include expenditures on human capital, research, and efficiency on different levels and therefore creates knowledge. Knowledge creation offers investment opportunities in the short term and higher incomes and productivity in the long term (Van Ark et al., 2009). This is confirmed through an empirical study by Seo & Kim (2020). They analyzed the relationship between 3 forms of intangible resource investments (human capital, advertising, R&D) and their impact on profitability and firm value. In all three examples, the relationship appeared to be positive and significant, which clearly shows that investing in intangible resources pays off.

As mentioned earlier, the resource-based view is one of the most frequent theories related to R&D and firm performance. In most cases, there is a positive result between R&D investments and the firm's performance. According to Sher & Yang (2005), an organization’s innovation capability is one of its most important resources. This enables the ability to create and offer varied and distinctive products that provide a competitive advantage. Their research among Taiwanese semiconductor manufacturers proves this statement. There is a positive and significant linear relationship between innovative capacity and R&D intensity which shows that competitive advantage can be achieved through investing in R&D. This is very much in line with the findings of Jaisinghani (2016), which points out that R&D is a very important resource to develop new and innovative products.

According to Ho et al. (2005), companies need to invest in resources that enhance their core competencies. They strengthen the resource-based view with their conclusion that especially manufacturing firms benefit the most from R&D. This has to do with the fact that manufacturing firms can distinguish themselves by developing innovative products, which can be stimulated by R&D investments. Non-manufacturing or service firms could invest better in marketing and advertising (Ho et al., 2005). Lome et al. (2016) found a positive relationship between R&D and revenue growth before and during recessions. Companies with high R&D can continue to grow during a recession and have a competitive advantage after the downturn due to product development. And thus, according to the resource-based theory, R&D is a unique resource. However, the resource-based view often reveals a clear positive and significant relationship between R&D and companies' performance, but it provides several interesting and relevant arguments.

2.2.2 Knowledge-based theory

The knowledge-based theory is very similar to the resource-based view and is seen as an outgrowth of this theory. In this theory, knowledge is the primary resource of a firm (Grant, 1996b). The accumulation of knowledge is seen as a crucial factor in the long term, while technology can be especially distinctive in the short term (Chen et al., 2016). Through knowledge, companies try to gain a competitive advantage that depends on the ultimately integrated knowledge. Organizations with a high level of knowledge can react faster to environmental changes (Nonaka, 1994). To differentiate from its competitors, a company needs to acquire unique knowledge. According to Nickerson & Zenger (2004), there are two ways to acquire new knowledge. This can be done either by absorbing existing external knowledge or by developing new unique knowledge. The last one often arises after a serious problem has emerged, after which a solution must be found. This generates new knowledge that provides a valuable solution. By developing this unique knowledge, a company can gain a competitive advantage by using knowledge. Especially tacit knowledge, which is more difficult to transfer and integrate and therefore harder to imitate through competitors, can be a tremendous competitive advantage (Grant, 1996b). Cuervo-Cazurra et al. (2018) discovered that R&D sources controlled by the company (insourced and onshore) create more knowledge.

Vithessonthi & Racela (2016) built their research on the knowledge-based theory. They see knowledge as a unique resource and expect that R&D investments increase this knowledge. Something that has been proven by Fey & Birkinshaw (2005), who stated that knowledge could form the basis for superior R&D performance. Investing in knowledge would not pay off in the short term, but it would generate a competitive advantage in the long term. Their results about the negative relationship between R&D and operating performance in the short term but a positive relationship between R&D and firm value confirms this. Vithessonthi and Racela (2016) argue that knowledge leads to competitive advantage, increasing this firm value. Their findings empirically support the knowledge-based theory.

Attracting external knowledge and technology through collaborations can be beneficial for a company as well. By working together, it is possible to create more knowledge and value. Collaborating allows a company to leverage the strengths of its partners (Inkpen, 1996). Wang et al. (2015) concluded in their research, based on the knowledge-based theory, that external knowledge increases innovation capabilities, which in turn leads to competitive advantage and enhanced firm performance. A conclusion also given by Chen et al. (2016). Attracting external knowledge has a strong and significant influence on innovation performance, which positively impacts firm performance. From a knowledge-based theory perspective, knowledge leads to improved products and performance and more robust and better collaborations. The presence of more excellent knowledge makes it more attractive for institutions, other companies, and universities to cooperate. This strengthens the process of attracting new external knowledge, innovation, and improved performance (Chen et al., 2016).

2.2.3 Absorptive capacity theory

A theory that is regularly linked to R&D investments is the absorptive capacity theory. Cohen and Levinthal (1990) are seen as the pioneers of this theory and describe the absorptive capacity as the "ability to recognize the value of new information, assimilate it, and apply it to a commercial end" (p. 128). Developing and maintaining absorptive capacity is crucial for long-term success as it broadens, reinforces, and improves the knowledge of a company (Lane et al., 2006). The level of prior knowledge, consisting of basic skills as well as knowledge of the latest scientific and technological developments, determines the level of absorptive capacity. It is an organizational and individual process. Organizations are alert to acquire new external knowledge to use, but individuals within the organizations need to utilize it (Cohen & Levinthal, 1990). As described in the previous theories, the available resources and knowledge could generate a competitive advantage. The absorptive capacity of the organization and the staff can make a difference in applying the resources. Something dependent on how efficiently a company can acquire, store, process, and integrate knowledge. Or how easily a firm can convert the input into valuable output (Grant, 1996a).

Cohen and Levinthal (1989) concluded that there is a dual role for R&D, which simultaneously establishes a clear link between this theory and the relationship between R&D and firm performance. R&D contributes to the innovation of products and services but also increases the absorptive capacity of an organization. According to these findings, an organization could therefore improve its absorptive capacity by increasing its R&D expenditures. Two factors affect the importance of investing in absorptive capacity through R&D expenditures. This depends on the quantity and difficulty of the knowledge that must be assimilated and utilized. Together with the consideration of industry-level determinants such as technological opportunities, appropriability, and possible spillover effects, a company can weigh up the amounts to be invested in R&D and thereby absorptive capacity (Cohen & Levinthal, 1990). This is important because, in addition to increased innovation and the exploitation of knowledge, it also results in a competitive advantage and higher firm performance (Volberda et al., 2010).

Especially that last point has been tested empirically several times. It is a statement that is confirmed by a recent investigation among Brazilian manufacturing firms. According to Paula & Silva (2018), a firm's absorptive capacity increases its innovation performance. However, the authors examined whether this higher ability to innovate would lead to better financial performance in the future, but the result was negative. They explained that the two-year lag they used was too short to see the financial results of the improved innovation. Griffith et al. (2004) confirmed the statement of Volberda et al. (2010) and examined the relationship between the role of human capital and R&D effectiveness. Based on a sample of 12 OECD countries, it was concluded that human capital could stimulate innovation and absorptive capacity. This in turn increases the effect of R&D on the growth of the total factor productivity (TFP).

Several studies also study the relationship between R&D and firm performance and consider the absorptive capacity as an important factor in the success of R&D investments. Both Lin et al. (2012) and Jaisinghani (2016) find a positive and significant correlation. Lin et al. (2012) consider everything from the absorptive capacity perspective and conclude that R&D intensity has a positive influence on the innovation performance of an organization. They measured innovation performance in the number of co-patents which represents the number of patents they hold together with other corporations. This suggests that entering collaborations and collecting knowledge in this way is crucial. Jaisinghani (2016) briefly mentions the subject and indicates that R&D should enhance the absorptive capacity, ultimately leading to higher returns. The research he conducted in the Indian pharmaceutical industry shows that increasing R&D has a positive impact on return on assets (ROA) and return on sales (ROS), which confirms his expectation.

2.3 Empirical evidence

There is a lot of research available concerning the impact of R&D on the performance of companies. There are many differences in measurement methods, the used variables, the used samples, etc. Because of this, there are also some differences in the results, and there are many different arguments for a given relationship. There is a large amount of diversity in the previous research.

Most of the existing research has found a positive and significant relationship between R&D and firm performance. Agustia et al. (2020) argue that a higher R&D intensity leads to higher sales and more efficient processes. This results in higher revenues and reduced costs, which increases the operating performance of the company. Chen et al. (2019) came to the same conclusion but extends this with the fact that larger companies can start this process faster. Larger companies have more resources and can therefore invest more in R&D and technology. This also applies to companies with better accounting performance. Research by Gui-long et al. (2017) shows that R&D contributes much more to companies with a better firm performance than with lower firm performance. It is suggested that this will be due to more financial resources and reserves from the better-performing companies. Capasso et al. (2015) and Falk (2012) investigated the impact of R&D investments on employment growth. Both found a positive relationship, which means that employment growth increases with a higher R&D intensity. Falk (2012) also investigated the influence of R&D intensity on revenue growth, just like Lome et al. (2016). Both see an evident boost in revenues when the R&D intensity increases.

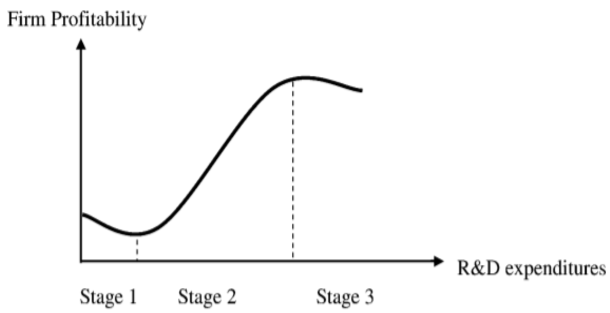
Lee (2020), Seo and Kim (2020), and Vithessonthi and Racela (2016) all find a positive relationship between investing in R&D and the market value of the company. Vithessonthi and Racela (2016) argue that there is a poorer performance in the short term, but that the firm value is higher because the investment opportunities are reflected in the value. The company is valued higher because people know that the R&D investments are likely to lead to positive developments in the future. Another way to investigate the relationship between R&D and firm value is through the value of stocks. Chan et al. (1990) found that increasing R&D expenditures lead to higher stock returns and prices. In other words, investors look at the long term and are not frightened by the costs that must be made.

Based on the literature, infinite investment in R&D may seem attractive on a financial level. However, R&D investments are not always profitable and carry various risks. It can take years before an R&D investment is recouped (Hartmann et al., 2006). There is no guarantee that an R&D project will be a success. The probability of commercialization or commercial success varies greatly and determines the likelihood of success of R&D investments (Mansfield & Wagner, 1975). Consequently, investing in R&D does involve risks. Something that is also reflected in the literature. Studies with a negative relationship between R&D investments and firm performance are in the minority, although they are enough studies with this conclusion, and therefore, they raise questions.

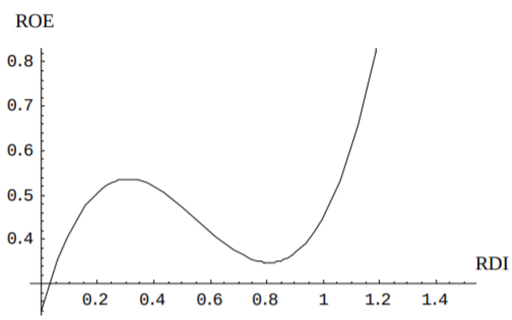
Chen et al. (2019) and Xu et al. (2019) tested the relationship between R&D intensity and financial firm performance and both concluded that there is a lagged effect. Initially, they found a negative relationship between R&D and current financial firm performance. A positive relationship occurs after implementing a lag. Paula & Silva (2018) even concluded that a two-year lag was not enough. The period of 2 years is not sufficient to see the improved innovation reflected in the financial results. A lagged effect is a delay between economic activity and its consequence. The tricky thing about lagged effects is that the effects come later than the earlier action, and this period is difficult to predict (Lee, 2020). It is a subject that always needs to be considered when researching R&D investments and their profitability. Implementing a lag increases the reliability of a study on this topic. It is related to the earlier mentioned risk that an R&D investment will generate revenues later. This moment can not be estimated beforehand. The company will also have to consider the fact that they may be too late with their developments. This can cause distress risk because the investment has probably cost them a lot of money (Zhang, 2005). Vithessonthi & Racela (2016) studied the same issue among listed firms from the United States. They came to a few interesting findings. Both ROA and ROS are negatively and significantly associated with R&D intensity, suggesting an increased amount of R&D leads to lower firm performance. But they also found that this is only true for companies with a high R&D. These companies would take too much risk in the field of R&D which frequently leads to loss-making projects. There is a positive relationship for companies with low R&D, which means that R&D would slightly improve the firm's performance.

In addition to linear positive and linear negative relationships, there are also several types of non-linear relationships discovered in the studies regarding R&D. According to Hartmann et al. (2006), there is a cut-off point after which it is no longer profitable to invest in R&D. Guo et al. (2018) confirmed this and found an inverted U-shaped relationship. Therefore, up to a certain threshold, investing in R&D is profitable and results in higher performance. After this point, these numbers decrease. Guo et al. (2018) argue that this point is around 6% to 8%, which means that the optimal amount of R&D spent is 6% to 8% of total assets. There would be a threshold because it is mainly in the beginning that most benefit is derived from innovation. Over time, the costs take over the benefits, and a threshold value arises. Booltink & Saka-Helmhout (2018) conducted a similar study, mainly focused on non-high-tech SMEs. They also found an inverted U-shape relationship. But the threshold value until which it was profitable to invest money in R&D was higher, namely 9.8%. Yeh et al. (2010) found thresholds for ROA (1.6%), ROE (0.6%), and net profit growth rate (9.8%). Below these thresholds, the coefficients were positive, and above, they change to negative. Therefore, the threshold value is the turning point, and the percentages are the optimal levels of R&D investments. So again, there is an inverted U-shape relationship.

Another curvilinear relationship is the S-curve. With the S-curve, there is also a point where it is no longer profitable to invest in R&D, but the losses are less heavy than with an inverted U-shaped relationship. Yang et al. (2010) talk about a three-stage S-curve. During the introduction of new technology, there is still a loss or a negative slope. After this, there will be a period of growth, a positive slope. Finally, there will be a point of maturity after which the profit will turn into a loss again. It is a point where the product is outdated, or a competitor introduces an improved product. Wang (2009) found another variation on the S-curve, the inverse S-shaped relationship. Again, they made a distinction between 3 different stages, but everything occurred exactly the other way around than the S-curve. R&D has a positive impact on performance on the initial stage, a negative impact during the second stage, and a positive impact during the final stage. Between the last two stages, there is again a threshold in which the advantages outweigh the disadvantages.



S-shaped relationship (Yang et al. 2010)



Inverse S-shaped relationship (Wang, 2009)

2.4 Hypotheses development

The previously mentioned theories all indicate that R&D positively influences firm performance in the longer term. As well the resource-based view as the knowledge-based view indicates that R&D is a unique way for companies to distinguish themselves and create a competitive advantage. Investing in R&D increases the available knowledge, and the knowledge of a company is difficult to duplicate by competitors. The last discussed theory, the absorptive capacity theory, also plays a role in this. It has been shown that a company can distinguish itself when it is able to recognize, absorb and apply knowledge faster than its competitors. This will increase the possibilities for innovation.

Existing empirical evidence also shows that there are mainly positive relationships found between R&D investments and firm performance. The investments, therefore, lead to better financial results in the long term because of improved products or more efficient processes. Some differences may arise due to different industries, the type of knowledge, or the size of companies. Because the theories and most of the empirical evidence demonstrate a positive relationship, the following hypothesis is formulated:

Hypothesis 1: Investments in Research & Development (R&D) have a positive impact on firm performance.

Both linear and non-linear relations were identified during the literature review. Most of the results show a positive and linear relationship, which means that investing in R&D is profitable at any level. Negative linear relationships were also found but to a lesser extent. The benefits do not outweigh the costs, especially in the short term. Furthermore, several interesting studies have led to a curvilinear relationship between R&D and firm performance. Examples are the inverted U-shape and S-curve. In these cases, there are threshold values that can be important when a company or organization is considering investing in R&D. These results indicate that it is financially profitable to invest in R&D up to a certain point. In this case, investing endlessly would therefore lead to losses. Given these results, it is interesting to look closer at this. Despite the curvilinear relationships, most of the empirical evidence is linear, and therefore the following hypothesis is formulated:

Hypothesis 2: The relationship between Research & Development (R&D) and firm performance is linear.

3. Methodology

3.1 Research methods

This section will analyze different methods that can be used to answer the research question. After weighing up the pros and cons, a choice can be made for a specific research method. The different methods are briefly discussed, and the relevant studies that have used the respective methods are examined.

3.1.1 Ordinary Least Squares (OLS) Regression

Ordinary Least Squares (OLS) is one of the most widely used statistical methods during research. OLS is a form of linear regression and attempts to correlate an independent and dependent variable. When the relationship between multiple independent and one dependent variable is tested, there is multiple regression. The advantage of OLS is that it makes the sum of the squared residuals as small as possible. In other words, the difference between the observed and predicted values is minimized. This makes OLS one of the most reliable statistical methods (Veaux et al., 2015). The OLS estimator is the most accurate and unbiased estimator (White & Macdonald, 1980).

However, some assumptions need to be met. If this does not happen, the reliability of the OLS regression can be questioned. The first assumption is related to linearity. There must be a linear relationship between the independent and dependent variables. In the case of, for example, a U-shaped relationship (non-linear), this can be achieved by adding a quadratic predictor variable. The second assumption is related to multicollinearity. In multiple regression, this would be the case if there is a linear relationship between the independent variables. Existing studies often test this using the Variance Inflation Factor (VIF). A VIF < 10 is often seen as an assumption that there is no or low multicollinearity. The third assumption that must be met and what becomes a problem when this does not happen concerns homoscedasticity. In the case of homoscedasticity, the variances of the residuals in the model are constant at all levels. When this is not the case, it is called heteroscedasticity. The model could now make incorrect estimates, and the reliability decreases. By using a scatterplot of the residuals, it is possible to see if the data is homoscedastic. Other statistical methods such as Weighted Least Squares (WLS) or Generalized Linear Model (GLM) should be considered if there is too much heteroscedasticity. The last assumption states that the residuals should be normally distributed. A normal distribution of the residuals ensures reliable results. However, this assumption is more relevant for smaller samples. A Q-Q plot can be used to see if this assumption is met (Williams et al., 2013).

As mentioned before, OLS is one of the most known and used statistical methods. OLS regression has also been widely used to investigate the relationship between R&D and firm performance (Xu & Sim, 2018; Seo & Kim, 2020; Xu et al., 2019; Guo et al., 2018; Vithessonthi & Racela, 2016; Coombs & Bierly, 2006). Unfortunately, there is very little explanation by the authors about their reasons for choosing this method. In most cases, the research is done over a certain period, and they work with a balanced panel dataset which makes it logical to choose for OLS. Gui-long et al. (2017) conducted a Pooled OLS regression because they had an unbalanced dataset. Using an unbalanced dataset would be an addition to the already existing knowledge, and it would give more robust statistical results than a balanced dataset. Furthermore, like most others, they used longitudinal data, so observing a sample over time. This would lead to a solution of bias by unnoticed heterogeneity and lower multicollinearity. As a result, the reliability of the regression estimates would increase. Capasso et al. (2015) also performed a Pooled OLS regression to estimate the average effect of R&D. It is an average effect because OLS uses the conditional mean. They also performed a quantile regression to obtain more insight. This is a well-known method that will be discussed in the following section.

3.1.2 Quantile Regression

Quantile regression is a type of regression that can be used when linear regression requirements are not met. Most regression models estimate the conditional mean. The quantile regression will use an estimated conditional median (Koenker & Hallock, 2001). A sample median is more robust for outlying observations than a sample mean for estimations, especially for contaminated data (Yu et al., 2003). Quantile regressions often use distributions. By plotting this for each group, it is possible to get a much more complete picture than just displaying the mean or median. For example, the total sample can be divided into groups of 4 (quartiles), groups of 5 (quintiles), or even more groups (quantiles or percentiles) (Koenker & Hallock, 2001). The most crucial advantage of quantile regression is that it also enables relationships with deviating data. This is also possible with regressions such as OLS, but by using quantiles, it is easier to understand because medians are used instead of averages. This is especially useful for samples with, for example, non-linear or skewed relationships. Knowledge of the underlying reasons for such relationships is often limited (Koenker & Hallock, 2001). OLS regressions are exceptionally efficient when the random variables are distributed normally, using a Gaussian. A significant advantage and the purpose of the quantile regression is that it remains robust when the distribution is unknown (Capasso et al., 2015). The quantile regression is robust and will remain the same even though the error term is heteroscedastic (Falk, 2012).

An example of a study that used quantile regression as a research method is that of Capasso et al. (2015). They investigated the effect of R&D on firm employment growth. Their reasoning for choosing quantile regression was that they wanted to determine the impact of different levels of R&D on growth would be. For this, quantile regression is an appropriate method. There will be changes in the quantiles when the level of R&D is adjusted. In contrast to a linear regression where only the mean will shift. A positive relationship between R&D and firm employment growth was found, especially in the higher quantiles. Thus, the pattern is not completely symmetric and will give high-growth firms an additional boost relative to low-growth firms. It was also concluded that the short-run and long-run effects of R&D

are overlapping and converging. Such detailed results emerge by using quantile regression. Falk (2012) conducted similar research among Austrian businesses to “explore the parameter heterogeneity in the relationship between R&D and firm growth across the conditional growth distribution” (p. 20). In other words, they were investigating and testing the degree of variability in this relationship. Again, it is concluded that R&D investments mostly pay off for faster-growing firms. Significant and positive coefficients are only found for the middle and upper quantiles. For shrinking or slow-growing companies, the relationship is negative, so it is not worthwhile to invest in R&D.

Hölzl (2009) also utilized quantile regression to examine the relationship between R&D and growth, only as a robustness test this time. Again, this was argued because a more complete picture can be created by this method. However, they also mentioned explicitly that they were looking for determinants of this growth. The country in which the company is located would be one of these determinants. Through quantile regression, it was found that firms in countries closer to the technological frontier would benefit more from R&D. Hölzl (2009) suggests that this is because there are more opportunities rather than using existing solutions. Coad & Rao (2008) also used quantile regression to identify determinants of relationships. In this case, determinants can explain the strong growth of firms. The authors concluded that innovation (R&D) is a strong determinant of firm growth, especially in fast-growing companies. The fast-growing companies owe a lot to being innovative and investing in R&D. Segarra & Teruel (2014) also used quantile regression to identify determinants of innovation. In general, innovation and R&D have a positive impact on growth, but they came to an important finding. Internal R&D is fundamental in the higher quartiles and external R&D in the quartiles up to the median. Something that suggests that investing in internal R&D is primarily an essential activity for fast-growing companies, a conclusion we have encountered many times before.

3.1.3 Fixed & Random Effects

Fixed and Random Effects Models are also types of regression analysis that can be used in similar studies. There is a clear difference between the two, which can be found in the independent variables used. Fixed effects are constant across individuals. Random effects vary between individuals (Gelman, 2005). Green and Tukey (1960) describe the difference as “When a sample exhausts the population, the corresponding variable is fixed; when the sample is a small (i.e., negligible) part of the population the corresponding variable is random” (p. 131). This means that the averages of the corresponding variables in a fixed-effects model are based on a more representative sample. For example, a mean group was created based on a survey of 1000 random Dutch people. This is representative of the entire population. This is the case with the fixed-effects model. The variance is low, and this method is appropriate for studies with little heterogeneity. In a random-effects model, as indicated, the average is based on several smaller samples. The same question is then asked again to the Dutch population. However, the samples are now divided: 800 times to the elderly, 100 times to children, and 100 times to other persons. In the end, the 1000 answers will be different, but regardless of the question, the group mean will be completely different. This is what happens with a random-effects model, although the example is radical. There will be more variation. A random-effects model is appropriate for studies that use samples with high heterogeneity. It can be used well when it is expected that there are differences across individuals.

At first, Lee (2020) used both a fixed-effects model and a random-effects model to study the impact of R&D investments on the market value of Chinese companies. Subsequently, a Hausman test was used to check which model is statistically more robust. The Hausman test is the most common test that determines whether a fixed-effects model or a random-effects model is more appropriate. In this study, the fixed-effects model was chosen, and it was concluded that R&D investments have a strong impact on the market value of Chinese companies. Sohn et al. (2010) applied a fixed-effects model and a between-effects model in their research. The coefficients of the between-effects model were larger, suggesting that the variance between firms is greater than the variance within each firm. Another important point was mentioned, the time-invariant variables have been removed from the fixed-effects model. These are variables that do not vary over time, and therefore they are removed from the fixed-effect model. Only independent variables have used that change because otherwise, they would not have any influence on the dependent variable. These can be, for example, gender, race, or education (Beck, 2011). However, these variables can be used in a random-effects model.

3.1.4 Generalized Method of Moments (GMM)

Generalized Method of Moments (GMM) is another method of estimating statistical results. It is often used in cases of dynamic panel data. GMM is based on population moment conditions. A moment condition is created using the parameters of interest and is a notation set to 0. The sample data is then searched for the persons or companies in the sample that most resemble this moment condition and thus come closest to zero. By using population moment conditions, it is possible to estimate the actual parameters in the sample (Hansen, 1982). According to Roodman (2009), GMM is an appropriate method if the data used has specific characteristics. In the first place, GMM can be a convenient method when there are few periods but many individuals in the sample. Second, the variables are dynamic but not strictly exogenous. At last, there may be heteroskedasticity and autocorrelation within individuals but not between them.

There are two different types of GMM. The standard method is called the difference GMM. The second type is an extension and is called system GMM. In this, the instrument variables are assumed to have no relationship to the fixed effects. This allows more instrument variables to be created, which increases efficiency. Because in this way, two systems are created, it is called system GMM (Roodman, 2009). This difference is also indicated as one-step GMM (difference GMM) or two-step GMM (system GMM). Windmeijer (2005) describes the difference in the following way "One-step GMM estimators use weight matrices that are independent of estimated parameters, whereas the efficient two-step GMM estimator weighs the moment conditions by a consistent estimate of their covariance matrix" (p. 26).

There are also several studies related to the subject of R&D that use GMM as an estimator. Both Jaisinghani (2016) and Sharma (2012) investigated the correlation between R&D and firm performance in the Indian pharmaceutical industry. Both used system GMM and argued their choice in the same way. In system GMM, it is allowed to add instrument variables. This increases the reliability of the results and reduces bias. Both Jaisinghani (2016) and Sharma (2012) give this as their justification for system GMM. Poldahl (2011) also utilized system GMM to add instrumental variables to decrease bias. Only he researched the effect of R&D on growth among Swedish manufacturing firms. Chen et al. (2019) also used two-step GMM because it generates more efficient and exact results, avoids endogeneity issues, and creates more insight into the model and variables over time. System GMM is much more widely used than difference GMM because of its increased efficiency and reliability.

3.2 Variables

In this section, the used variables will be discussed, and the choices for these will be explained. There is a distinction between dependent variables, independent variables, and control variables.

3.2.1 Dependent variables

The dependent variable, a firm performance measurement that changes when the amount of money invested in R&D is adjusted upward or downward, is determined in many ways in the existing literature. However, they can be divided into two categories. A category of market-based measurements and a type of accounting-based measurements. The first category of variables is classified into market value, meaning R&D investments affect a company's market value. It is a well-known method that is widely used (Lee, 2020; Seo & Kim, 2020; Vithessonthi & Racela, 2016). The market value is then often measured in Tobin's Q. It is an excellent way to determine if your company is undervalued or overvalued and compare your company with other competing ones in the market. The second category, based on accounting-based measurements, often looks at the effect of R&D on firm performance, growth, or profitability. In many cases, these are methods that measure how efficiently a company is operating. This can be done in many ways, but the most used variables are the ROA (Chen et al., 2019), ROE (Coombs & Bierly, 2006), profit margin (Seo & Kim, 2020), earnings per share (Yang et al., 2010), or different growth rates such as firm employment growth (Capasso et al., 2015), sales growth (Falk, 2012), or total factor productivity (TFP) (Sharma, 2012). Primarily the ROA is a widely used method regarding the subject of R&D.

3.2.2 Independent variables

Most studies use R&D intensity as independent variable. It reflects how much is invested in R&D in relative terms. A high R&D intensity means that a company invests relatively much in developing its products. But there are different ways to calculate this R&D intensity. A well-known method is the total R&D investment divided by the company's total sales (Lee, 2020; Seo & Kim, 2020; Chen et al., 2019; Falk, 2012). It is a way that properly reflects how much of the incoming capital is simultaneously reinvested in developing the company's products or services. Another well-known way of calculating R&D intensity is to divide the total amount of R&D expenditures by total assets (Vithessonthi & Racela, 2016; Lin et al., 2012; Eberhart et al., 2004). This is not dependent on sales and is perhaps easier to use as a guideline. For example, a company that wants to invest 5% or 10% of the value of its total assets in R&D each year. Another way of calculating the intensity is by dividing the total investment amount by the number of employees in the company (Sher & Yang, 2005).

An interesting and important issue regarding this type of studies is the time lag that is used. The economic effect with a time lag implies that the actual consequences of an investment in R&D occur at a later period. However, the exact length of the period is unknown and depends on the investment and the project (Lee, 2020). The authors' opinions regarding similar studies are mixed. Gui-long et al. (2017) made a very deliberate choice to use dependent variables without lag in their research. This was argued because due to the product's varying industry or life cycle stage, the time span can be divergent per R&D investment. Nevertheless, a positive relationship between R&D intensity and firm performance was found. The quantile regression shows that the coefficients become higher in the higher quantiles. Firms that invest more in R&D also benefit more in relative terms. They also cited a study by Yeh et al. (2010). Yeh et al. (2010) also researched R&D intensity and the effects on firm performance and extensively mentioned the topic of time lag. They also used no lag and indicated that using different time lags creates mixed results and would remove the focus from the actual issue. Besides, the topic would now be less relevant because the life cycles are shorter, and therefore the R&D effects are noticeable more quickly.

However, most of the existing literature does make use of a time lag. The period does vary across studies. A part of the studies conducted both a regression with lag and without lag to make comparisons with each other (Xu & Sim, 2018; Lee, 2020; Chen et al., 2019). However, most studies use a time lag that usually varies between 1 and 5 years (Jaisinghani, 2016; Vithessonthi & Racela, 2016; Falk, 2012; Capasso et al., 2015; Booltink & Saka-Helmhout, 2018; Seo & Kim, 2020). Adding a lag avoids endogeneity problems and allows authors to make better and more accurate estimates (Seo & Kim, 2020). There may be a negative relationship between R&D investment and growth or firm performance without lag, while there is a positive relationship in the longer term (Paula & Silva, 2018). Therefore, a lag can ensure that better estimates can be made. Again, because the economic impact of the earlier investments will occur later. Thus, implementing lagged dependent variables or not is an important issue and choice to make.

3.2.3 Control variables

In addition to R&D intensity, many other factors can affect firm performance. Many of them have been discussed and used in the existing literature. In this chapter, several of these factors that can be used as control variables will be discussed.

The control variable that is most used is probably the firm size. This has to do with the fact that the size of a firm has a massive impact on the income that is available to be spent. A large firm can invest much more in R&D than a small one. This would also allow them to benefit more from it. They represent a greater value than smaller companies, but this does not mean that they invest proportionally more in R&D than a smaller company. A small pharmaceutical company generally has a higher R&D intensity than a large multinational, while in amounts of money this would be reversed. On the other hand, a larger company often has greater access to complementary resources because of their network, image, and experience (Coombs & Bierly, 2006). Altogether, it is a topic that should be considered while investigating the subject of R&D. This can be calculated in several ways; natural logarithm of total assets (Seo & Kim, 2020), firm employment plus one (Capasso et al., 2015), or natural logarithm of annual sales (Gui-long et al., 2017). Using the natural logarithm of total assets is the most common way to measure firm size.

Another control variable that has some financial influence is the leverage of a company. This is the ratio of debt to equity. Again, this indicates how much a company has to spend and how much it could benefit from R&D investments. High leverage indicates a certain level of risk. Companies with higher leverage have higher and stricter payment obligations that restrict them from investing in R&D (Guo et al., 2018). As a result, high or low leverage does affect a company's R&D intensity. Therefore, it is an interesting variable for a study like this. Leverage is mostly measured by the ratio of total liabilities to total assets (Guo et al., 2018; Xu & Sim, 2018).

Firm age is another factor that could affect the cost-effectiveness of R&D investments for several reasons. Lin et al. (2012) illustrate this by presenting several arguments from the literature. On one side, an older firm might innovate well because of the more experienced and stable organization. On the other side, it is also said that aging counteracts innovation. A fresh and new company that looks at a specific product with a renewed perspective might be more capable of innovating a product or process. Therefore, it is interesting to add firm age as a control variable. It can be calculated by simply adding up the number of years since it was founded or using a natural logarithm (Seo & Kim, 2020; Gui-long et al., 2017). Another way is to create a dummy variable that distinguishes old and young firms (Booltink & Saka-Helmhout, 2018).

Another interesting issue is the differences between industries. It is well known that, for example, high-tech companies and pharmaceutical companies are more dependent on the latest developments and consequently on R&D. In the case of a large sample study with many different industries, there will be a difference in this as well. Coombs & Bierly (2006) also indicate that the industry characteristics and environment are very influential on the ultimate development possibilities. Booltink &

Saka-Helmhout (2018) created a dummy variable that differentiated service firms from manufacturing firms assuming that manufacturing firms would benefit more from R&D. A finding previously made by Ho et al. (2005) as well. Manufacturing firms benefit more through R&D by introducing more innovative products into the market. Service firms could better spend on advertising and marketing. This separation makes it an interesting topic. Capasso et al. (2015) created even 51 dummy variables so that each 1 industry has its dummy variable. This allowed for an analysis of each industry separately. Coad & Rao (2008) also used an industry dummy as a control variable.

Another influential factor related to industry characteristics is country characteristics. The environments and facilities needed for development differ significantly from one country to another. According to Pindado et al. (2015), different country-level characteristics can influence the success of R&D. These include factors such as investor protection, the country's financial system, control mechanisms, and corporate governance. Alam et al. (2020) successfully analyzed the influence of investor protection and country governance on the relationship between R&D and firm performance. The impact of investor protection on this relationship is positive. But the relationship of country governance, measured by factors such as the level of corruption, the rule of law, and political stability, is negative. Hillier et al. (2010) concluded that both better developed financial systems and more robust corporate control mechanisms reduce the R&D to cash flow sensitivity. As a result, the efficiency of R&D can vary greatly. However, the relevance of this control variable depends on the countries used in the sample. The contrast between an EU country and an underdeveloped African country is much larger than between EU countries.

Two other factors used as control variables are the export intensity and import intensity (Sharma, 2012; Jaisinghani, 2016). Sharma (2012) suggests that firms are more productive and efficient when they enter foreign markets. This, in turn, would have a positive impact on the results of R&D investments. Firms with a higher import intensity would also benefit from this because they receive more technology and inputs, increasing productivity. These inputs and technology can in turn be used for the R&D of their products. Usually, import intensity and export intensity are calculated by dividing total imports and exports by sales.

Multiple dummies can be added to control for fixed effects. Year or time dummies are added to control for the impact of economic changes in the environment (Xu & Sim, 2018; Gui-Long et al., 2017). This can be convenient and important in a longitudinal study like this one. To control for differences between countries in the sample, also related to the country characteristics mentioned above, a country dummy can be added (Booltink & Saka-Helmhout, 2018).

3.3 Model Specification

OLS regression will be utilized to estimate the impact of R&D on firm performance. There are several reasons for this. The estimates from an OLS estimator are the most accurate, and therefore it reduces the potential bias. This makes it a widely used method, which is also reflected in the existing literature on the subject. The OLS regression is the most used method to measure the impact of R&D on firm performance in a variety of ways. The equation of the OLS regression will look like this:

$$\text{FIRM\_PERFORMANCE}_{it} = \beta_0 + \beta_1\text{R\&D}_{it} + \beta_2\text{R\&D}^2_{it} + \beta_3\text{CONTROL}_{it} + \varepsilon_{it}$$

Where:

$\text{FIRM\_PERFORMANCE}_{it}$	= Firm performance for firm i in year t
β_0	= Intercept
$\beta_1\text{R\&D}_{it}$	= R&D intensity of firm i in year t
$\beta_2\text{R\&D}^2_{it}$	= R&D intensity^2 of firm i in year t
$\beta_3\text{CONTROL}_{it}$	= Control variables of firm i in year t.
ε_{it}	= Error term

The return on assets (ROA), the profit margin, and the earnings per share (EPS) will be used to measure firm performance. The firm size, firm age, firm leverage, and an industry dummy are used as control variables. A year dummy and country dummy are added to control for the time-invariant effects of these factors. The definitions and calculation methods of the variables in this study can be found in Appendix 2.

In addition to the first regression model, several robustness tests will be conducted. All regressions will be performed without lag, with a lag of one year, and with a lag of two years. The value of implementing a lag has been discussed in detail before. Besides this, the sample will be divided based on the R&D intensity of the firms. This will investigate whether there is a difference between low R&D and high R&D companies. Lastly, alternative dependent and independent variables are used to support or refute previous results. In brief, there are different ways to test the same hypotheses and to create certainty in this way.

4. Data

The required data for this research is collected through ORBIS. This is a database with financial information of firms worldwide. Almost all required data can be found in ORBIS. For some cases, the annual reports of the companies in the sample were used. This to fill in some missing values. It involves a sample of listed firms located in the most developed European countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, The Netherlands, Portugal, Spain, and Sweden. The time period of the study is set from 2011 to 2019.

The sample size is affected significantly because there is filtered on companies that publish R&D expenditures continuously throughout the sample period. Many companies publish the costs of a multi-year R&D project once instead of spreading them evenly over the years. Or they incorporate the R&D costs somewhere else in their financial statements. This strongly reduces the sample size, although it is difficult to estimate to what extent. After screening listed companies that are from the EU and publish annual R&D data, a sample size of 521 firms is presented.

After this, all the annual data for the required variables must be retrieved from the database. The ROA, profit margin, and EPS can be extracted directly from ORBIS. The gearing ratio, which is used to measure leverage, can also be obtained from ORBIS. The total assets and the total R&D expenditures are required to calculate the R&D intensity for each year. The total assets were also necessary for the log of firm size, and the year of incorporation was used to apply firm age as a control variable. The industry dummy, which distinguishes service and manufacturing firms, could not be obtained directly from ORBIS. However, it was known in which industry they operate, and the NACE Rev. 2 code attached to it. Using this code, a differentiation can be made. According to Eurostat, the European Statistical Office, and part of the European Parliament, all NACE Rev. 2 codes below 3300 can be assigned as manufacturing companies. Companies with a NACE Rev. 2 code above 4500 are mainly service companies. 22 companies fall in between, and these are assigned to service or manufacturing based on the main activity that is shown on ORBIS.

After adding all the necessary variables, 49 firms were removed for varying reasons. A few companies were too small because they had less than 5 employees. After an extensive data check, several more companies were removed for various reasons. In some cases, extreme values in ORBIS did not match with the data of the annual reports. Observations with missing values and companies that were not in the dataset for at least three years were also removed. In the end, a sample size of 472 firms with 3955 firm-year observations remained.

Table 1: Data selection and exclusion

Reason for excluding firms	Number of excluded companies
Firms located in the European Union (14)	375,463
- Unlisted firms	- 364,268
- Firms without annual R&D expenses	- 10,672
European listed firms that publish R&D costs annually	= 521
- Firms with < 5 employees	- 4
- Firms removed after data check	- 45
Final sample size	= 472

If looking at the distribution of R&D-intensive companies across Europe, there are a few remarkable things to be noticed. Germany (155) has by far the largest share of the sample, followed by France (72). These are also the largest countries in terms of population. The most striking thing is the large share of the Scandinavian countries in the sample size compared to the southern located countries. Denmark (20), Finland (44), and Sweden (64) together have 128 firms in the sample. This while Greece (19), Italy (5), Portugal (0), and Spain (9) have only 33. This is remarkable when comparing this to the total population numbers. This shows that the Scandinavian countries are more developed and much more involved in R&D. The other countries, Austria (24), Belgium (19), Ireland (17), Luxembourg (4), and the Netherlands (20) do not show any remarkable results. The division between manufacturing and service firms is also clear. A total of 361 firms are seen as manufacturing firms and 111 of the firms are focused primarily on providing services.

Before starting the statistical part, it is crucial to make decisions about how to deal with outliers. A well-known way to deal with outliers is through winsorizing. A technique that, according to Tukey (1962), does not remove the outliers but "replacing its original value by the nearest value of an observation not seriously suspect" (p. 18). Because the variables are modified but not deleted, the data points remain in the dataset. Winsorizing makes the dataset more robust by reducing the extreme values. According to Aggarwal (2013), there are outliers when it looks like specific observations were generated differently and are therefore so anomalous. That is not what is happening here. Therefore, the dependent and independent variables are winsorized at the 2.5th and 97.5th percentiles. The leverage and R&D intensity are also transformed to their natural logarithm to correct for skewness and increase the reliability of the results. There is significant variance among the firms in the dataset and the logarithms ensure that the distribution is normal. The logarithms were calculated using LN(1+X). The +1 is added to prevent negative log values. Something that would lead to incorrect results when squaring the values.

5. Results

5.1 Descriptive Statistics

Table 2: Descriptive Statistics Full Sample

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	3955	0.061	0.281	-0.153	0.079	0.025	0.059	0.097
Profit Margin	3955	0.069	0.299	-0.241	0.101	0.026	0.066	0.118
EPS (€)	3955	1.807	15.62	-2.720	3.245	0.167	0.850	2.293
Independent variables								
R&D Intensity (%)	3955	0.039	0.177	0.000	0.044	0.008	0.023	0.056
Log_RD Intensity	3955	1.260	2.929	0.047	0.817	0.567	1.178	1.889
Log_RD Intensity²	3955	2.254	8.577	0.002	2.374	0.321	1.387	3.570
Control variables								
Size (€ in millions)	3955	8.616	104.3	0.018	20.47	0.168	1.047	5.841
Log (Size)	3955	13.84	18.46	9.821	2.274	12.03	13.86	15.58
Age	3955	66.17	186.0	14.00	48.00	28.00	45.00	98.00
Log (Age)	3955	3.953	5.231	2.708	0.718	3.367	3.829	4.595
Leverage	3955	0.846	3.601	0.010	0.766	0.320	0.670	1.080
Log (Leverage)	3955	0.544	1.526	0.010	0.355	0.277	0.513	0.732
Industry_Control	3955	0.770	1.000	0.000	0.419	1.000	1.000	1.000

The table above shows the descriptive statistics of the used variables. All variables show 3955 firm-year observations. The average ROA is 6.11%, which is slightly higher but similar to the ROA of 4.5% of the study by Guo et al. (2018) but a lot below the ROA of 15.55% reported in the research of Jaisinghani (2016). Guo et al. (2018) used a sample of Chinese listed firms, which could explain the difference. It was also found that ROA in our study increased a bit after using winsorization, which is not discussed in the study by Guo et al. (2018). Jaisinghani (2016) used a small sample with only pharmaceutical companies that generally have high ROA and R&D Intensity. Something that is also reflected among the pharmaceutical companies in our sample. The profit margin shows an average value of 6.94%. This also does not differ much from the 7.2% profit margin found in the study by Guo et al. (2018). With both ROA and profit margin, the median hardly deviates from the mean, suggesting an absence of skewness. From this, it can be concluded that there is a normal distribution in both cases. The median ROA is 5.94%, and that of the profit margin deviates a little bit more and is 6.62%. The situation is slightly different for EPS. The mean of EPS is €1.81 with a maximum of €15.62 and a minimum of €-2.72. The negative minimum value demonstrates that there are also loss-making companies in the sample. The median of €0.85 is below the mean, which leads to a small right-skewed distribution.

The average R&D intensity after winsorizing is 3.97%, with a maximum of 17.70%, and a minimum of 0.05%. The found percentage is similar but slightly higher to the mean found in the studies by Guo et al. (2018) and Gui-long et al. (2017). Average values of 3.2% and 3.47% were found there. Both studies used a sample with Chinese companies, which could cause the difference. However, Gui-long et al. (2017) specifically choose electronic manufacturing firms. Looking at the natural logarithms of R&D intensity, we find a mean of 1.260. The median is 1.178, so there is almost no skewness. There is significantly more skewness when logarithms are not used. The mean logarithm of R&D intensity² is 2.254. The median logarithm is 1.387, which causes a small right-skewed distribution. The skewness is amplified by squaring the variable. In all three variables, there is a slight skewness to the right. This is caused by the relatively high number of companies with very low R&D intensity. Therefore, it was decided to create two subsamples: a sample with low R&D-intensive companies and a sample with high R&D-intensive companies. It is an additional robustness test, and it may yield interesting findings. The low R&D-intensive sample consists of companies that have an average R&D intensity of less than 1%. It has been found that mainly this group provides the skewness. Thus, the high R&D-intensive sample consists of companies with an average R&D intensity of higher than 1%.

Logarithms have been used in all control variables, and therefore no skewness can be found. The descriptive statistics of the logarithm of size are as follows: the mean is 13.84, the maximum is 18.46, and the minimum is 9.821. Which is equivalent to an average value of total assets of €8,616,335. The maximum is €104,253,000, and the lowest value is only €18,418. These statistics also clearly show the differences between the firms in the dataset. The year of incorporation used for the logarithm of firm age also reveals no unusual information. The mean of the natural logarithm of age is 3.95, equivalent to a firm age of 66.17 years. The youngest company is 14 years old, and the oldest company is 186 years old. The mean logarithm of leverage is 0.54 and is comparable to a leverage of almost 0.85%. This means that for every €1 of shareholders' equity, there is also €0.85 of debt on the other side. This is a high average, which can cause problems with raising new capital from investors. The maximum logarithm is 1.53 (3.60%), and the minimum is 0.010 (0.10%).

The descriptive statistics where the 1-year and 2-year lagged dependent variables are used do not show much unusuality. Since there are few striking issues in these tables, they are listed in Appendix 3. The ROA and profit margin remain almost the same when using the lagged variables. The EPS increases somewhat during each sample. The differences are still minimal. The R&D Intensity remains the same in all three samples. Size rises a little each time, which is logical and explainable. The oldest firm-year observations are removed during the implementation of the lag. In general, firms grow, and then these firm-year observations are the smallest in size. Removing them will therefore increase the average size, something that is applicable here. Leverage also shows a minimal but not meaningful decrease.

Table 3: Descriptive Statistics Low R&D-Intensive Sample

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	1112	0.055	0.197	-0.067	0.057	0.020	0.052	0.084
Profit Margin	1112	0.070	0.278	-0.103	0.077	0.024	0.062	0.110
EPS (€)	1112	1.481	11.88	-4.741	2.799	0.189	0.869	2.147
Independent variables								
R&D Intensity (%)	1112	0.004	0.011	0.000	0.003	0.002	0.003	0.006
Log_RD Intensity	1112	0.323	0.748	0.009	0.209	0.141	0.296	0.499
Log_RD Intensity²	1112	0.148	0.559	0.000	0.153	0.020	0.088	0.249
Control variables								
Size (€ in millions)	1112	13.82	139.9	0.033	29.98	0.325	2.053	9.856
Log (Size)	1112	14.47	18.76	10.39	2.220	12.69	14.53	16.10
Age	1112	70.77	186.0	14.00	48.25	30.00	56.00	110.0
Log (Age)	1112	4.031	5.231	2.708	0.716	3.434	4.043	4.710
Leverage	1112	1.067	4.308	0.090	0.929	0.440	0.820	1.268
Log (Leverage)	1112	0.648	1.669	0.086	0.377	0.365	0.598	0.819
Industry_Control	1112	0.740	1.000	0.000	0.441	0.000	1.000	1.000

The subsample of low R&D-intensive companies consists of 132 companies, 98 are manufacturing, and 34 are service companies. Which is not a striking distribution. More remarkable is that 22 of the 33 companies from Southern European countries are in this sample. This while the Scandinavian countries are only represented with 29 of their 128 companies in this sample. This confirms the previous findings of the development of these regions. Also, the generally well-developed countries Germany (32), Belgium (3), Ireland (4), and the Netherlands (5) have a significantly lower presence. This is reflected in Figure 1, which can be found later. The low R&D-intensive subsample consists of 1112 firm-year observations, which is 28.12% of the total amount.

The mean ROA in this subsample is 5.47% which is lower than the ROA in the entire sample. Also, the EPS is much lower with a value of €1.48 against a value of €1.81. For the profit margin, this is not the case. The average profit margin is 7% and therefore even slightly higher. The much lower dependent variables could be a sign, which shows that companies with a lower R&D intensity also obtain poorer results. The medians show that the skewness for ROA and profit margin is still minimal and reduced for the EPS. The R&D intensity is of course very low in this sample. Written out, the mean is 0.41%, and the median is 0.34%. That the mean is still only 0.41% shows the influence of the many low values in the whole sample. The R&D intensity corresponds to the 0.72% found for large firms and 0.83% for small firms in the study by Xu et al. (2019). They use Korean firms and indicate that the R&D intensity in better-developed countries fluctuates in the range of 4% to 10%.

The mean size measured in total assets is remarkable and significantly larger than in the full sample. The mean is €13,817,134 and thus more than €5,000,000 larger than in the full sample. Also, the firm age of 70.77 is several years higher than the 66.17 in the full sample. An explanation for this could be that the smaller and younger startups are in the other subsample. Small startups are more often high-tech companies with higher R&D intensity, but these pull down the average age and size in the sample. They would then be missing in this subsample, and therefore, both the age and size would be higher. The industry dummy is 0.74, which is slightly lower than in the full sample. This means that there are slightly more service companies that have joined. The descriptive statistics with the lagged variables do not show significant differences and can be found in Appendix 3.

Table 4: Descriptive Statistics High R&D-Intensive Sample

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	2843	0.063	0.324	-0.202	0.093	0.026	0.063	0.101
Profit Margin	2843	0.067	0.302	-0.329	0.116	0.027	0.068	0.121
EPS (€)	2843	1.919	16.60	-2.087	3.415	0.160	0.840	2.340
Independent variables								
R&D Intensity (%)	2843	0.054	0.204	0.009	0.048	0.019	0.035	0.074
Log_RD Intensity	2843	1.632	3.062	0.615	0.663	1.096	1.513	2.137
Log_RD Intensity²	2843	3.103	9.373	0.379	2.384	1.201	2.288	4.568
Control variables								
Size (€ in millions)	2843	6.693	86.12	0.015	16.17	0.129	0.754	4.433
Log (Size)	2843	13.59	18.27	9.631	2.253	11.77	13.53	15.30
Age	2843	64.39	186.0	15.00	47.77	27.00	41.00	97.00
Log (Age)	2843	3.924	5.231	2.773	0.715	3.332	3.738	4.585
Leverage	2843	0.759	3.244	0.010	0.678	0.280	0.610	1.020
Log (Leverage)	2843	0.504	1.456	0.010	0.337	0.247	0.476	0.703
Industry_Control	2843	0.790	1.000	0.000	0.410	1.000	1.000	1.000

The high R&D-intensive sample is a lot larger than the low R&D-intensive sample and consists of 2843 firm-year observations. This represents the other 71.88% of the sample. The subsample consists of 340 firms, of which 263 are classified as manufacturing companies, and the remaining 77 are mainly focused on services. This ratio is hardly different than in the full sample. The distribution is the opposite of the other subsample, which is obvious. This means that the remaining 99 Scandinavian countries are in this sample, and only 11 Southern European countries are represented in it. Other well-developing countries such as Germany (123), Belgium (16), Ireland (13), and the Netherlands (15) are also overwhelmingly more present. A clear illustration of this distribution is shown in Figure 1, which can be found below.

Many results from the descriptive statistics are opposite compared to the low R&D-intensive subsample. When a variable goes up in the low R&D sample, it often goes down in the high R&D sample and vice versa. Therefore, I will discuss this a bit more briefly. With the dependent variables, this pattern is immediately apparent. Where these were lower in the low R&D-intensive subsample, this is not the case here. The ROA of 6.34% is higher than that of 6.11% in the full sample, and the EPS of €1.91 is higher than the €1.81 that was found. However, the profit margin of 6.71% is somewhat lower than the 6.95% found earlier. Again, this may be a small sign of support for hypothesis 1, but this can only be observed after conducting the regressions.

The average R&D intensity is 5.44%, which is a lot higher than in the full sample. The difference between the low R&D sample and the high R&D sample is more than 5%, which will undoubtedly give different results. The mean of 5.44% is slightly lower but similar to the 6.57% found by Vithessonthi & Racela (2016). They used US companies which are also well developed like most European countries and also used winsorization. This makes it an excellent study to compare results with. The natural logarithms of both R&D intensity and the squared term show that the mean and median are close to each other and do not cause problems in terms of skewness. As indicated earlier, the average size of a firm in the low R&D-intensive sample was quite a bit larger than in the full sample. Here it is slightly lower with a mean of €6,692,804 compared to a mean of €8,616,335. The firm age, as expected, is somewhat lower with a mean of 64.39 years old. The same applies to the leverage of 0.759. In all cases, this is as expected, and the results are not striking. The averages of the used natural logarithms are close to the medians and ensure no skewness.

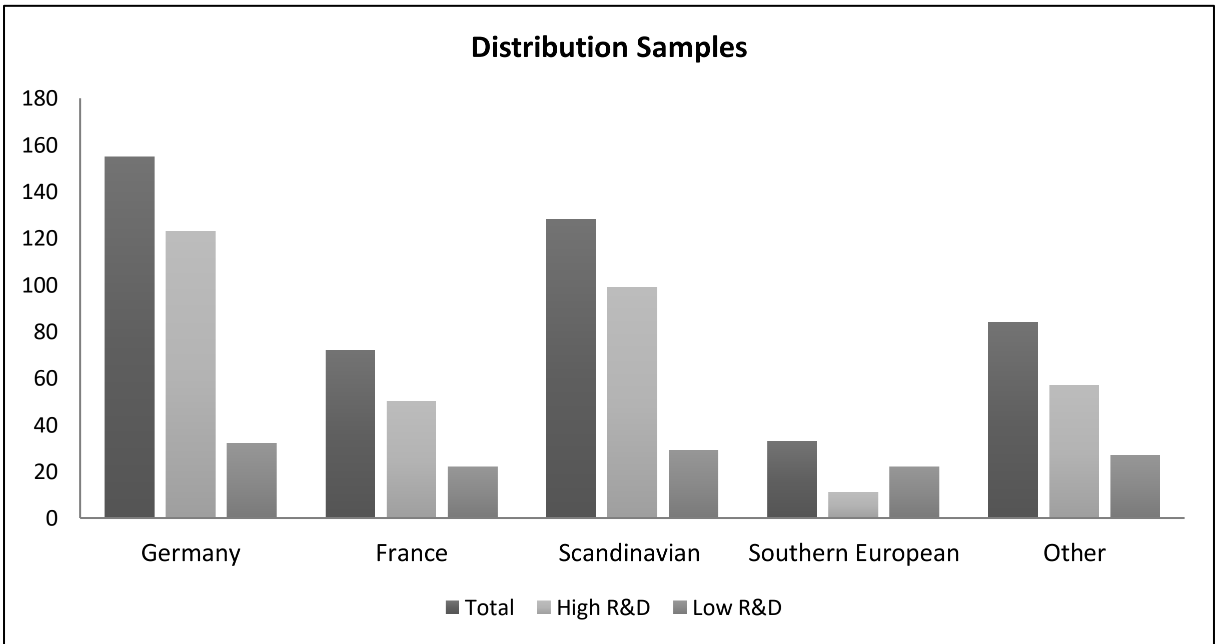


Figure 1: Distribution of the countries among the samples

5.2 Bivariate Analysis

5.2.1 Pearson's Correlation Matrix

The correlation matrices of the different samples can be found in the tables below. In a correlation matrix, the strength of the relationship between the variables in the study can be tested. The most striking and noteworthy results from these matrices will be discussed. The most important matrix is that of the full sample, which can be found in Table 5. An important point is that all dependent variables are positively and significantly correlated with each other. These correlations were expected because they all measure the same thing, namely the firm performance of the firms in the sample. The striking thing is the significant difference between the correlations in the ROA. The ROA and profit margin are highly correlated with a coefficient of 0.884** while the coefficient for EPS is only 0.316**. A significant difference, therefore. A possible explanation of the high correlation between ROA and profit margin is that both variables use EBIT in their calculation while EPS does not.

The correlations between the R&D intensity and the different dependent variables are mixed and, therefore, interesting. The coefficient of ROA (0.059**) is positive but not very high. This indicates a positive but weak relationship between R&D intensity and firm performance. Also, because of the negative coefficient of EPS (-0.086**), the results are not convincing. The correlations between R&D intensity^2 and the dependent variables are similar, which is not surprising. A squared term was used, which is visible in the high coefficient of 0.961** between R&D intensity and R&D intensity^2. The high coefficient could indicate multicollinearity. However, a high and strong correlation between the linear and quadratic term is logical in curvilinear relationships. The high correlation between the linear term and squared term is not relevant because it does not affect the estimation of the whole curve (Disatnik & Sivan, 2014). Therefore, multicollinearity is not an issue in this situation and can be ignored.

Looking at the correlations of the different control variables, there are few strange things. There is a clear high and significant correlation between size and EPS of 0.316**, which suggests that larger companies get better results when measured in EPS. Also, an older company performs better if looking at the significant coefficient of 0.239** that can be found between age and EPS. A larger company also performs better in terms of ROA (0.048**) and profit margin (0.181**), but not as significantly as it does in EPS. The amount of issued shares also matters here. More striking are the negative coefficients of size (-0.330**) and age (-0.188**) with R&D intensity. This would imply that firms with more capital and experience invest proportionally less in R&D. This is consistent with the previously cited research by Lin et al. (2012), which argues that aging counteracts innovation. Stable and experienced organizations are more likely to like things the way they are. Here are opportunities for start-ups that then invest relatively heavily in R&D. In addition, there is a clear positive relationship between size and age (0.250**), which is not surprising. An older company had the opportunity to grow for a longer period, so they are often larger, resulting in this positive correlation.

Furthermore, all three dependent variables ROA (-0.304**), profit margin (-0.216**), EPS (-0.036*) and the two independent variables regarding R&D intensity (-0.327**) are negatively and significantly correlated with leverage. Higher leverage, meaning more debt, results in both weaker firm performance and lower R&D investment. A well-known pattern because high debt creates difficulties in attracting capital that could subsequently be invested in R&D. The higher default risk causes these difficulties among companies with higher leverage. The dependent variables are negative because high debt incurs costs that reduce the net income and EBIT. Factors that are used to calculate the dependent variables. There is also a clear positive coefficient of 0.349** between leverage and size which means that larger firms have higher leverage. This could be explained by the fact that larger companies prefer to use debt rather than issuing new shares when making new investments. Smaller companies are or not always able to do this. They are forced to issue shares and give up part of their equity. This is consistent with, for example, the pecking order theory. This could also explain the high relationship between EPS and size. Larger firms carry more debt and consequently fewer shares. As a result, the earning per share is relatively high and this firm performance measurement stands out in comparison to ROA and profit margin. The coefficients of the industry dummy with size (0.161**) and age (0.198**) suggest that manufacturing companies tend to be larger and older than service companies. However, the negative coefficients of R&D intensity (-0.069**) and R&D intensity^2 (-0.147**) show that service firms are relatively more likely to invest in R&D. Something consistent with a previously found conclusion, claiming that smaller and younger companies invest proportionally more in R&D.

The correlation matrices of the split samples can be found in Appendix 4. The main differences will be discussed, starting with the low R&D-intensive sample. The mutual coefficients between ROA and profit margin have decreased slightly. The coefficients of EPS have clearly and significantly increased. The correlations between the dependent variables and R&D intensity are somewhat changed. This applies mostly to the coefficients of the EPS. These were negative and significant but are slightly positive in this sample, which would mean that it is rewarding for low R&D firms to invest in R&D. Obviously, the correlation between R&D intensity and R&D intensity^2 is still positive and robust (0.963**).

The correlations of the control variables in the low R&D subsample do reveal some notable differences. The correlations of age and size with the dependent variables are strongly reduced. It can be concluded that the influence of age and size on firm performance is less influential in this sample. EPS is an exception here, even in this sample it appears that an older and larger firm will have a higher firm performance. Another remarkable change is that the strong positive and significant relationship between size and age is no longer there (-0.002). This is not logical, but after a data check, it can be seen that the oldest companies in this sample are also among the smallest companies. This subsample is the smallest and most sensitive to these kinds of outliers, which explains this coefficient and the other changes. The relationship between the industry dummy and R&D investments is positive (0.341**), signifying that manufacturing companies invest more in R&D than service companies. The correlation with the size of the firm (-0.232**) has also reversed and shows that service firms in this sample are on average larger than manufacturing firms. Both findings from the industry dummy were the opposite in the full sample.

The correlation matrix of the high R&D-intensive sample has much more similarity with the matrix of the full sample, which is probably caused by the size of the subsample. The dependent variables are almost equal compared to Table 5. The most striking finding is that all correlations between the R&D intensity variables and the dependent variables decreased. This while the R&D intensity among the firms is higher, and it is expected that this would improve the firm performance. This would suggest that R&D would not improve firm performance. A finding that is not in line with hypothesis 1. The control variables were slightly up or down in places but showed no significant differences from the full sample.

Table 5: Pearson's Correlation Matrix Full Sample

	ROA	Profit Margin	EPS	R&D Intensity	R&D Intensity²	Log (Size)	Log (Age)	Leverage	Industry Control
ROA	1								
Profit Margin	0.884**	1							
EPS	0.316**	0.303**	1						
R&D Intensity	0.059**	-0.008	-0.086**	1					
R&D Intensity²	0.044**	-0.019	-0.128**	0.961**	1				
Log (Size)	0.048**	0.181**	0.316**	-0.330**	-0.343**	1			
Log (Age)	0.038*	0.026	0.239**	-0.188**	-0.228**	0.250**	1		
Log (Leverage)	-0.304**	-0.216**	-0.036*	-0.327**	-0.330**	0.349**	0.113**	1	
Industry_Control	0.025	-0.030	0.167**	-0.069**	-0.147**	0.161**	0.198**	0.013	1

*** Correlation is significant at the 0.01 level (2-tailed).*
** Correlation is significant at the 0.05 level (2-tailed).*

5.3 OLS Results

5.3.1 Full Sample Results

Table 6: OLS Results Full Sample

	ROA		Profit Margin		EPS	
	1	**2**	**1**	**2**	**1**	**2**
Constant	2.161*	1.874	-5.044***	-5.518***	-5.513***	-5.731***
	(1.85)	(1.60)	(-3.45)	(-3.77)	(-12.17)	(-12.68)
R&D Intensity	-0.018	0.158***	-0.021	0.209***	-0.026	0.302***
	(-1.04)	(2.77)	(-1.25)	(3.68)	(-1.60)	(5.54)
R&D Intensity²		-0.185***		-0.242***		-0.345***
		(-3.23)		(-4.26)		(-6.32)
Log (Size)	0.180***	0.178***	0.314***	0.312***	0.313***	0.311***
	(9.99)	(9.94)	(17.57)	(17.52)	(18.19)	(18.14)
Log (Age)	0.029*	0.023	-0.001	-0.009	0.127***	0.117***
	(1.78)	(1.43)	(-0.09)	(-0.54)	(8.19)	(7.52)
Log (Leverage)	-0.361***	-0.363***	-0.322***	-0.325***	-0.200***	-0.204***
	(-21.56)	(-21.69)	(-19.42)	(-19.62)	(-12.48)	(-12.78)
Industry_Control	-0.011	-0.025	-0.078***	-0.095***	0.093***	0.068***
	(-0.74)	(-1.56)	(-5.13)	(-6.08)	(6.35)	(4.51)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	14.6%	14.8%	15.9%	16.3%	21.8%	22.6%
N	3955	3955	3955	3955	3955	3955

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total assets. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 6 shows the results of the OLS regression performed when using the full sample. The results of R&D intensity are mixed but simultaneously show a pattern. When using just the linear term in the model, a negative but insignificant coefficient appears in all cases. This indicates that R&D investments have no impact on firm performance. However, the models including the squared term are clearer and all coefficients are significant at the 1% level. For all three independent variables, the linear terms are positive, and the squared terms are negative. This could indicate an inverted U-shaped relationship. This means that investing in R&D is profitable up to a certain point after which a maximum is reached, and it would no longer be beneficial. According to Haans et al. (2015), this is only true if the β of the linear term is positive and significant and the β of the quadratic term is negative and significant. The findings from this regression met these conditions. This provides actual evidence for an inverted U-shaped relationship. A relationship also found by Booltink & Saka-Helmhout (2018). They tested it in the same way and used a sample with European enterprises. However, they focused specifically on SMEs and only used a lag of 2 years. When this would occur more frequently, hypothesis 2 would be rejected. This hypothesis states that there is a linear relationship between R&D and firm performance.

Accordingly, in the case of an inverted U-shaped relationship, there is a threshold up to which it is profitable to invest in R&D. In the case of a threshold or turning point, it is interesting to examine the height of this point. This can be done using the following equation: exp ($-\beta_1$ / $2\beta_2$) (Haans et al., 2015; Lind & Mehlum, 2010; Song et al., 2008). In this sample, this results in a turning point of 1.53% for ROA, 1.54% for profit margin, and 1.55% for EPS. On average, a turning point of 1.54%. A percentage that is very similar to the optimal level of R&D of 1.60% found by Yeh et al. (2010). They also used the ROA, although the turning point was found by using an advanced threshold model. In practice, this means that investing in R&D is profitable up to 1.54% of the asset value. Higher investments will not result in better firm performance. It is not a high percentage, but this could explain the slight negative coefficients when testing the linear terms only. However, it should be noted that this applies to firms in this sample and this turning point will vary by industry, country, and type of firm.

The positive and significant results found in size are not surprising after previously analyzing the correlation matrices. The results show, as expected, that a larger company generally achieves better results which is reflected in the ROA, profit margin, and EPS. A conclusion also made in similar research by Xu et al. (2019). According to Pervan & Višić (2012), this is due to factors such as higher market power, more market experience, and the economies of scale that a larger company has. Also, because of these economies of scale, larger companies would have advantages in the R&D process as they are more efficient. Additionally, larger companies are in more stable environments and are more likely to avoid risky R&D activities. Smaller companies will take more risks in this area with potentially serious consequences (Xu & Sim, 2018). The results regarding age are mixed and not as convincing as those of size. The coefficient of ROA in model 1 is positive, but only significant at the 10% level, which makes it not very strong. The results of EPS are more evident. The highest β is 0.127***, suggesting that an increasing age also leads to increasing earnings per share. A more senior firm has the advantage of benefiting from earlier learning processes which can result in better performance (Gui-long et al., 2017). Consequently, a positive relationship between the firm performance measurements and age is observed.

Further non-surprising results are the negative and highly significant coefficients at leverage. As explained in more detail earlier, higher debt leads to higher costs which are often reflected in the financial performance of a company. Something that can be seen here. The coefficients in the full models of ROA (β = -0.363***), profit margin, (β = -0.325***) and EPS (β = -0.204***) are clear and confirm these expected findings. These findings are consistent with the existing literature (Xu et al., 2019; Chen et al., 2019; Guo et al., 2018). Xu et al. (2019) also concluded that companies with lower leverage are more engaged in research activities. The results of the industry dummy are mixed. While the negative and significant β of -0.095*** in the full model of profit margin suggests that service firms generally achieve a higher profit margin, it is just the opposite for EPS. In fact, the β of 0.068*** shows that manufacturing firms achieve a higher EPS.

There are also regressions conducted in which the dependent variables are lagged by 1 and 2 years. This was done because similar research has regularly shown that implementing a lag can certainly show different results. This is because the impact of an investment in R&D is often not yet visible in the first year, but it is also difficult to determine when this will happen. For example, Chen et al. (2019) found a negative relationship between R&D intensity and current business performance, but a positive and significant relationship between R&D intensity and business performance with a lag of 2 years. There are numerous examples available in which this occurs. Thus, these positive results could have been achieved by R&D investments made earlier. To keep it organized, the results of the regressions with lagged dependent variables have been moved to Appendix 5. The number of firm-year observations decreases with each lag that is added. The main differences in comparison with Table 6 will be discussed.

Although the R&D intensity coefficients increase moderately during the lags, the conclusion stays the same. In this sample, R&D investments do not affect firm performance. Despite the lag, a significant inverted U-shaped relationship is still found in all full models. However, the coefficients of R&D intensity² have weakened during the two lags, especially for ROA and the profit margin. This changed for ROA from β = -0.185*** to β = -0.123* and for profit margin from β = -0.242*** to β = -0.166**. This could suggest that the R&D intensity increases at a slower rate to the maximum, and the same applies to the rate of decrease after the turning point (Haans et al., 2015). In this case, the U-shape would have a more moderate shape. The average turning points have increased slightly but not noticeably. The optimal level of R&D at a 1-year lag is 1.59% and at a 2-year lag 1.64%. It is in line with the positive coefficients because the higher threshold implies that more can be invested in R&D without losing money. The evidence to reject Hypothesis 2 stands because non-linear relationships were observed again. The control variables do not present any unusual results. The impact of size and leverage has slightly decreased but are still factors with a significant effect on firm performance.

5.3.2 Low R&D-Intensive Sample Results

Table 7: OLS Results Low R&D-Intensive Sample

	ROA		Profit Margin		EPS	
	1	**2**	**1**	**2**	**1**	**2**
Constant	12.51***	11.94***	9.566***	10.25***	-3.406***	-3.570***
	(6.92)	(6.49)	(3.80)	(4.00)	(-3.84)	(-3.96)
R&D Intensity	0.012	0.202*	-0.039	-0.206*	0.027	0.137
	(0.37)	(1.72)	(-1.24)	(-1.73)	(0.86)	(1.18)
R&D Intensity²		-0.189*		0.166		-0.110
		(-1.69)		(1.46)		(-0.99)
Log (Size)	0.004	0.006	0.127***	0.126***	0.232***	0.232***
	(0.12)	(0.16)	(3.62)	(3.58)	(6.76)	(6.78)
Log (Age)	-0.037	-0.037	-0.027	-0.027	0.089***	0.089***
	(-1.24)	(-1.22)	(-0.88)	(-0.89)	(2.97)	(2.98)
Log (Leverage)	-0.420***	-0.415***	-0.328***	-0.332***	-0.232***	-0.229***
	(-13.23)	(-13.04)	(-10.15)	(-10.24)	(-7.36)	(-7.25)
Industry_Control	-0.062**	-0.079**	-0.221***	-0.207***	0.086***	0.076**
	(-2.02)	(-2.43)	(-7.04)	(-6.29)	(2.79)	(2.37)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	18.8%	19.0%	16.2%	16.3%	20.2%	20.2%
N	1112	1112	1112	1112	1112	1112

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total assets. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 7 shows the results of the OLS regression of the low R&D-intensive subsample. The most significant differences between this sample and the full sample will be discussed. As in the full sample, the first models with the linear terms do not detect any significant coefficients. This means that R&D investments do not impact firm performance for firms with low R&D. In these firms, R&D is not a distinguishing factor, which can explain the results. However, R&D intensity² changed considerably and found much less significant coefficients. In the full model of the ROA, there is still a significant positive linear term and a significant negative squared term. Again, this means an inverted U-shaped relationship, this time with a turning point at 1.71%. This higher percentage suggests that it is worthwhile to do R&D investments up to a higher amount. Something consistent with the positive coefficients of the linear terms. The results of the profit margin reveal the exact opposite. Consequently, the linear term in the full model is negative (β = -0.206*) and the squared term positive (β = 0.166). These two coefficients are indicators of a regular U-shaped relationship between R&D intensity and firm performance (Haans et al., 2015). In the case of such a relationship, investments are not profitable at first, but it changes over time. After reaching a threshold, profits are made. In this situation, a competitive advantage is achieved after investing a significant amount in R&D. The inverted U-shaped relationship of EPS is no longer there.

When analyzing the control variables, it is striking that the impact of the firm size is strongly reduced. There are no significant coefficients found at ROA, and the remaining coefficients have decreased. With leverage, it is just the opposite, and the results are even lower than they already were. Firms with relatively low R&D intensity and high leverage have a lower firm performance. High leverage also makes it more difficult to invest in R&D due to limited financial resources. This higher leverage among low R&D intensity firms may also be a major reason why their R&D is at such a low level. Firms may deliberately choose to invest less in R&D, but it may also be that the options are simply limited, and they have difficulties getting out of this position. The impact of age is not much changed. More striking is the clear strengthening of the coefficients of the industry dummy. The coefficients of ROA have become significant, and the others have been strengthened.

As in the full sample, the lagged regressions of this sample have been moved to Appendix 5. The linear terms in the first models are still similar and as before, significance is lacking. An inverted U-shaped relationship was found in both full models of ROA and EPS, but they are no longer significant. After two years, the U-shaped relationship of profit margin did become significant and slightly stronger with a negative linear term of $\beta = -0.274^{**}$ and a positive squared term of $\beta = 0.250^{*}$. Overall, the results from the low R&D-intensive sample are a lot weaker due to missing significance. This can be explained by the fact that companies in this sample are much less dependent on investments in R&D. For these companies, R&D is not a tool to distinguish themselves. This is reflected in their R&D intensity percentages and subsequently in the regression results.

The changing impact of firm size and leverage continues with the use of a lag. As in the full sample and this subsample without lag, the only thing that can be said conclusively about age is that there is a positive relationship with EPS. In both tables, the strong effect of the industry dummy decreased. It is still negative and significant for profit margin and positive and significant for EPS. This suggests that service companies have a higher profit margin while manufacturing companies have a higher EPS.

5.3.3 High R&D-Intensive Sample Results

Table 8: OLS Results High R&D-Intensive Sample

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	-1.296	-3.171*	-8.541***	-12.91***	-4.893***	-4.733***
	(-0.79)	(-1.73)	(-4.24)	(-5.75)	(-8.56)	(-7.41)
R&D Intensity	-0.027	0.194**	-0.049**	0.366***	-0.105***	-0.157*
	(-1.34)	(1.97)	(-2.49)	(3.76)	(-5.52)	(-1.68)
R&D Intensity²		-0.228**		-0.429***		0.053
		(-2.29)		(-4.36)		(0.56)
Log (Size)	0.248***	0.245***	0.362***	0.356***	0.345***	0.346***
	(11.18)	(10.99)	(16.50)	(16.20)	(16.37)	(16.36)
Log (Age)	0.017	0.013	-0.017	-0.025	0.105***	0.106***
	(0.85)	(0.64)	(-0.88)	(-1.26)	(5.65)	(5.67)
Log (Leverage)	-0.346***	-0.345***	-0.322***	-0.320***	-0.197***	-0.197***
	(-17.66)	(-17.61)	(-16.62)	(-16.57)	(-10.57)	(-10.58)
Industry_Control	-0.018	-0.024	-0.057***	-0.068***	0.048***	0.050***
	(-0.95)	(-1.24)	(-2.98)	(-3.53)	(2.64)	(2.69)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	15.6%	15.7%	17.5%	18.0%	24.0%	23.9%
N	2843	2843	2843	2843	2843	2843

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total assets. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 8 contains the results of the OLS regression of the subsample containing all firms with an average R&D intensity above 1%. R&D plays the most prominent role in these companies, which also makes these results very meaningful. It is a relatively large sample with 2843 observations. The impact of this division in firms can be seen immediately when analyzing the coefficients of the independent variables. The observed β of -0.049** of profit margin and the β of -0.105*** of EPS are very convincing. Accordingly, an increase in R&D intensity will result in a decrease in these 2 firm performance measurements. Results that contradict the earlier formulated hypothesis 1. While R&D investments have no impact on low R&D firms, this is different for firms with a higher level of R&D. The coefficients are significantly and convincingly negative, which means that it does not reward to invest such percentages in R&D. According to Vithessonthi & Racela (2016), high R&D companies are frequently investing in large R&D projects that span a longer period and therefore carry more risk. This risk implies that companies invest large amounts of capital that are not recovered in the end. Something consistent with the results above. Low R&D companies prefer shorter projects and quickly benefit from the results, which could make R&D profitable for them.

When adding the quadratic terms, an inverted U-shaped relationship is again observed for ROA and profit margin. Both with a turning point of 1.53%. The negative relationships in the first models can be explained since most companies in this subsample have an average R&D intensity above this threshold. Above this percentage, it is no longer profitable to invest in R&D, resulting in a negative coefficient. The full model of EPS shows a negative significant linear term (β = -0.157*) along with an insignificant and low positive squared term (β = 0.053). It indicates a U-shaped relationship, but the squared term does not provide a convincing threshold. It looks like a negative relationship that is slowly weakening in strength. The implementation of lagged variables can possibly shed some more light on this.

The difference in the impact of firm size is not that big compared to the low R&D-Intensive subsample. This is probably caused by the fact that the sample size is higher and deviates less from the full sample. In the other subsample, it had dropped significantly. Here it is slightly higher compared to the full sample, which is logical. The impact of leverage is almost the same and has decreased somewhat. However, the impact of leverage on firm performance is still strong in a negative way. The coefficients of the industry dummy are somewhat more strongly downscaled. This is consistent with the other subsample, where the opposite pattern was seen, and all coefficients were significant. The adjusted R^2 rates are similar to those in the full sample and peak here because of the higher reliance on R&D.

The regressions including the lagged dependent variables can be found in Appendix 5. One thing that immediately stands out is that the negative coefficients of R&D intensity in the first models increase during the lags. The increases are not strong, but it can be concluded that the R&D investments have a small effect on the firm performance. After 2 years, R&D investments only hurt the EPS of companies. The same applies to the full models. The patterns and relationships that were found remain the same, but it mainly loses strength and significance. It is only the inverted U-shaped relationship of profit margin that is significant and robust after 2 years. So, the implementation of lagged variables does create a lagged effect. The control variables hardly change after the use of a lag. They go up or down slightly, but no conclusions can be drawn from this, so it will not be discussed further.

The results of the high R&D-intensive subsample provide clear evidence for a negative relationship between R&D and firm performance. Evidence that claims investing in R&D is not profitable for companies. In this case, evidence that would reject hypothesis 1. However, the evidence is slightly weakened after applying the lag. There is clear evidence that indicates a curvilinear relationship between R&D and profit margin as well as between R&D and ROA. This would also reject hypothesis 2.

5.4 Results Robustness Tests

5.4.1 Robustness Tests Alternative Independent Variable

Table 9: OLS Results Robustness Test Full Sample

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	4.946***	4.090***	-4.659***	-5.659***	-5.254***	-5.495***
	(4.42)	(3.68)	(-3.31)	(-4.03)	(-12.04)	(-12.61)
R&D Intensity	-0.112***	0.348***	-0.034**	0.391***	-0.051***	0.268***
	(-6.70)	(6.24)	(-2.06)	(7.02)	(-3.18)	(4.98)
R&D Intensity²		-0.485***		-0.448***		-0.336***
		(-8.64)		(-8.00)		(-6.21)
Log (Size)	0.168***	0.169***	0.314***	0.315***	0.312***	0.313***
	(9.60)	(9.74)	(17.94)	(18.13)	(18.51)	(18.64)
Log (Age)	0.010	-0.009	-0.005	-0.023	0.121***	0.108***
	(0.59)	(-0.57)	(-0.32)	(-1.39)	(7.73)	(6.86)
Log (Leverage)	-0.381***	-0.392***	-0.326***	-0.335***	-0.205***	-0.213***
	(-22.90)	(-23.69)	(-19.59)	(-20.29)	(-12.83)	(-13.32)
Industry_Control	-0.016	-0.037**	-0.079***	-0.099***	0.092***	0.077***
	(-1.02)	(-2.45)	(-5.17)	(-6.47)	(6.29)	(5.20)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	15.6%	17.1%	16.0%	17.3%	22.0%	22.7%
N	3955	3955	3955	3955	3955	3955

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total sales. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

A few robustness tests were conducted to see if this would confirm previous results or possibly lead to new insights. The first one was done by calculating the R&D intensity differently. In this robustness test, the R&D intensity is calculated by dividing the total R&D expenditures by total sales. This is a well-known method in the existing literature (Booltink & Saka-Helmhout, 2018). Table 9 shows the results of the OLS regression of the full sample where the R&D intensity is based on sales. The main thing that has changed is the significance of the R&D intensity in the first models. These were also negative with the other R&D calculation but are now all significant. There is a clear negative impact of R&D investments on firm performance. It provides strong evidence to reject hypothesis 2. Especially, the ROA (β = -0.112***) is very powerful. Probably caused by the fact that ROA is no longer analyzed by the R&D based on the same assets, but by the R&D in sales. As in Table 6, all 3 dependent variables show a significant inverted U-shaped relationship. The mean threshold of 1.49% is slightly lower than previously, which is in line with the reduced R&D intensity in the first models. As expected, the control variables show few notable differences.

Table 10: OLS Results Robustness Test Low R&D-Intensive Sample

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	14.38***	13.52***	7.436***	6.061***	-2.541**	-3.029***
	(9.28)	(8.58)	(3.76)	(3.02)	(-2.29)	(-2.68)
R&D Intensity	0.008	0.253***	0.070**	0.386***	0.098***	0.290***
	(0.27)	(2.78)	(2.43)	(4.11)	(3.51)	(3.20)
R&D Intensity²		-0.250***		-0.322**		-0.196**
		(-2.84)		(-3.53)		(-2.23)
Log (Size)	-0.049	-0.044	0.134***	0.140***	0.165***	0.168***
	(-1.60)	(-1.45)	(4.22)	(4.42)	(5.40)	(5.52)
Log (Age)	-0.042	-0.037	-0.026	-0.020	0.036	0.040
	(-1.49)	(-1.32)	(-0.90)	(-0.69)	(1.29)	(1.43)
Log (Leverage)	-0.417***	-0.421***	-0.312***	-0.318***	-0.197***	-0.201***
	(-14.09)	(-14.27)	(-10.18)	(-10.41)	(-6.68)	(-6.80)
Industry_Control	-0.028	-0.047	-0.215***	-0.238***	0.114***	0.100***
	(-1.01)	(-1.61)	(-7.34)	(-7.97)	(4.05)	(3.47)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	21.3%	21.8%	15.7%	16.5%	21.6%	21.9%
N	1211	1211	1211	1211	1211	1211

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total sales. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 10 shows the results of the robustness test of the low R&D-intensive subsample. The number of firm-year observations increased slightly because the distribution is now based on a different R&D intensity. This R&D intensity based on sales leads to some small changes in the sample sizes. The number of observations in this sample is 1211. The findings of this robustness test are similar to the results from Table 7. However, in this case, all the linear terms in the first models are positive. The other calculation method strengthens the results, which leads to significant coefficients for profit margin (β = 0.070**) and EPS (β = 0.098***). An R&D investment and consequently an increase in R&D intensity would also lead to an increase in profit margin and EPS. This provides evidence for a positive relationship between R&D intensity and firm performance. With this, these variables provide uncommon support for hypothesis 1. An inverted U-shaped relationship was found in all 3 dependent variables. Noteworthy, since there were signs of a regular U-shaped relationship for profit margin in the first test with this sample. However, these were not very strong. The turning points are higher than previously observed, namely 1.66%, 1.82%, and 2.10%. However, this is consistent with this subsample and the positive linear terms in the first models. The control variables show some changes, such as the reduced impact of size and age, although they are not very influential.

Table 11: OLS Results Robustness Test High R&D-Intensive Sample

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	3.591**	-4.579**	-6.545***	-18.17***	-5.728***	-6.869***
	(2.22)	(-2.52)	(-3.21)	(-7.97)	(-12.21)	(-12.87)
R&D Intensity	-0.152***	0.727***	-0.101***	0.887***	-0.073***	0.326***
	(-7.58)	(7.56)	(-5.06)	(9.31)	(-3.85)	(3.56)
R&D Intensity²		-0.900***		-1.012***		-0.409***
		(-9.34)		(-10.60)		(-4.44)
Log (Size)	0.259***	0.257***	0.386***	0.383***	0.421***	0.420***
	(11.61)	(11.69)	(17.37)	(17.60)	(19.99)	(20.01)
Log (Age)	-0.028	-0.041**	-0.049**	-0.063***	0.115***	0.109***
	(-1.39)	(-2.03)	(-2.44)	(-3.18)	(5.96)	(5.68)
Log (Leverage)	-0.385***	-0.391***	-0.347***	-0.354***	-0.208***	-0.211***
	(-19.37)	(-19.99)	(-17.55)	(-18.27)	(-11.10)	(-11.28)
Industry_Control	-0.033*	-0.038**	-0.065***	-0.071***	0.032*	0.030
	(-1.71)	(-2.01)	(-3.40)	(-3.77)	(1.76)	(1.64)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	18.1%	20.6%	18.9%	22.1%	27.1%	27.6%
N	2744	2744	2744	2744	2744	2744

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total sales. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 11 shows the results of the high R&D-intensive subsample with the new calculated R&D intensity. The size of this subsample dropped a bit and has a total of 2744 observations. As in the earlier regression in Table 8, all coefficients of R&D intensity are negative, and the sign is clear. Again, the R&D based on sales provides more significance and confirms the previous findings. All dependent variables and thereby firm performance measures would drop if a firm from this sample would increase their R&D. The results are strong enough to reject hypothesis 1. However, it should be kept in mind that this only applies to high R&D firms. The well-known inverted U-shaped relationship can also be found here. However, the very high linear terms and very low squared terms for the ROA and profit margin are remarkable. This indicates a strong increase until the threshold and subsequently a steep decline (Haans et al., 2015). The height of this turning point is not striking with an average of 1.51%. The impact of age is higher than has been observed before, with significant negative relationships at ROA and profit margin. Which in practice means that younger firms are more likely to perform higher on these firm performance measures. The other control variables have not changed much.

5.4.2 Robustness Tests Alternative Dependent Variable

Table 12: OLS Results Robustness Test Full Sample ROE & Tobin's Q

	ROE		**Tobin's Q**	
	1	**2**	**3**	**4**
Constant	-5.619**	-6.909***	1.380***	1.454***
	(-2.08)	(-2.56)	(9.14)	(9.64)
R&D Intensity	-0.023	0.306***	0.266***	-0.009
	(-1.25)	(5.07)	(17.69)	(-0.18)
R&D Intensity²		-0.345***		0.289***
		(-5.72)		(5.84)
Log (Size)	0.228***	0.226***	0.009	0.010
	(11.94)	(11.90)	(0.58)	(0.67)
Log (Age)	0.047***	0.037**	-0.079***	-0.071***
	(2.73)	(2.20)	(-5.64)	(-5.09)
Log (Leverage)	-0.192***	-0.196***	-0.310***	-0.306***
	(-10.84)	(-11.13)	(-21.39)	(-21.20)
Industry_Control	-0.026	-0.051***	-0.083***	-0.062***
	(-1.58)	(-3.06)	(-6.28)	(-4.52)
Year Dummy	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes
Adjusted R²	8.1%	8.9%	38.4%	38.9%
N	3829	3829	3829	3829

*The dependent variables are Return on Equity (ROE) and Tobin's Q. R&D intensity is the ratio of R&D expenditures to total assets. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Another way to do a robustness check is by using other dependent variables. As with the alternative independent variable, the use of an alternative dependent variable can validate previous results as well as lead to possible new findings. Another dependent variable used in the literature is the Return on Equity (Coombs & Bierly, 2006; Yeh et al., 2010). The ROE can be extracted directly from ORBIS and is calculated by dividing the P/L before tax by the total equity of the company. Tobin's Q is also employed as an alternative dependent variable. This is a market-based indicator of the market value of a company (Vithessonthi & Racela, 2016). This can provide different insights since the other variables are mainly focused on the efficiency and operating performance of a firm. Tobin's Q is a ratio that is calculated by dividing the total market capital by the total assets. The R&D intensity based on the assets is used as the independent variable. Furthermore, the same variables were used. The sample size has decreased very slightly because, in the recent period, a few firms dissolved, merged with another firm, or decided to no longer make their information available in the database. The total number of firm-year observations is still 3829, so this minimal difference will not impact the results.

Table 12 shows the full sample results with the ROE and Tobin’s Q as the dependent variables. The results of the ROE in the full model again confirm the inverted U-shaped relationship. The turning point of 1.56% is also not deviating from previous results. The control variables do not show shocking results either. The impact of age is slightly strengthened while the effect of leverage is slightly weakened. Tobin's Q, on the other hand, does provide remarkable and interesting results. The β of 0.266*** provides clear evidence that there is a positive relationship between R&D investment and Tobin's Q. Alternatively, R&D investments increase Tobin's Q and consequently the market value of the firm. A conclusion that had not been established before. Something that can be explained by examining the market value and not the operating performance of a company. Variables such as ROA, profit margin, and ROE, for example, are more focused on this. This difference can be explained by the fact that operating performance is reduced by the high cost of R&D investments. The market value rises because outsiders know that the company is investing in new knowledge and products. It is expected that this will lead to growth of the company. As a result, shares are valued higher which ultimately leads to a higher market value and Tobin’s Q (Vithessonthi & Racela, 2016). The positive quadratic term in the full model supports this finding. There is no inverted U-shape but a clear growth in market value when R&D intensity increases.

The results of the control variables also differ from the previously used dependent variables. For example, the size has no impact on Tobin's Q where this impact was always positive and significant before. The negative coefficients of age suggest that younger firms have higher Tobin's Q. An explanation for this could be that smaller firms have fewer assets yet, which increases Tobin's Q. Similarly, service companies would have a benefit when it is about Tobin's Q. The high adjusted R-squared values of 38.4% and 38.9% are also remarkable and noteworthy.

Table 13: OLS Results Robustness Test Full Sample ROE & Tobin's Q

	ROE		Tobin's Q	
	1	2	3	4
Constant	0.694	-1.657	1.684***	1.719***
	(0.27)	(-0.65)	(11.51)	(11.70)
R&D Intensity	-0.120***	0.395***	0.236***	0.121**
	(-6.79)	(6.71)	(16.09)	(2.45)
R&D Intensity²		-0.543***		0.121**
		(-9.17)		(2.44)
Log (Size)	0.216***	0.218***	-0.023	-0.024
	(11.62)	(11.88)	(-1.51)	(-1.55)
Log (Age)	0.027	0.007	-0.068***	-0.064***
	(1.58)	(0.42)	(-4.82)	(-4.48)
Log (Leverage)	-0.214***	-0.226***	-0.314***	-0.311***
	(-12.12)	(-12.93)	(-21.48)	(-21.24)
Industry_Control	-0.030*	-0.055***	-0.086***	-0.080***
	(-1.85)	(-3.38)	(-6.42)	(-5.93)
Year Dummy	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes
Adjusted R²	9.2%	11.1%	37.6%	37.7%
N	3829	3829	3829	3829

*The dependent variables are Return on Equity (ROE) and Tobin's Q. R&D intensity is the ratio of R&D expenditures to total sales. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 13 shows the results of the same robustness test but using the alternative independent variable. The negative coefficient (β = -0.120***) validates previous results and shows a negative relationship between R&D investments and ROE. As in Table 9, it is confirmed that this way of calculating R&D intensity gives the results some extra power. The inverted U-shaped relation also remains with this variable. However, this time with a lower threshold of only 1.44%. In summary, the test with the ROE confirms the previous results, stating that an increase in R&D intensity would lead to a decrease in firm performance. This is contrary to hypothesis 1, which states that there would be a positive relationship between R&D investment and firm performance. As a result, the evidence to reject hypothesis 1 becomes stronger and stronger.

The results in models (3) and (4) support the conclusions that were drawn based on the results from Table 12. Namely, that there is a positive relationship between R&D investments and the market value of a company. The full model shows two positive and significant R&D intensity coefficients which indicate a positive and linear relationship between R&D investments and Tobin's Q. Therefore, there would be no threshold or turning point in this case.

5.4.3 Robustness Tests Exclusion High R&D Industries

Table 14: OLS Results Robustness Test Sample Without High R&D Industries

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	4.647***	4.521***	-0.579	-0.701	-4.730***	-4.780***
	(4.11)	(4.02)	(-0.43)	(-0.52)	(-8.28)	(-8.39)
R&D Intensity	-0.014	0.264***	-0.044**	0.182***	0.023	0.242***
	(-0.75)	(4.53)	(-2.30)	(3.08)	(1.21)	(4.20)
R&D Intensity²		-0.297***		-0.242***		-0.234***
		(-5.05)		(-4.06)		(-4.02)
Log (Size)	0.177***	0.169***	0.302***	0.295***	0.285***	0.279***
	(8.60)	(8.20)	(14.50)	(14.16)	(14.00)	(13.67)
Log (Age)	-0.002	-0.008	-0.016	-0.021	0.088***	0.084***
	(-0.13)	(-0.43)	(-0.85)	(-1.10)	(4.82)	(4.59)
Log (Leverage)	-0.429***	-0.431***	-0.360***	-0.362***	-0.210***	-0.211***
	(-22.89)	(-23.09)	(-19.00)	(-19.14)	(-11.32)	(-11.43)
Industry_Control	0.041**	0.008	-0.076***	-0.102***	0.124***	0.098***
	(2.28)	(0.42)	(-4.20)	(-5.34)	(7.05)	(5.26)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	19.9%	20.6%	18.0%	18.4%	21.5%	22.0%
N	2836	2836	2836	2836	2836	2836

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total assets. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

An indication of whether a company invests a lot or little in R&D can also be made based on their industry. According to OECD (2016), there are several industries, based on NACE. Rev 2 codes, which can be seen as high R&D industries. This concerns the following industries: Air and spacecraft and related machinery (303), Computer, electronic and optical products (26), Pharmaceuticals (21), Scientific research and development (72), and Software publishing (582). When checking the data set, especially the pharmaceutical companies stand out because they are present in large numbers and cause outliers. As an additional robustness test, an OLS regression was performed where these five industries were excluded from the sample. This is to see how influential these industries are and whether this would lead to different results. After removing these five industries, there are more than 1000 observations less, and the sample consists of 2836 firm-year observations. The 5 sectors are thus well represented when putting it in relation to the total of 158 different industries in the full sample. Adding to this, the fact that the ranges after winsorizing are strongly reduced compared to the full sample shows how influential these five industries are.

The results do not reveal a lot of surprising new facts. It is most straightforward to compare the results with those in Table 6 because this was the full sample, and the same R&D intensity was used. When comparing the tables, the directions of almost all linear and quadratic terms are the same. Again, an inverted U-shaped relationship is found in all 3 full models. The average turning point is 1.56%, which is very similar to the 1.54% found in the sample including the high R&D industries. Additionally, the linear terms in the first models reveal some differences. The profit margin is the only linear term in the first models that is significant. The negative coefficient (β = -0.044**) indicates a negative relationship between R&D investments and the profit margin of firms from the remaining industries. In Table 6, which included the high R&D industries, there were no significant linear terms found. The lower variance among the variables in this sample could explain the additional significance. The control variables reveal no unexpected results. The influence of firm size and leverage is evident as always, and the influence of firm age and industry are mixed. These mixed results show the same directions as usual. The adjusted R^2 rates are quite high when compared to Table 6.

Table 15: OLS Results Robustness Test Sample Without High R&D Industries

	ROA		**Profit Margin**		**EPS**	
	1	**2**	**1**	**2**	**1**	**2**
Constant	5.746***	5.821***	-1.857	-1.787	-4.498***	-4.473***
	(5.25)	(5.35)	(-1.40)	(-1.35)	(-8.11)	(-8.08)
R&D Intensity	-0.065***	0.218***	-0.004	0.217***	0.006	0.187***
	(-3.51)	(3.85)	(-0.22)	(3.77)	(0.34)	(3.32)
R&D Intensity²		-0.306***		-0.239***		-0.195***
		(-5.28)		(-4.06)		(-3.40)
Log (Size)	0.168***	0.156***	0.314***	0.305***	0.280***	0.272***
	(8.34)	(7.76)	(15.37)	(14.88)	(14.02)	(13.59)
Log (Age)	-0.006	-0.014	-0.015	-0.022	0.088***	0.083***
	(-0.30)	(-0.77)	(-0.81)	(-1.16)	(4.80)	(4.50)
Log (Leverage)	-0.434***	-0.439***	-0.355***	-0.359***	-0.212***	-0.215***
	(-23.27)	(-23.62)	(-18.76)	(-18.99)	(-11.45)	(-11.63)
Industry_Control	0.033*	0.003	-0.072***	-0.095***	0.123***	0.104***
	(1.85)	(0.15)	(-3.96)	(-5.02)	(6.94)	(5.58)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	20.3%	21.0%	17.8%	18.3%	21.5%	21.8%
N	2836	2836	2836	2836	2836	2836

*The dependent variables are Return on Assets (ROA), Profit Margin, and Earnings per Share (EPS). R&D intensity is the ratio of R&D expenditures to total sales. Log (Size), Log (Age), and Log (Leverage) are the natural logarithms of the firms' total assets, number of years since incorporation, and debt to equity ratio. Industry_Control is a dummy variable, with a value of 1 for manufacturing companies. ***, **, * denote significance at the 1%, 5% and 10% levels. T-values are reported in parentheses.*

Table 15 contains the results of the sample without high R&D industries, but with the alternative independent variable. So, the independent variable is based on sales. Again, there is one significant linear term and this one is negative as well. It is the R&D intensity coefficient of ROA in the first model (β = -0.065***). Based on this independent variable, there is a negative relationship between R&D investment and the ROA of companies. As seen earlier, the use of the alternative variable gives much stronger results on ROA. The full models are also not affected after changing the independent variable. Three significant inverted U-shaped relationships were found with an average threshold of 1.54%. The coefficients of the control variables are also almost equal to the results from Table 14.

The results found after excluding the five industries are not very surprising. In both situations, there is only one significant R&D coefficient observed. This one is negative in both cases, a pattern that is familiar now. In practice, this means that an increase in R&D intensity leads to reduced firm performance. Excluding high R&D industries could have caused more positive relationships, but the impact is limited. In addition, the fact that there is a non-linear relationship between investment in R&D and firm performance is again confirmed. This is compelling evidence to reject hypothesis 2.

6. Conclusion

6.1 Conclusion

There is a lot of existing research available about the effect of investing in R&D on firm performance. However, the results of the findings are various and very dependent on the circumstances of the study. Because of this ambiguity about the findings, this paper examined this issue in more detail. Using a sample of companies from OECD countries, a comprehensive answer is given to this question. The relationship between investments in Research & Development and the performance of a business is extensively investigated. The data from the study is collected from 472 firms located in 14 well-developed OECD countries and covers the years 2011 to 2019. The R&D intensity is used to see the effect on the ROA, profit margin, EPS, ROE, and Tobin's Q of a company. Due to the existing literature and theories, the established hypotheses expected a positive and linear relationship between R&D and firm performance.

Through the implementation of a lag, the analysis of different subsamples, and the use of several robustness checks, some clear patterns emerged. The most frequent and convincing relationship in both the full sample and robustness tests is the inverted U-shaped relationship. This relationship has been found for each variable and indicates that there is a turning point to where it is profitable to invest in R&D. This turning point varies from sample to sample, where it is slightly higher in the low R&D sample and a bit lower in the high R&D sample. The average turning point is approximately 1.60%. This non-linear relationship means that hypothesis 2 is rejected. The expected positive and significant relationship was also found in a few cases, but this was mainly in the subsample with companies with low R&D intensity. Their R&D intensity is below the identified turning points, which would make it profitable for low R&D companies to invest in R&D. Therefore, there is support for hypothesis 1 when dealing with low R&D firms. There is also a positive linear relationship observed between R&D investments and the market value of a firm. R&D investments are expressed in the firm value before they have been recouped. A much clearer pattern is a negative relationship between firms with high R&D intensity and firm performance. In the full sample where R&D intensity was calculated based on assets, this had been observed for profit margin and EPS. Using the alternative R&D intensity by sales confirmed these results, and all variables showed a strong negative relationship between investing in R&D and its firm performance. Their average R&D intensity is in most cases above the established turnings points, which makes it unprofitable to invest in R&D. Hypothesis 1 is thereby rejected for high R&D firms.

In summary, there is a clear negative relationship between firms with higher R&D intensity and their firm performance. Thus, it is not rewarding for these firms to invest in R&D. Throughout the full sample, an inverted U-shaped relationship emerges. So, up to a certain point, it is profitable to invest in R&D. This is also the area where R&D investments can be a way for companies to distinguish themselves from the competition. Then it reaches a maximum and starts a decline, seen in all negative significant coefficients of the squared term.

6.2 Discussion

The formulated hypotheses expected a positive and linear relationship. The results turned out differently, and there are several explanations and arguments for a different relationship. In this case, an inverted U-shaped relationship across all firms and a negative relationship for mainly high R&D firms. Most arguments apply to both relationships.

A logical explanation is the fact that it can take a long period from the first R&D activity until the new or improved product enters the market. Usually, an R&D investment is not directly visible in the results (Xu et al., 2019). Investing in R&D involves costs and risks, and this causes lower performance especially in the beginning (Vithessonthi & Racela, 2016). An important reason that studies concerning R&D often use a lag. Wang (2009) indicates that this technological uncertainty together with market uncertainty means that R&D investments are not always recouped. It is simply not possible to say that an R&D investment will also lead to commercial success (Mansfield & Wagner, 1975). However, a lag was used in this study, and this did not change the results significantly.

Some other firm-specific factors can make the results of R&D investments uncertain. According to Teece (1992), companies must enter collaborations to stimulate innovation. However, research by Wales et al. (2012) found an inverted U-shaped relationship between a firm's absorptive capacity and its performance. Thus, at a certain point, collaborations are less efficient, and the cost and time of the collaboration overtake the benefits. Firms have to carefully select and manage their collaborations, or it becomes a loss-making R&D project (Booltink & Saka-Helmhout, 2018). Therefore, this should be considered when investing in R&D. Also, because these kinds of projects are often performed in partnerships. Yeh et al. (2010) indicate that other operating activities such as marketing intensity and labour productivity as well as factors as firm size and debt structure, which are included in this analysis, should be considered. When R&D investments do not complement other operating activities, a company cannot absorb and exploit all R&D potential. A threshold arises because the R&D activities are not used optimally.

Another characteristic of a firm that could be influential on R&D success is its human capital. This is manifested in the theory of the growth of the firm, written by Penrose (1959). This theory states that managerial capability is a determining factor in the growth of a company. The Penrose effect implies that the limited availability of managers with firm-specific experience holds back the growth of a firm. Primarily, individuals who are hired externally do not have the experience that is needed for long-term growth. A long period of close interactions between the people and resources in a company is essential for long-term growth. When this does not happen, it can be a limiting factor that prevents growth (Kor et al., 2016). Wang (2009) also found a curvilinear relationship between R&D and firm performance and refers to bounded rationality. Bounded rationality is the physical limit of human capital in the ability to process information and make decisions based upon that information (Wang, 2009). Taken together, the limits of human capital can cause those R&D investments are partially or not earned back, resulting in an inverted U-shaped relationship between R&D and firm performance. In summary, firm-specific factors such as collaborations with other companies, their other operational activities, and human capital can affect the results of R&D investments. Ultimately, this can lead to a reduction in the success of these investments, resulting in an inverted U-shaped relationship or even a negative relationship between R&D intensity and firm performance. A threshold value arises because the potential of the R&D investment is not being fully exploited.

An alternative explanation for the results can be found in the S-curve theory of Foster (1986). The S-shaped relationship was discussed earlier and shows many similarities to the inverted U-shaped relationship. When the company is the first to enter the market with a new product, it can achieve what is known as a first-mover advantage. In this phase, a competitive advantage can be gained quickly which is reflected in the firm performance. At some point, the new technology of the product becomes outdated or there is a competitor with a better product. At this point, a threshold value often arises. Competitors introduce a better product, resulting in a loss of market share (Lieberman & Montgomery, 1988). This is also reflected in the firm performance, which causes the inverted U-shaped relationship. To stay ahead of the competition, there will be a need to reinvest in R&D and the same pattern will occur.

6.3 Limitations and recommendations

As seen in the conclusion, the significant variance in the full sample creates different patterns and results. More conclusive conclusions can be drawn when choosing a more specific sample. For example, specifically SMEs or just large companies. Or when a threshold is used, which states that the average R&D intensity must be at least 3% or 5%. In that case, the differences in firm size and R&D intensity will not be as large as in this sample. By using a more specific sample, the results will likely be less different and consequently more robust. However, through the creation of subsamples, there are still clear patterns formed. This becomes more difficult when using only a full sample with this much variety. However, using a substantial full sample and several subsamples also has advantages. As in this study, using different samples results in interesting findings, such as an inverted U-shaped relationship. This will be more difficult in a sample with many similar companies.

In addition, there are several reasons why representativeness must be considered. As mentioned earlier, Scandinavian countries have a much higher relative presence than Southern European countries. It was discussed earlier that the success of R&D is influenced by country characteristics, such as country governance and investor protection. This should therefore be taken into account. Also, only listed companies are considered, which means that private-owned companies are not included. This was necessary for the earnings per share but affected the sample size. This together ensures that it is not entirely representative of all companies in OECD countries. Still, a good picture can be created because of the large sample size, but it should always be considered when generalizing the results.

The limitations, the results, and the methods lead to some interesting recommendations for further research. The limitations show that the results are not generalizable to listed and private companies from areas outside of Europe. R&D is also rapidly growing in Asia and the United States which makes these interesting areas for further research. The results could be different in which the country characteristics might have a significant impact.

Even when the same sample is used, there are some recommendations for further research. It has been shown more than once that there are differences between the southern European countries and the Scandinavian countries. It was not examined in more detail in this study because the focus was elsewhere. For further research, it can be investigated whether these differences are truly there and what the possible explanations are. In addition, this study only made a distinction between service firms and manufacturing firms. Therefore, there is no classification of different industries. All industries in the sample can be divided into different sub-industries to gain more insight. The average R&D intensities per industry are very diverse and this will undoubtedly have an impact on the results. Exploring this in depth could provide interesting information. Similarly, the principal focus during this research was on the relationship between R&D and the operating performance of firms. The relationship between Tobin's Q, which represents the market value, provided other insights. For further research, it may be interesting to use a similar sample and explore this in more detail. For example, by looking at differences in low R&D firms and high R&D firms in terms of firm value.

Lastly, the entire study was conducted during the COVID-19 pandemic. The collected data from the companies are from the prior period. The impact of this pandemic on business is very large. However, this varies greatly by country, company, and industry. Where some companies have experienced strong growth, other industries were closed for months. After the COVID-19 pandemic, it may be interesting to examine the actual impact of this period. It seems logical that the rapidly growing companies have invested in R&D and that the companies that were affected the most by the pandemic have cut back on these types of expenditures. Conducting post-pandemic research on the effect of R&D on firm performance can provide diverse results.

References

Aboody, D., & Lev, B. (2000). Information Asymmetry, R&D, and Insider Gains. The Journal of Finance, 55(6), 2747–2766. https://doi.org/10.1111/0022-1082.00305

Aggarwal, C. C. (2013). An Introduction to Outlier Analysis. Outlier Analysis, 1–40. https://doi.org/10.1007/978-1-4614-6396-2_1

Agustia, D., Permatasari, Y., Fauzi, H., & Sari, M. N. A. (2020). RESEARCH AND DEVELOPMENT INTENSITY, FIRM PERFORMANCE, AND GREEN PRODUCT INNOVATION. *Journal of Security and Sustainability Issues*, 1039–1049. https://doi.org/10.9770/jssi.2020.9.3(27)

Alam, A., Uddin, M., Yazdifar, H., Shafique, S., & Lartey, T. (2020). R&D investment, firm performance and moderating role of system and safeguard: Evidence from emerging markets. *Journal of Business Research*, 106, 94–105. https://doi.org/10.1016/j.jbusres.2019.09.018

Anagnostopoulou, S. C., & Levis, M. (2008). R&D and performance persistence: Evidence from the United Kingdom. *The International Journal of Accounting*, 43(3), 293–320. https://doi.org/10.1016/j.intacc.2008.06.004

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99–120. https://doi.org/10.1177/014920639101700108

Beck, N. (2011). Of Fixed-Effects and Time-Invariant Variables. *Political Analysis*, 19(2), 119–122. https://doi.org/10.1093/pan/mpr010

Beladi, H., Deng, J., & Hu, M. (2021). Cash flow uncertainty, financial constraints and R&D investment. *International Review of Financial Analysis*, *76*. Published. https://doi.org/10.1016/j.irfa.2021.101785

Bettis, R. A., & Hitt, M. A. (1995). The new competitive landscape. *Strategic Management Journal*, 16(S1), 7–19. https://doi.org/10.1002/smj.4250160915

Bloom, N., Griffith, R., & Van Reenen, J. (2002). Do R&D tax credits work? Evidence from a panel of countries 1979–1997. *Journal of Public Economics*, 85(1), 1–31. https://doi.org/10.1016/s0047-2727(01)00086-x

Booltink, L. W. A., & Saka-Helmhout, A. (2018). The effects of R&D intensity and internationalization on the performance of non-high-tech SMEs. *International Small Business Journal: Researching Entrepreneurship*, 36(1), 81–103. https://doi.org/10.1177/0266242617707566

Brouwer, E., Kleinknecht, A., & Reijnen, J. O. N. (1993). Employment growth and innovation at the firm level. *Journal of Evolutionary Economics*, 3(2), 153–159. https://doi.org/10.1007/bf01213832

Capasso, M., Treibich, T., & Verspagen, B. (2015). The medium-term effect of R&D on firm growth. *Small Business Economics*, 45(1), 39–62. https://doi.org/10.1007/s11187-015-9640-6

Chan, L. K. C., Lakonishok, J., & Sougiannis, T. (2001). The Stock Market Valuation of Research and Development Expenditures. *The Journal of Finance*, 56(6), 2431–2456. https://doi.org/10.1111/0022-1082.00411

Chan, S. H., Martin, J. D., & Kensinger, J. W. (1990). Corporate research and development expenditures and share value. *Journal of Financial Economics*, 26(2), 255–276. https://doi.org/10.1016/0304-405x(90)90005-k

Chen, C.-J., Huang, Y.-F., & Lin, B.-W. (2012). How firms innovate through R&D internationalization? An S-curve hypothesis. *Research Policy*, 41(9), 1544–1554. https://doi.org/10.1016/j.respol.2012.06.008

Chen, J., Jiao, H., & Zhao, X. (2016). A knowledge-based theory of the firm: managing innovation in biotechnology. *Chinese Management Studies*, 10(1), 41–58. https://doi.org/10.1108/cms-11-2015-0273

Chen, M. C., & Li, H. Y. (2018). The effects and economic consequences of cutting R&D tax incentives. China Journal of Accounting Research, 11(4), 367–384. https://doi.org/10.1016/j.cjar.2018.07.003

Chen, T.-, Guo, D.-Q., Chen, H.-M., & Wei, T.-. (2019). Effects of R&D intensity on firm performance in Taiwan's semiconductor industry. *Economic Research-Ekonomska Istraživanja*, 32(1), 2377–2392. https://doi.org/10.1080/1331677x.2019.1642776

Chung, A., & Choi, M. (2017). The Effects Of Business Strategy On The Association Between R&D Expenditure And Future Firm Performance. *Journal of Applied Business Research (JABR)*, 33(5), 1035–1046. https://doi.org/10.19030/jabr.v33i5.10025

Coad, A., & Rao, R. (2008). Innovation and firm growth in high-tech sectors: A quantile regression approach. *Research Policy*, 37(4), 633–648. https://doi.org/10.1016/j.respol.2008.01.003

Coad, A., & Rao, R. (2010). Firm growth and R&D expenditure. *Economics of Innovation and New Technology*, 19(2), 127–145. https://doi.org/10.1080/10438590802472531

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R & D. *The Economic Journal*, 99(397), 569–596. https://doi.org/10.2307/2233763

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128–152. https://doi.org/10.2307/2393553

Coombs, J. E., & Bierly, P. E. (2006). Measuring technological capability and performance. *R and D Managemen*t, 36(4), 421–438. https://doi.org/10.1111/j.1467-9310.2006.00444.x

Cuervo-Cazurra, A., Nieto, M. J., & Rodríguez, A. (2018). The impact of R&D sources on new product development: Sources of funds and the diversity versus control of knowledge debate. *Long Range Planning*, 51(5), 649–665. https://doi.org/10.1016/j.lrp.2017.06.004

Diefenbach, T. (2006). Intangible resources: a categorial system of knowledge and other intangible assets. *Journal of Intellectual Capital*, 7(3), 406–420. https://doi.org/10.1108/14691930610681483

Disatnik, D., & Sivan, L. (2014). The multicollinearity illusion in moderated regression analysis. Marketing Letters, 27(2), 403–408. https://doi.org/10.1007/s11002-014-9339-5

Dumont, M. (2013). The impact of subsidies and fiscal incentives on corporate R&D expenditures in Belgium (2001-2009). *Reflets et Perspectives de La Vie Économique*, LII(1), 69–91. https://doi.org/10.3917/rpve.521.0069

Eberhart, A. C., Maxwell, W. F., & Siddique, A. R. (2004). An Examination of Long-Term Abnormal Stock Returns and Operating Performance Following R&D Increases. *The Journal of Finance*, 59(2), 623–650. https://doi.org/10.1111/j.1540-6261.2004.00644.x

Eurostat. (n.d.). Eurostat indicators on High-tech industry and Knowledge – intensive services Annex 3 – High-tech aggregation by NACE Rev.2. Retrieved from https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an3.pdf

Falk, M. (2012). Quantile estimates of the impact of R&D intensity on firm performance. *Small Business Economics*, 39(1), 19–37. https://doi.org/10.1007/s11187-010-9290-7

Fey, C. F., & Birkinshaw, J. (2005). External Sources of Knowledge, Governance Mode, and R&D Performance. *Journal of Management*, 31(4), 597–621. https://doi.org/10.1177/0149206304272346

Foster, R. N. (1986). *Innovation: The Attacker's Advantage*. London: Summit Books.

Galbreath, J. (2005). Which resources matter the most to firm success? An exploratory study of resource-based theory. *Technovation*, 25(9), 979–987. https://doi.org/10.1016/j.technovation.2004.02.008

Gelman, A. (2005). Analysis of variance—why it is more important than ever. *The Annals of Statistics*, 33(1), 1–53. https://doi.org/10.1214/009053604000001048

Gharbi, S., Sahut, J.-M., & Teulon, F. (2014). R&D investments and high-tech firms' stock return volatility. *Technological Forecasting and Social Change*, 88, 306–312. https://doi.org/10.1016/j.techfore.2013.10.006

Grant, R. M. (1991). The Resource-Based Theory of Competitive Advantage: Implications for Strategy Formulation. *California Management Review*, 33(3), 114–135. https://doi.org/10.2307/41166664

Grant, R. M. (1996a). Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science*, 7(4), 375–387. https://doi.org/10.1287/orsc.7.4.375

Grant, R. M. (1996b). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109–122. https://doi.org/10.1002/smj.4250171110

Green, B. F., & Tukey, J. W. (1960). Complex analyses of variance: General problems. *Psychometrika*, 25(2), 127–152. https://doi.org/10.1007/bf02288577

Griffith, R., Redding, S., & Reenen, J. V. (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *Review of Economics and Statistics*, 86(4), 883–895. https://doi.org/10.1162/0034653043125194

Gui-long, Z., Yi, Z., Kai-hua, C., & Jiang, Y. (2017). The impact of R&D intensity on firm performance in an emerging market: Evidence from China's electronics manufacturing firms. *Asian Journal of Technology Innovation*, 25(1), 41–60. https://doi.org/10.1080/19761597.2017.1302492

Guo, B., Wang, J., & Wei, S. X. (2018). R&D spending, strategic position and firm performance. *Frontiers of Business Research in China*, 12(1), 1–19. https://doi.org/10.1186/s11782-018-0037-7

Haans, R. F. J., Pieters, C., & He, Z. L. (2015). Thinking about U: Theorizing and testing U- and inverted U-shaped relationships in strategy research. *Strategic Management Journal*, 37(7), 1177–1195. https://doi.org/10.1002/smj.2399

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50(4), 1029. https://doi.org/10.2307/1912775

Hartmann, G. C., Myers, M. B. and Rosenbloom, R. S. (2006) Planning your firm's R&D investment, *Research Technology Management*, 25–36.

Hillier, D., Pindado, J., Queiroz, V., & Torre, C. (2010). The impact of country-level corporate governance on research and development. *Journal of International Business Studies*, 42(1), 76–98. https://doi.org/10.1057/jibs.2010.46

Hitt, M. A., Bierman, L., Shimizu, K., & Kochhar, R. (2001). Direct and Moderating Effects of Human Capital on Strategy and Performance in Professional Service Firms: A Resource-Based Perspective. *Academy of Management Journal*, 44(1), 13–28. https://doi.org/10.5465/3069334

Ho, Y. K., Keh, H. T., & Ong, J. M. (2005). The Effects of R&D and Advertising on Firm Value: An Examination of Manufacturing and Nonmanufacturing Firms. *IEEE Transactions on Engineering Management*, 52(1), 3–14. https://doi.org/10.1109/tem.2004.839943

Hölzl, W. (2009). Is the R&D behaviour of fast-growing SMEs different? Evidence from CIS III data for 16 countries. *Small Business Economics*, 33(1), 59–75. https://doi.org/10.1007/s11187-009-9182-x

Hottenrott, H., & Lopes-Bento, C. (2016). R&D Partnerships and Innovation Performance: Can There Be too Much of a Good Thing? *Journal of Product Innovation Management*, 33(6), 773–794. https://doi.org/10.1111/jpim.12311

Huang, C., & Liu, C. (2005). Exploration for the relationship between innovation, IT and performance. *Journal of Intellectual Capital*, *6*(2), 237–252. https://doi.org/10.1108/14691930510592825

Inkpen, A. C. (1996). Creating Knowledge through Collaboration. *California Management Review*, *39*(1), 123–140. https://doi.org/10.2307/41165879

Jaisinghani, D. (2016). Impact of R&D on profitability in the pharma sector: an empirical study from India. *Journal of Asia Business Studies*, 10(2), 194–210. https://doi.org/10.1108/jabs-03-2015-0031

Koenker, R., & Hallock, K. F. (2001). Quantile Regression. *Journal of Economic Perspectives*, 15(4), 143–156. https://doi.org/10.1257/jep.15.4.143

Kor, Y. Y., Mahoney, J. T., Siemsen, E., & Tan, D. (2016). Penrose's The Theory of the Growth of the Firm: An Exemplar of Engaged Scholarship. *Production and Operations Management*, 25(10), 1727–1744. https://doi.org/10.1111/poms.12572

Lane, P. J., Koka, B. R., & Pathak, S. (2006). The Reification of Absorptive Capacity: A Critical Review and Rejuvenation of the Construct. *Academy of Management Review*, 31(4), 833–863. https://doi.org/10.5465/amr.2006.22527456

Lee, J. W. (2020). Lagged Effects of R&D Investment on Corporate Market Value: Evidence from Manufacturing Firms Listed in Chinese Stock Markets. *The Journal of Asian Finance, Economics and Business*, 7(8), 69–76. https://doi.org/10.13106/jafeb.2020.vol7.no8.069

Li, X. (2012). R&D Intensity and firm performance: Evidence from Chinese manufacturing firms. *2012 IEEE International Conference on Management of Innovation & Technology (ICMIT)*, 45–50. https://doi.org/10.1109/icmit.2012.6225777

Li, W., & Du, J. (2016). Tax incentives, adjustment costs, and R&D investment in China. *China Journal of Accounting Studies*, *4*(4), 433–455. https://doi.org/10.1080/21697213.2016.1252088

Liao, Z., & Cheung, M. T. (2002). Do competitive strategies drive R&D? *The Journal of High Technology Management Research*, *13*(2), 143–156. https://doi.org/10.1016/s1047-8310(02)00052-4

Lieberman, M. B., & Montgomery, D. B. (1988). First-Mover Advantages. *Strategic Management Journal*, 9, 41–58. Retrieved from https://www.jstor.org/stable/2486211

Lin, B.-W., Lee, Y., & Hung, S.-C. (2006). R&D intensity and commercialization orientation effects on financial performance. *Journal of Business Research*, 59(6), 679–685. https://doi.org/10.1016/j.jbusres.2006.01.002

Lin, C., Wu, Y.-J., Chang, C. C., Wang, W., & Lee, C.-Y. (2012). The alliance innovation performance of R&D alliances—the absorptive capacity perspective. *Technovation*, 32(5), 282–292. https://doi.org/10.1016/j.technovation.2012.01.004

Lind, J. T., & Mehlum, H. (2010). With or Without U? The Appropriate Test for a U-Shaped Relationship*. *Oxford Bulletin of Economics and Statistics*, *72*(1), 109–118. https://doi.org/10.1111/j.1468-0084.2009.00569.x

Lome, O., Heggeseth, A. G., & Moen, Ø. (2016). The effect of R&D on performance: Do R&D-intensive firms handle a financial crisis better? *The Journal of High Technology Management Research*, 27(1), 65–77. https://doi.org/10.1016/j.hitech.2016.04.006

Mansfield, E., & Wagner, S. (1975). Organizational and Strategic Factors Associated with Probabilities of Success in Industrial R & D. *The Journal of Business*, *48*(2), 179–198. https://doi.org/10.1086/295734

Montresor, S., & Vezzani, A. (2016). Intangible investments and innovation propensity: Evidence from the Innobarometer 2013. *Industry and Innovation*, 23(4), 331–352. https://doi.org/10.1080/13662716.2016.1151770

Nickerson, J. A., & Zenger, T. R. (2004). A Knowledge-Based Theory of the Firm—The Problem-Solving Perspective. *Organization Science*, 15(6), 617–632. https://doi.org/10.1287/orsc.1040.0093

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), 14–37. https://doi.org/10.1287/orsc.5.1.14

OECD, Galindo-Rueda, F., & Verger, F. (2016). *OECD Taxonomy of Economic Activities Based on R&D Intensity*. https://doi.org/10.1787/18151965

OECD. (2020). Gross domestic spending on R&D. *OECD Data*. https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm

Paula, F., & Silva, J. (2018). The impact of alliances and internal R&D on the firm's innovation and financial performance. *Brazilian Business Review*, 15(6), 533–550. https://doi.org/10.15728/bbr.2018.15.6.2

Penrose, E. T. (1959). *The Theory of the Growth of the Firm.* New York, United States: John Wiley and Sons.

Pervan, M., & Višić, J. (2012). INFLUENCE OF FIRM SIZE ON ITS BUSINESS SUCCESS. *Croatian Operational Research Review (CRORR)*, 3, 213–223.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal*, 14(3), 179–191. https://doi.org/10.1002/smj.4250140303

Pindado, J., de Queiroz, V., & de la Torre, C. (2015). How do country-level governance characteristics impact the relationship between R&D and firm value? *R&D Management*, 45(5), 515–526. https://doi.org/10.1111/radm.12115

Poldahl, A. (2011). The Two Faces of R&D: Does Firm Absorptive Capacity Matter? *Journal of Industry, Competition and Trade*, 12(2), 221–237. https://doi.org/10.1007/s10842-010-0094-x

Roodman, D. (2009). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal: Promoting Communications on Statistics and Stata*, 9(1), 86–136. https://doi.org/10.1177/1536867x0900900106

Segarra, A., & Teruel, M. (2014). High-growth firms and innovation: an empirical analysis for Spanish firms. *Small Business Economics*, 43(4), 805–821. https://doi.org/10.1007/s11187-014-9563-7

Seo, H. S., & Kim, Y. J. (2020). INTANGIBLE ASSETS INVESTMENT AND FIRMS' PERFORMANCE: EVIDENCE FROM SMALL AND MEDIUM-SIZED ENTERPRISES IN KOREA. *Journal of Business Economics and Management*, 21(2), 421–445. https://doi.org/10.3846/jbem.2020.12022

Sharma, C. (2012). R&D and firm performance: evidence from the Indian pharmaceutical industry. *Journal of the Asia Pacific Economy*, 17(2), 332–342. https://doi.org/10.1080/13547860.2012.668094

Sher, P. J., & Yang, P. Y. (2005). The effects of innovative capabilities and R&D clustering on firm performance: the evidence of Taiwan's semiconductor industry. *Technovation*, 25(1), 33–43. https://doi.org/10.1016/s0166-4972(03)00068-3

Sohn, D. -. W., Hur, W., & Kim, H. J. (2010). Effects of R&D and patents on the financial performance of Korean venture firms. *Asian Journal of Technology Innovation*, 18(2), 169–185. https://doi.org/10.1080/19761597.2010.9668697

SONG, T., ZHENG, T., & TONG, L. (2008). An empirical test of the environmental Kuznets curve in China: A panel cointegration approach. *China Economic Review*, *19*(3), 381–392. https://doi.org/10.1016/j.chieco.2007.10.001

Teece, D. J. (1992). Competition, cooperation, and innovation. *Journal of Economic Behavior & Organization*, 18(1), 1–25. https://doi.org/10.1016/0167-2681(92)90050-l

Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319–1350. https://doi.org/10.1002/smj.640

Tukey, J. W. (1962). The Future of Data Analysis. The Annals of Mathematical Statistics, 33(1), 1–67. Retrieved from https://www.jstor.org/stable/2237638

Van Ark, B., Hao, J., Corrado, C., & Hulten, C. (2009). Measuring intangible capital and its contribution to economic growth in Europe (No. 1). *EIB Papers*. https://www.eib.org/attachments/efs/eibpapers/eibpapers_2009_v14_n01_en.pdf#page=64

Veaux, R. D. D., Velleman, P., & Bock, D. E. (2015). *Stats: Data and Models, Global Edition* (4th ed.). Pearson.

Vithessonthi, C., & Racela, O. C. (2016). Short- and long-run effects of internationalization and R&D intensity on firm performance. *Journal of Multinational Financial Management*, 34, 28–45. https://doi.org/10.1016/j.mulfin.2015.12.001

Volberda, H. W., Foss, N. J., & Lyles, M. A. (2010). PERSPECTIVE—Absorbing the Concept of Absorptive Capacity: How to Realize Its Potential in the Organization Field. *Organization Science*, 21(4), 931–951. https://doi.org/10.1287/orsc.1090.0503

Wales, W. J., Parida, V., & Patel, P. C. (2012). Too much of a good thing? Absorptive capacity, firm performance, and the moderating role of entrepreneurial orientation. *Strategic Management Journal*, 34(5), 622–633. https://doi.org/10.1002/smj.2026

Wang, C. -H. (2009). Clarifying the Effects of R&D on Performance: Evidence from the High Technology Industries. *Asia Pacific Management Review*, 16(1), 51–64.

Wang, C.-H., Chang, C.-H., & Shen, G. C. (2015). The effect of inbound open innovation on firm performance: Evidence from high-tech industry. *Technological Forecasting and Social Change*, 99, 222–230. https://doi.org/10.1016/j.techfore.2015.07.006

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171–180. https://doi.org/10.1002/smj.4250050207

White, H., & Macdonald, G. M. (1980). Some Large-Sample Tests for Nonnormality in the Linear Regression Model. *Journal of the American Statistical Association*, 75(369), 16–28. https://doi.org/10.1080/01621459.1980.10477415

Williams, M. N., Grajales, C. A. G., & Kurkiewicz, D. (2013). Assumptions of multiple regression: Correcting two misconceptions. *Practical Assessment, Research & Evaluation*, 18(11), 1–14. https://scholarworks.umass.edu/pare/vol18/iss1/11/

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25–51. https://doi.org/10.1016/j.jeconom.2004.02.005

Xu, J., Liu, F., & Chen, Y.-. (2019). R&D, Advertising and Firms' Financial Performance in South Korea: Does Firm Size Matter? *Sustainability*, 11(14), 3764. https://doi.org/10.3390/su11143764

Xu, J., & Sim, J.-W. (2018). Characteristics of Corporate R&D Investment in Emerging Markets: Evidence from Manufacturing Industry in China and South Korea. *Sustainability*, 10(9), 3002. https://doi.org/10.3390/su10093002

Yang, K.-P., Chiao, Y.-C., & Kuo, C.-C. (2010). The Relationship Between R&D Investment and Firm Profitability Under a Three-Stage Sigmoid Curve Model: Evidence From an Emerging Economy. *IEEE Transactions on Engineering Management*, 57(1), 103–117. https://doi.org/10.1109/tem.2009.2023452

Yeh, M.-L., Chu, H.-P., Sher, P. J., & Chiu, Y.-C. (2010). R&D intensity, firm performance and the identification of the threshold: fresh evidence from the panel threshold regression model. *Applied Economics*, 42(3), 389–401. https://doi.org/10.1080/00036840701604487

Yu, K., Lu, Z., & Stander, J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3), 331–350. https://doi.org/10.1111/1467-9884.00363

Zhang, W. (2015). R&D investment and distress risk. *Journal of Empirical Finance*, 32, 94–114. https://doi.org/10.1016/j.jempfin.2015.03.009

Appendixes

Appendix 1: Empirical Evidence

Relationship	**Method of Measurement**	**Supporting Literature**
Positive	ROA	Sher & Yang (2005)
Positive	ROA	Chen et al. (2019)
Positive	ROA	Agustia et al. (2020)
Positive	ROS	Gui-long et al. (2017)
Positive	Tobin’s Q	Xu & Sim (2018)
Positive	Employment Growth	Capasso et al. (2015)
Positive	Employment / Revenue Growth	Falk (2012)
Positive	Revenue Growth	Lome et al. (2016)
Positive	ROA / ROS	Jaisinghani (2016)
Positive	Profit Margin / Firm Value	Seo & Kim (2020)
Positive	Firm Value	Lee (2020)
Positive	Tobin’s Q	Vithessonthi & Racela (2016)
Positive	ROA / ROE / Tobin’s Q	Guo et al. (2018)
Negative	ROA	Chen et al. (2019)
Negative	ROA	Xu et al. (2019)
Negative	Innovation Performance	Paula & Silva (2018)
Negative	EPS / DCR / ROE / Growth Rate	Li (2012)
Negative	ROA / ROS	Vithessonthi & Racela (2016)
Negative	Employment Growth	Brouwer et al. (1993)
Inverted U-shape	ROA / ROE	Guo et al. (2018)
Inverted U-shape	Revenue / Profit Margin / Firm Value	Booltink & Saka-Helmhout (2018)
Inverted U-shape	ROA / EPS	Yang et al. (2010)
Inverted U-shape	ROA / ROE / Net Profit Growth Rate	Yeh et al. (2010)
Inverted U-shape	ROA / ROS	Huang & Liu (2005)
S-shape	ROA / EPS	Yang et al. (2010)
S-shape	Innovation Performance	Chen et al. (2012)
Inverted S-shape	Firm Growth / ROA / ROE	Wang (2009)

Appendix 2: Definitions & Measurements Variables

Variable	**Measurement**	**Supporting Literature**
Dependent Variables		
ROA	EBIT / Total Assets	Agustia et al., 2020; Xu et al., 2019; Chen et al., 2019; Guo et al., 2018; Jaisinghani, 2016; Vithessonthi & Racela, 2016; Yeh et al., 2010; Yang et al., 2010
Profit Margin	EBIT / Operating Revenues	Seo & Kim, 2020; Guo et al., 2018; Booltink & Saka-Helmhout, 2018
Earnings Per Share	Net Income / Outstanding Shares	Yang et al., 2010; Li, 2012
Independent Variables		
R&D Intensity	Total R&D Expenses / Total Assets	Vithessonthi & Racela, 2016; Lin et al., 2012; Eberhart et al., 2004; Anagnostopoulou & Levis, 2008; Lin et al., 2006
R&D Intensity^2	R&D Intensity * R&D Intensity	Guo et al., 2018; Booltink & Saka-Helmhout, 2018; Wang, 2009; Yang et al., 2010; Chen et al., 2012: Huang & Liu, 2005
Control Variables		
Log (Size)	Natural Logarithm of Total Assets	Xu & Sim, 2018; Seo & Kim, 2020; Xu et al., 2019; Guo et al., 2018; Vithessonthi & Racela, 2016
Log (Age)	Natural Logarithm of (Current Year – Year of Incorporation of the Firm)	Seo & Kim, 2020; Gui-long et al., 2017; Lin et al., 2012; Chen et al., 2012
Log (Leverage)	Natural Logarithm of Gearing Ratio: (Non-current liabilities + Loans) / Shareholders funds	Xu & Sim, 2018; Xu et al., 2019; Chen et al., 2019; Guo et al., 2018; Vithessonthi & Racela, 2016
Industry_Control	Dummy variable based on NACE. Rev 2 codes.	Booltink & Saka-Helmhout, 2018
Year Dummy	Year dummies	Xu & Sim, 2018; Gui-Long et al., 2017; Guo et al., 2018; Vithessonthi & Racela, 2016
Country Dummy	Country dummies	Booltink & Saka-Helmhout, 2018

Appendix 3: Descriptive Statistics Lagged Samples

Descriptive Statistics Full Sample 1-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	3475	0.063	0.281	-0.153	0.079	0.026	0.060	0.098
Profit Margin	3475	0.070	0.299	-0.241	0.099	0.027	0.066	0.118
EPS (€)	3475	1.796	15.62	-2.720	3.136	0.170	0.850	2.284
Independent variables								
R&D Intensity (%)	3475	0.039	0.177	0.000	0.044	0.008	0.022	0.056
Log_RD Intensity	3475	1.258	2.929	0.047	0.813	0.599	1.173	1.883
Log_RD Intensity²	3475	2.243	8.577	0.002	2.364	0.359	1.375	3.544
Control variables								
Size (€ in millions)	3475	8.736	104.7	0.019	20.59	0.172	1.086	5.990
Log (Size)	3475	13.88	18.47	9.854	2.267	12.05	13.90	15.61
Age	3475	66.38	186.0	14.00	48.02	28.00	46.00	99.00
Log (Age)	3475	3.957	5.231	2.708	0.718	3.367	3.850	4.605
Leverage	3475	0.840	3.601	0.010	0.755	0.320	0.670	1.080
Log (Leverage)	3475	0.543	1.526	0.010	0.352	0.278	0.513	0.732
Industry_Control	3475	0.770	1.000	0.000	0.418	1.000	1.000	1.000

Descriptive Statistics Full Sample 2-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	2995	0.063	0.281	-0.153	0.079	0.025	0.061	0.098
Profit Margin	2995	0.069	0.299	-0.241	0.098	0.026	0.066	0.117
EPS (€)	2995	1.744	15.62	-2.720	3.076	0.170	0.832	2.230
Independent variables								
R&D Intensity (%)	2995	0.039	0.177	0.000	0.043	0.008	0.022	0.055
Log_RD Intensity	2995	1.251	2.929	0.047	0.808	0.596	1.170	1.869
Log_RD Intensity²	2995	2.219	8.577	0.002	2.339	0.355	1.369	3.493
Control variables								
Size (€ in millions)	2995	8.864	104.7	0.021	20.69	0.179	1.131	6.108
Log (Size)	2995	13.91	18.47	9.942	2.256	12.09	13.94	15.63
Age	2995	66.66	186.0	14.00	48.05	28.00	47.00	100.0
Log (Age)	2995	3.961	5.231	2.708	0.718	3.367	3.871	4.615
Leverage	2995	0.834	3.601	0.010	0.747	0.320	0.670	1.070
Log (Leverage)	2995	0.541	1.526	0.010	0.348	0.278	0.513	0.728
Industry_Control	2995	0.780	1.000	0.000	0.417	1.000	1.000	1.000

Descriptive Statistics Low R&D-Intensive Sample 1-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	978	0.055	0.197	-0.067	0.057	0.022	0.052	0.085
Profit Margin	978	0.069	0.278	-0.103	0.077	0.024	0.061	0.110
EPS (€)	978	1.478	11.88	-4.741	2.766	0.180	0.869	2.126
Independent variables								
R&D Intensity (%)	978	0.004	0.011	0.000	0.003	0.002	0.003	0.006
Log_RD Intensity	978	0.322	0.748	0.009	0.207	0.144	0.297	0.494
Log_RD Intensity²	978	0.147	0.559	0.000	0.151	0.021	0.088	0.244
Control variables								
Size (€ in millions)	978	13.82	139.1	0.033	29.81	0.321	2.130	9.892
Log (Size)	978	14.49	18.75	10.39	2.215	12.68	14.57	16.11
Age	978	70.82	186.0	14.00	48.24	30.00	56.00	111.3
Log (Age)	978	4.033	5.231	2.708	0.715	3.434	4.043	4.721
Leverage	978	1.054	4.308	0.090	0.927	0.430	0.800	1.250
Log (Leverage)	978	0.641	1.669	0.086	0.377	0.357	0.588	0.811
Industry_Control	978	0.740	1.000	0.000	0.440	0.000	1.000	1.000

Descriptive Statistics Low R&D-Intensive Sample 2-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	844	0.055	0.197	-0.067	0.058	0.022	0.052	0.086
Profit Margin	844	0.069	0.278	-0.103	0.077	0.024	0.061	0.109
EPS (€)	844	1.433	11.88	-4.741	2.737	0.161	0.840	2.054
Independent variables								
R&D Intensity (%)	844	0.004	0.011	0.000	0.003	0.002	0.003	0.006
Log_RD Intensity	844	0.322	0.748	0.009	0.207	0.146	0.299	0.488
Log_RD Intensity²	844	0.147	0.559	0.000	0.151	0.021	0.090	0.238
Control variables								
Size (€ in millions)	844	13.84	137.9	0.034	29.67	0.316	2.235	9.914
Log (Size)	844	14.50	18.74	10.44	2.208	12.66	14.62	16.11
Age	844	70.86	184.4	14.00	48.12	31.00	56.00	113.8
Log (Age)	844	4.034	5.222	2.708	0.714	3.466	4.043	4.743
Leverage	844	1.037	4.308	0.090	0.926	0.410	0.790	1.228
Log (Leverage)	844	0.632	1.669	0.086	0.377	0.344	0.582	0.801
Industry_Control	844	0.740	1.000	0.000	0.440	0.000	1.000	1.000

Descriptive Statistics High R&D-Intensive Sample 1-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	2497	0.065	0.324	-0.202	0.093	0.027	0.064	0.102
Profit Margin	2497	0.068	0.303	-0.329	0.114	0.028	0.068	0.121
EPS (€)	2497	1.926	16.60	-2.088	3.407	0.170	0.842	2.339
Independent variables								
R&D Intensity (%)	2497	0.054	0.204	0.009	0.047	0.019	0.035	0.074
Log_RD Intensity	2497	1.627	3.062	0.615	0.662	1.095	1.508	2.134
Log_RD Intensity²	2497	3.085	9.373	0.379	2.371	1.198	2.274	4.554
Control variables								
Size (€ in millions)	2497	6.838	86.77	0.017	16.38	0.131	0.799	4.670
Log (Size)	2497	13.64	18.28	9.731	2.245	11.79	13.59	15.36
Age	2497	64.66	186.0	15.00	47.81	27.00	41.00	97.00
Log (Age)	2497	3.928	5.231	2.773	0.715	3.332	3.738	4.585
Leverage	2497	0.758	3.244	0.010	0.667	0.285	0.610	1.020
Log (Leverage)	2497	0.504	1.446	0.010	0.333	0.251	0.476	0.703
Industry_Control	2497	0.790	1.000	0.000	0.408	1.000	1.000	1.000

Descriptive Statistics High R&D-Intensive Sample 2-year lag

Variable	N	Mean	Max.	Min.	Std. Dev.	Q1	Median	Q3
Dependent variables								
ROA	2151	0.066	0.324	-0.202	0.093	0.028	0.065	0.103
Profit Margin	2151	0.068	0.303	-0.329	0.113	0.028	0.068	0.121
EPS (€)	2151	1.868	16.60	-2.088	3.339	0.170	0.831	2.281
Independent variables								
R&D Intensity (%)	2151	0.053	0.204	0.009	0.046	0.019	0.035	0.073
Log_RD Intensity	2151	1.618	3.602	0.615	0.658	1.089	1.501	2.119
Log_RD Intensity²	2151	3.049	9.373	0.379	2.341	1.186	2.252	4.491
Control variables								
Size (€ in millions)	2151	6.997	87.06	0.018	16.56	0.138	0.837	4.866
Log (Size)	2151	13.68	18.28	9.772	2.238	11.84	13.64	15.40
Age	2151	65.02	186.0	15.00	47.87	27.00	41.00	97.00
Log (Age)	2151	3.934	5.231	2.773	0.715	3.332	3.738	4.585
Leverage	2151	0.756	3.244	0.010	0.657	0.290	0.620	1.010
Log (Leverage)	2151	0.505	1.446	0.009	0.329	0.255	0.482	0.698
Industry_Control	2151	0.790	1.000	0.000	0.407	1.000	1.000	1.000

Appendix 4: Correlation Matrices Split Samples

Pearson's Correlation Matrix Low R&D-Intensive Sample

	ROA	Profit Margin	EPS	R&D Intensity	R&D Intensity²	Log (Size)	Log (Age)	Leverage	Industry Control
ROA	1								
Profit Margin	0.778**	1							
EPS	0.449**	0.394**	1						
R&D Intensity	0.068*	-0.095**	0.023	1					
R&D Intensity²	0.050	-0.064*	0.006	0.963**	1				
Log (Size)	-0.100**	0.098**	0.170**	-0.317**	-0.286**	1			
Log (Age)	-0.033	-0.041	0.165**	0.022	0.031	-0.002	1		
Log (Leverage)	-0.391**	-0.223**	-0.137**	-0.198**	-0.149**	0.293**	0.071*	1	
Industry_Control	0.037	-0.196**	0.110**	0.341**	0.244**	-0.232**	0.018	-0.260**	1

*** Correlation is significant at the 0.01 level (2-tailed).*
** Correlation is significant at the 0.05 level (2-tailed).*

Pearson's Correlation Matrix High R&D-Intensive Sample

	ROA	Profit Margin	EPS	R&D Intensity	R&D Intensity²	Log (Size)	Log (Age)	Leverage	Industry Control
ROA	1								
Profit Margin	0.905**	1							
EPS	0.281**	0.279**	1						
R&D Intensity	0.026	-0.027	-0.193**	1					
R&D Intensity²	0.015	-0.046*	-0.204**	0.983**	1				
Log (Size)	0.107**	0.218**	0.376**	-0.335**	-0.356**	1			
Log (Age)	0.064**	0.064**	0.265**	-0.250**	-0.275**	0.340**	1		
Log (Leverage)	-0.274**	-0.205**	0.004*	-0.342**	-0.343**	0.348**	0.121**	1	
Industry_Control	0.026*	0.027**	0.182**	-0.240**	-0.269**	0.341**	0.281**	0.152**	1

*** Correlation is significant at the 0.01 level (2-tailed).*
** Correlation is significant at the 0.05 level (2-tailed).*

Appendix 5: OLS Results Lagged Split Samples

OLS Results Full Sample 1-year Lag

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	1.793	1.506	-5.318***	-5.753***	-5.604***	-5.834***
	(1.45)	(1.22)	(-3.47)	(-3.75)	(-12.11)	(-12.64)
R&D Intensity	-0.002	0.175***	-0.011	0.204***	-0.009	0.349***
	(-0.13)	(2.83)	(-0.60)	(3.33)	(-0.51)	(6.00)
R&D Intensity²		-0.186***		-0.225***		-0.376***
		(-3.00)		(-3.68)		(-6.45)
Log (Size)	0.184***	0.184***	0.315***	0.314***	0.322***	0.320***
	(9.54)	(9.50)	(16.48)	(16.46)	(17.63)	(17.63)
Log (Age)	0.021	0.015	-0.004	-0.011	0.128***	0.117***
	(1.19)	(0.87)	(-0.25)	(-0.64)	(7.79)	(7.11)
Log (Leverage)	-0.332***	-0.334***	-0.300***	-0.303***	-0.187***	-0.192***
	(-18.45)	(-18.59)	(-16.90)	(-17.08)	(-11.05)	(-11.39)
Industry_Control	-0.005	-0.019	-0.071***	-0.088***	0.101***	0.073***
	(-0.30)	(-1.09)	(-4.37)	(-5.20)	(6.48)	(4.57)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	13.2%	13.4%	15.1%	15.4%	22.7%	23.6%
N	3475	3475	3475	3475	3475	3475

OLS Results Full Sample 2-year Lag

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	1.984	1.795	-5.087***	-5.407***	-5.413***	-5.623***
	(1.49)	(1.34)	(-3.10)	(-3.29)	(-11.07)	(-11.53)
R&D Intensity	0.011	0.128*	0.005	0.164**	-0.001	0.332***
	(0.58)	(1.91)	(0.24)	(2.46)	(-0.07)	(5.27)
R&D Intensity²		-0.123*		-0.166**		-0.350***
		(-1.82)		(-2.50)		(-5.55)
Log (Size)	0.172***	0.171***	0.303***	0.303***	0.310***	0.308***
	(8.21)	(8.19)	(14.65)	(14.64)	(15.72)	(15.74)
Log (Age)	0.019	0.015	-0.006	-0.011	0.130***	0.120***
	(0.98)	(0.78)	(-0.33)	(-0.59)	(7.33)	(6.74)
Log (Leverage)	-0.306***	-0.307***	-0.277***	-0.280***	-0.169***	-0.174***
	(-15.72)	(-15.80)	(-14.43)	(-14.55)	(-9.22)	(-9.53)
Industry_Control	-0.004	-0.013	-0.068***	-0.080***	0.101***	0.075***
	(-0.20)	(-0.68)	(-3.84)	(-4.37)	(5.99)	(4.32)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	11.9%	12.0%	13.9%	14.1%	22.1%	22.9%
N	2995	2995	2995	2995	2995	2995

OLS Results Low R&D-Intensive Sample 1-year Lag

	ROA		Profit Margin		EPS	
	1	**2**	**1**	**2**	**1**	**2**
Constant	12.56***	12.18***	9.735***	11.03***	-3.170***	-3.219***
	(6.43)	(6.11)	(3.63)	(4.04)	(-3.39)	(-3.37)
R&D Intensity	0.025	0.140	-0.020	-0.311**	0.050	0.080
	(0.74)	(1.12)	(-0.58)	(-2.44)	(1.49)	(0.65)
R&D Intensity²		-0.115		0.289**		-0.030
		(-0.95)		(2.37)		(-0.25)
Log (Size)	-0.003	-0.001	0.119***	0.115***	0.213***	0.213***
	(-0.07)	(-0.02)	(3.14)	(3.03)	(5.78)	(5.78)
Log (Age)	-0.046	-0.046	-0.033	-0.034	0.084***	0.084***
	(-1.42)	(-1.40)	(-1.01)	(-1.05)	(2.63)	(2.64)
Log (Leverage)	-0.387***	-0.385***	-0.307***	-0.312***	-0.207***	-0.206***
	(-11.23)	(-11.14)	(-8.78)	(-8.94)	(-6.10)	(-6.07)
Industry_Control	-0.051	-0.061*	-0.214***	-0.189***	0.091***	0.089**
	(-1.53)	(-1.75)	(-6.29)	(-5.32)	(2.76)	(2.56)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	16.6%	16.6%	14.4%	14.8%	19.3%	19.2%
N	978	978	978	978	978	978

OLS Results Low R&D-Intensive Sample 2-year Lag

	ROA		Profit Margin		EPS	
	1	**2**	**1**	**2**	**1**	**2**
Constant	13.23***	12.57***	10.82***	12.01***	-2.844***	-2.939***
	(6.09)	(5.66)	(3.69)	(4.01)	(-2.81)	(-2.84)
R&D Intensity	0.018	0.205	-0.023	-0.274**	0.022	0.078
	(0.50)	(1.50)	(-0.61)	(-1.99)	(0.60)	(0.58)
R&D Intensity²		-0.185		0.250*		-0.056
		(-1.42)		(1.90)		(-0.44)
Log (Size)	-0.028	-0.024	0.093**	0.088**	0.183***	0.184***
	(-0.67)	(-0.59)	(2.23)	(2.12)	(4.53)	(4.55)
Log (Age)	-0.057	-0.056	-0.041	-0.043	0.084**	0.084**
	(-1.59)	(-1.57)	(-1.15)	(-1.19)	(2.40)	(2.40)
Log (Leverage)	-0.336***	-0.334***	-0.273***	-0.276***	-0.163***	-0.162***
	(-8.86)	(-8.79)	(-7.12)	(-7.21)	(-4.37)	(-4.35)
Industry_Control	-0.040	-0.056	-0.202***	-0.180***	0.107***	0.102***
	(-1.10)	(-1.48)	(-5.48)	(-4.68)	(3.00)	(2.73)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	14.0%	14.1%	12.3%	12.6%	17.2%	17.1%
N	844	844	844	844	844	844

OLS Results High R&D-Intensive Sample 1-year Lag

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	-1.611	-3.000	-9.304***	-13.20***	-5.133***	-4.927***
	(-0.92)	(-1.53)	(-4.37)	(-5.56)	(-8.50)	(-7.30)
R&D Intensity	-0.018	0.148	-0.038*	0.340***	-0.098***	-0.165*
	(-0.84)	(1.39)	(-1.79)	(3.24)	(-4.84)	(-1.65)
R&D Intensity²		-0.171		-0.390***		0.069
		(-1.59)		(-3.68)		(0.69)
Log (Size)	0.251***	0.248***	0.368***	0.362***	0.348***	0.349***
	(10.51)	(10.38)	(15.69)	(15.43)	(15.58)	(15.59)
Log (Age)	0.011	0.008	-0.020	-0.027	0.109***	0.110***
	(0.53)	(0.39)	(-0.96)	(-1.28)	(5.50)	(5.54)
Log (Leverage)	-0.316***	-0.315***	-0.297***	-0.296***	-0.188***	-0.188***
	(-14.97)	(-14.94)	(-14.30)	(-14.26)	(-9.49)	(-9.50)
Industry_Control	-0.020	-0.024	-0.057***	-0.067***	0.051***	0.052***
	(-0.96)	(-1.16)	(-2.81)	(-3.28)	(2.61)	(2.68)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	14.2%	14.3%	16.8%	17.2%	24.6%	24.5%
N	2497	2497	2497	2497	2497	2497

OLS Results High R&D-Intensive Sample 2-year Lag

	ROA		Profit Margin		EPS	
	1	2	1	2	1	2
Constant	-1.755	-2.028	-9.466***	-12.00***	-5.038***	-4.994***
	(-0.93)	(-0.96)	(-4.16)	(-4.72)	(-7.89)	(-6.99)
R&D Intensity	0.002	0.035	-0.017	0.232**	-0.084***	-0.099
	(0.09)	(0.30)	(-0.76)	(2.02)	(-3.84)	(-0.91)
R&D Intensity²		-0.034		-0.256**		0.015
		(-0.29)		(-2.22)		(0.14)
Log (Size)	0.239***	0.239***	0.359***	0.356***	0.338***	0.338***
	(9.31)	(9.26)	(14.18)	(14.01)	(14.03)	(14.00)
Log (Age)	0.012	0.011	-0.020	-0.025	0.114***	0.114***
	(0.52)	(0.49)	(-0.91)	(-1.09)	(5.30)	(5.29)
Log (Leverage)	-0.294***	-0.294***	-0.275***	-0.274***	-0.174***	-0.174***
	(-12.88)	(-12.87)	(-12.22)	(-12.18)	(-8.16)	(-8.15)
Industry_Control	-0.016	-0.017	-0.053**	-0.060***	0.053**	0.054**
	(-0.72)	(-0.75)	(-2.40)	(-2.70)	(2.53)	(2.53)
Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R²	13.3%	13.3%	15.8%	16.0%	24.0%	23.9%
N	2151	2151	2151	2151	2151	2151