Stockholm School of Economics
Department of Accounting and Financial Management
Bachelor Thesis
May 2021

Innovative performance and bankruptcy risk in Swedish private firms

Anna Hedensiö
24479@student.hhs.se

Michaela Jungner
24373@student.hhs.se

Abstract

This study examines how innovative performance impacts the probability of bankruptcy. We use three different measures of innovative performance with the aim of capturing the intensity of innovative input, the quantity of innovative output as well as the innovative efficiency. These are R&D intensity, patent count and the patent-to-R&D ratio, respectively. We formulate a hypothesis for each of the three measures and its association to the probability of bankruptcy. To investigate the respective associations, we perform multiple logistic regressions using an unbalanced panel dataset consisting of Swedish private firms between 2002 and 2019. We find evidence indicating that the quantity of innovative output as well as innovative efficiency are negatively associated with the probability of bankruptcy. However, we do not find any statistically significant association between the intensity of innovative input and the probability of bankruptcy. Our findings are in line with previous studies that suggest that the quantity of innovative output and innovative efficiency are negatively related to bankruptcy risk, although our exact definition of innovative efficiency to the best of our knowledge is unprecedented in the literature on bankruptcy prediction.

Tutor: Ting Dong

Keywords: Bankruptcy prediction; Innovation; R&D spending; Patents; Innovative efficiency

Acknowledgements
We would like to express our gratitude to our tutor Ting Dong, Researcher at Stockholm School of Economics, for her valuable input and guidance throughout the writing of this thesis. We would also like to thank Katerina Hellström, Assistant Professor at Stockholm School of Economics, for providing meaningful insights and advice regarding the writing process. Lastly, we want to express our appreciation to family and friends for their support.

Table of contents

1. Introduction

1.1 Background

Corporate bankruptcy has for long been a major area of study due to all its negative consequences. The significant costs it entails is one of them. The direct costs of a bankruptcy, such as legal and filing fees, on average amount to 11-17% of firm value up to three years prior to bankruptcy (Altman, 1984). Apart from that, there are also indirect costs involved, such as the opportunity cost of the management's time as well as a deteriorated reputation among different stakeholders (Thornburn, 2000). Bankruptcy is not only costly for internal stakeholders but also for external ones. For example, creditors risk not being paid and suppliers as well as customers risk losing a meaningful partnership or not receiving an order. Due to the large costs which a bankruptcy implies, the ability to predict a firm's tendency to go bankrupt is of great interest to both internal and external stakeholders.

Scholars have conducted a vast amount of research on the topic bankruptcy, and several well-known bankruptcy prediction models have been introduced. Ohlson's O-score model and Altman's Z-score model are two examples of such models (Altman, 1968; Ohlson, 1980). These models are generally accounting-based models that include lagging accounting measures, but in more recent decades also models consisting of more market-based measures have been developed, such as Merton's DD model and the CHS model (Merton, 1974; Campbell et al., 2008). However, even these models tend to include historical financial numbers, and thus also lack in their forward-looking nature (Miao et al., 2018). Due to many traditional bankruptcy prediction models being insufficient in their scope, scholars have more recently aimed to incorporate non-financial measures in order to capture other factors which could have a potential impact on bankruptcy. One such factor is innovation, where some scholars have extended the research on bankruptcy prediction by examining how different measures of innovative performance relate to bankruptcy risk. As the traditional models using accounting or market-based financial numbers do not manage to capture valuable information about firms' innovative work (Bai and Tian, 2020), it can be argued that additional variables measuring innovative performance are needed.

There are several reasons as to why different measures of innovation would be of interest to incorporate in bankruptcy prediction models. First and foremost, innovative work infers an interesting trade-off for firms in terms of benefits versus risk. On the one hand, such activities are vital for firms to retain their competitive advantage towards other incumbents as well as potential entrants. In that sense, firms that do not prioritize innovation risk falling behind and experience a deterioration in operating performance in the long run (Christensen et al., 2008). On the other hand, innovation entails a lot of risk since the future outcome is extremely uncertain, which is another important aspect that various scholars emphasize (Eberhart et al., 2008; Buddelmeyer et al., 2010; Ericson and Pakes, 1995). Both these aspects suggest that innovation and bankruptcy risk could be related, though in different ways. The dual nature of innovation enhances the interest of understanding how innovative work impacts the probability of bankruptcy in practice. Moreover, the technological advancement as well as the upward trend in R&D activity over the last decades further support that innovation should be taken into account in modern bankruptcy prediction models (Eisdorfer and Hsu, 2011; Franzen et al., 2007).

Looking into previous research on how innovative performance relates to bankruptcy risk, the literature suggests several different measures for different parts of the innovation process. The two by far most common ones are R&D spending as a measure of innovative input as well as some patent-related measure as a measure of innovative output. As for the patent-related measures, previous studies' results point toward a negative relationship to bankruptcy risk (Eisdorfer and Hsu, 2011; Bai and Tian, 2020). However, previous studies on how R&D spending relates to the probability of bankruptcy are clearly showing contradictory results. Some scholars find a negative relationship (Eberhart et al., 2008; Li et al., 2010), while others find a positive relationship (Ericson and Pakes, 1995; Zhang, 2015). The lack of consensus advocates for further research on the topic. Furthermore, as to the best of our knowledge, no previous study has yet examined how the relationship between innovative input and output relates to bankruptcy risk, which implies an opening to contribute to the literature field by studying that specific topic further. In the light of this, we believe there is a great opportunity to provide valuable insights and contribute to the literature by investigating the relationship between innovative performance and the probability of bankruptcy further.

To investigate the relationship between innovative performance and the probability of bankruptcy, we perform multiple logistic regressions testing three different measures of innovative performance, namely R&D intensity, patent count and the patent-to-R&D ratio. Using an unbalanced panel dataset consisting of Swedish private firms over the years 2002 to 2019, we find that patent count as well as the patent-to-R&D ratio are negatively associated with the probability of bankruptcy on a 1% significance level. However, we find no significant association between R&D intensity and the probability of bankruptcy. These results prove to be robust in the majority of the robustness tests on meaningful significance levels.

1.2 Purpose and question formulation

The purpose of this study is to examine the relationship between innovative performance and the probability of bankruptcy. By using the three measures R&D intensity, patent count and the patent-to-R&D ratio, we aim to capture different aspects of the innovation process in terms of innovative input, output, and efficiency, respectively. The results of our study will hence provide evidence regarding how these measures relate to the probability of bankruptcy. The study aims to address the following research question:

How do firms' innovative performance impact their probability of bankruptcy?

1.3 Hypotheses

We formulate three hypotheses based on the literature review, as will be presented in section 2. All three hypotheses consider the relationship between innovative performance and the probability of bankruptcy, where each hypothesis relates to one of our three measures of innovative performance. The first hypothesis considers the intensity of innovative input. Given that prior research shows mixed results, we formulate our first hypothesis as a null hypothesis. The second hypothesis relates to the quantity of innovative output, while the third hypothesis relates to the innovative efficiency. Both of these hypotheses are directional, expecting negative associations in line with what literature suggests.

1.4 Contributions

With our research, we aim to take a different approach than existing research by examining the relationship between innovative performance and the probability of bankruptcy in the Swedish market specifically. To the best of our knowledge, no previous study has had that specific geographical scope, hence leaving a gap in the literature for us to examine. Furthermore, the available papers on the topic are in some aspects contradictory in their conclusions regarding the connection between innovative performance and bankruptcy likelihood. Therefore, the research area is of high relevance and interest and needs to be explored further. Moreover, we include a measure of innovative performance that aims to capture innovative efficiency, defined as patent count over R&D spending. As far as we are aware, this exact measure has not yet been used in research on bankruptcy prediction, and thus provides new insights.

1.5 Delimitation

The scope of the study is limited to Swedish private limited liability companies during the time period 2002-2019. However, the sample also includes observations from 1998-2001 due to the use of lagged variables in the regressions as well as calculations for variables requiring opening balance figures. The choice of not including listed firms has to do with the fact that their conditions as well as capabilities to innovate are likely to differ compared to private firms, and the reason for not looking at them exclusively is that there are rather few of them in Sweden and even fewer that go bankrupt - which would result in a too small sample. Furthermore, previous papers within the same area of research commonly study listed firms in other markets, which make it relevant and interesting to further examine the relationship in private firms as well.

Another important delimitation to our study is that it only includes firms that apply the cost of sales method in the accounting of their income statements, hence excluding all firms applying the alternative nature of expense method. The reason for excluding firms using the latter accounting method is that when applying it, the R&D expenditures are not reported on a separate line in the income statement, and thus not accessible in the database, which results in severely lacking data for the firms using it. Excluding those firms naturally leads to smaller firms complying with the K2 regulations being excluded too, since such firms are restricted and only allowed to use the nature of expense method (Bolagsverket, 2019). However, the K2 regulations were not allowed to be applied in the accounts of the annual reports until December 31, 2008 (Bokföringsnämnden, 2020). Before that, all firms could choose freely between the two accounting methods according to the Annual Reports Act (1995:1554) (Riksdagen, 2020, 3 kap. 3 § ÅRL). This means that smaller firms that have chosen to comply with the K2 regulations from 2008 and onwards are excluded from our sample but might be part of it prior to 2008 given that they chose the cost of sales method then. From 2008, only firms that comply with the K3 regulations as well as private firms that voluntarily choose to adopt the IFRS Standards appear in our sample, given that they choose to apply the cost of sales method.

1.6 Disposition

The study consists of six sections. Section 2 presents a relevant literature review, including the development of the hypotheses. Section 3 includes a description of the methodology, by outlining the research design as well as presenting the regression models and the variables used in the study. The section also ends with an illustration of the data collection and the sample selection process. Section 4 presents preliminary analysis as well as empirical results of the study. Section 5 further discusses the findings, while section 6 provides a summary of the findings, followed by a presentation of practical contributions, limitations, and suggestions for future research.

2. Theory and literature review

2.1 Summary of section

The following section describes the theory and literature of which we build our study on. Firstly, this section will discuss innovation in terms of its definition and characteristics. After this, a review of bankruptcy prediction theory will be provided. Next part of the section focuses on innovation in relation to bankruptcy risk by presenting previous research within this area of study, where the hypotheses also are presented.

2.2 Innovation

2.2.1 What innovation is and how it can be measured

There is no single, distinct definition of innovation employed in the empirical literature. Instead, various definitions are used. Though many studies have handled the topic, no generally accepted measure or set of measures of innovative performance have been established (Hagedoorn and Cloodt, 2003). Some examples of measures that are being commonly used in previous studies are R&D inputs, patent counts, patent citations and counts of new product announcements, where some studies use only one indicator while others use more. Hagedoorn and Cloodt (2003) provide a broad understanding of innovative performance, seeing it as the achievements in several stages, all the way from coming up with an idea to introducing an actual invention to the market. By emphasizing all stages from R&D to patenting as well as the introduction of new products, this broad definition captures all the commonly used indicators listed above.

R&D spending is a commonly used measure of innovative input (Bai and Tian, 2020; Pandit et al., 2011; Zhang, 2015; Ugur et al., 2016; Hsu et al., 2015) and captures both successful and unsuccessful innovative work. Hopefully, investments in R&D create intangible assets that yield long term payoff (Bai and Tian, 2020). In that way, they relate to the future health of a firm. Under US GAAP, firms are required to directly expense R&D (KPMG, 2021), meaning that the total R&D spending equals the R&D expenditure in the income statement. For Swedish firms, things are more complicated. According to the Annual Reports Act (1995:1554), Swedish firms have the possibility to capitalize development costs (Riksdagen, 2020, 4 kap. 2

§ ÅRL). The K2 regulations restrict this possibility, but firms following both the K3 regulations as well as the IFRS may capitalize under certain circumstances (Skatteverket, 2021). This means that to capture the full R&D spending in those firms, one cannot not only look at the R&D expenditure in the income statement but must also take the investments in the capitalized part into account. Most often, the R&D spending is scaled by some sort of base such as total assets, sales, or staff headcount in order to measure the intensity of it, referred to as R&D intensity.

Input measures such as R&D spending are often complemented by some sort of measure of innovative output. In the empirical literature, there are various ways to define innovative output, such as patents, trademarks, registered designs, and new product announcements (Buddelmeyer et al., 2010; Hagedoorn and Cloodt, 2003). However, the by far most common measure of innovative output used in previous studies is patents (Bai and Tian, 2020; Pandit et al., 2011; Eisdorfer and Hsu, 2011; Hsu et al., 2015). Patents serve as exclusive rights to use certain knowledge in an economy, implying that they help prevent competitors from using resembling technology, and reflect firms' intangible intellectual assets. Hence, patent counts contain valuable information about the output of firms' innovation (Hsu et al., 2015).

Measuring innovative performance by looking at patents has several advantages. To start with, it involves less uncertainty and inefficiency than input measures, such as R&D spending, since patents are realized inventions that affect future financial performance (Eisdorfer and Hsu, 2011). Moreover, information about patents is not subject to accounting distortion or manipulation in the same way as R&D expenditures might be (Hsu et al., 2015). Also, patents can in many ways be considered superior to the alternative innovative output measures mentioned above. For example, the legal threshold for receiving patents is much higher (Buddelmeyer et al., 2010), and hence the measure contains more valuable information. In addition, the screening process performed is often more thorough (Hagedoorn and Cloodt, 2003).

When using patents as a basis for innovative performance measurement, there are alternative measures to choose from other than looking at the raw patent count in terms of patent stock or patents granted in a year. One could also look at other types of measures such as patent citations or patents applied for in a year. Since the distribution of patent value is highly skewed (Buddelmeyer et al., 2010), patent citations can be a good complement to raw patent count since it takes the quality of patents into account (Hagedoorn and Cloodt, 2003). Patent applications, on the other hand, though more certain than R&D spending, is more of an input measure, and can be a good way to indicate the level of a firm's high-risk innovation. This interpretation of patent applications is for example made by Buddelmeyer et al. (2010).

Some previous studies also try to include some kind of measure of the effectiveness or efficiency of firms' innovative work. For example, Bai and Tian (2020) include a measure of firms' effectiveness in R&D investment called R&D productivity, which measures how changes in R&D expenditure influence revenue. Another example is Hirshleifer et al. (2013), who study the relationship between innovative efficiency and stock returns and define innovative efficiency as patents or citations scaled by R&D expenditures. Since investments in R&D may not lead to commercially successful innovation outcomes, and since patents scaled by R&D expenditures tend to fall as the R&D intensity increases (Ugur et al., 2016), the patent-to-R&D ratio can be a good, complementary measure to look at besides the two separate measures in isolation.

2.2.2 Characteristics and implications of innovative work

According to Christensen et al. (2008), innovation investments are vital for firms to keep their competitive advantage towards already existing actors as well as new entrants. The authors emphasize how underinvesting in innovation may result in worsened operating performance in the long run due to for example decreasing demand, declining market share and margin pressure. Thus, investing in innovation can be considered not to have an immediate effect on a firm's performance, but rather be a more strategic investment with a longer time horizon than regular, add-on investments. Furthermore, firms that are more competitive in innovation in terms of strong patent portfolios have shown to gain competitive advantage because the patents provide for a monopoly-like position in the market, making them more financially stable (Hsu et al., 2015). This positive relationship between a firm's future operating performance and its innovative work has been established by several other studies. For instance, Hirshleifer et al. (2013) find robust results suggesting that firms that have higher innovative efficiency have higher future earnings and stock prices. In addition, Pandit et al. (2011) find that more innovative firms, where innovation is measured both in terms of R&D expenses and patent as well as citation counts, have higher future cash flows and earnings.

However, innovative work does not necessarily result in improved financial performance. Innovative work also entails a lot of risk for a firm. Because the future outcome is very uncertain and long-term strategic investments in innovation are based on assumptions which may not turn out as expected, firms' innovative work can be argued to be positively related to risk (Buddelmeyer et al., 2010; Ericson and Pakes, 1995). In addition to the high risk, firms investing in R&D tend to face a higher cost of capital and hence be more exposed to financial constraints due to the increased information asymmetry between them and investors regarding the likelihood of success for such activities (Zhang, 2015). To summarize, innovative work entails a clear trade-off, where it on the one hand provides opportunities and thereby may improve a firm's performance, while it on the other hand can threaten the performance if the investment does not generate a good return.

Another characteristic of innovative work that is important to emphasize is that investments in innovation considerably differ between industries (Hirshleifer et al., 2013). One explanation to why this difference exists is that the level of competition in terms of creativity in an industry may vary, and this may prompt how large the investments in R&D tend to be for firms (Ugur et al., 2016). The usage of patents has also shown to vary across industries. For instance, patents are uncommon when other methods for protection easily can be used or where infringement in an innovation is difficult (Buddelmeyer et al., 2010). It is therefore important to be aware that a firm's level of innovative work could partly be explained by the industry as such.

2.3 Bankruptcy prediction

Many studies have examined which factors contribute to predict a firm's future tendency to go bankrupt, and as a result several bankruptcy prediction models have been developed. Two well-known models are Altman's Z-score model and Ohlson's O-score model (Altman, 1968; Ohlson, 1980). However, these models have been criticized for almost solely consisting of accounting-based variables. During the more recent decades, bankruptcy prediction models which to a greater extent emphasize market-based numbers have been developed with the aim of better capturing future performance. Hence, beyond accounting-based variables, those models also include more leading indicators like for example share price and other market valuations. Merton's DD model and the CHS model are examples of such market-based models (Merton, 1974; Campbell et al., 2008). However, even these models have been criticized for still being somewhat backward-looking since many of the market-based numbers include lagging data, such as information about historical volatility and stock returns (Miao et al., 2018).

To improve bankruptcy prediction, more recent studies have aimed to also include non-financial variables. One category of such measures, which has received rather limited focus in the literature, is measures of innovative performance of different kinds (Buddelmeyer et al., 2010). As discussed in section 2.2.2, there are several reasons to believe that innovation has an impact on financial performance and hence on the probability of bankruptcy. In a study by Bai and Tian (2020), the authors argue that accounting- as well as market-based models do not capture meaningful information about firms' innovative work. Their explanation to this is that accounting standards cause distortion in accounting numbers, while market numbers are being understated because of information asymmetry. This means that the effectiveness of firms' innovative work is not fully reflected in traditional bankruptcy prediction models, even though there are reasons for including such information.

2.4 Innovative performance and bankruptcy risk

2.4.1 Different measures of innovative performance in relation to bankruptcy risk

In previous theses, the results on the relationship between innovation and bankruptcy differ depending on the measures of innovative performance used. If using R&D spending, theory points in two directions. On the one hand, such investment may lead to patents as well as new products, hence increasing a firm's competitive position on the market. On the other hand, it can also be seen as a huge fixed cost that increases both leverage and volatility in earnings (Bai and Tian, 2020). In line with this two-directional theory, the results are conflicting when looking at previous studies that have examined the relationship between innovation in the form of R&D input and the probability of bankruptcy.

In a paper of Eberhart et al. (2008), the authors find a negative relationship between R&D intensity and the probability of default. They explain this by bond ratings, meaning that high R&D intensity in firms provides for lower required spreads as well as better bond ratings. Li et al. (2010) present similar findings using data on software firms, showing that R&D expenditures are associated with lower hazard rates. On the contrary, Ericson and Pakes (1995) reach the conclusion that R&D is positively related to the probability of bankruptcy since the innovative output may be unsuccessful, also concluding that failures in the innovation process are common. A study conducted by Zhang (2015) shows similar results, suggesting that R&D intensity increases firms' distress risk, especially in financially constrained firms. The arguments behind these results refer to R&D investments being inflexible and causing severe financial constraints. There are also papers that find an inverted U-shaped relationship between R&D intensity and firm survival. One of those is written by Ugur et al. (2016), who find that the relationship between R&D intensity and firm survival among R&D active firms is subject to diminishing scale effects, where firm survival increases with R&D intensity at decreasing rates up until an optimal level of R&D intensity, and then declines. Moreover, there are also studies that present mixed or insignificant results regarding the association between R&D input and bankruptcy, such as the study by Mahmood (2000). Based on the conflicting theory as well as the mixed results in previous studies for R&D input, the first hypothesis, formally stated as a null hypothesis, states that there is no association between the intensity of innovative input and the probability of bankruptcy:

$H_0$1: There is no association between the intensity of innovative input and the probability of bankruptcy.

Since patents are realized inventions and hence more certain than R&D spending, their effect on firm survival should be more positive. In theory, firms with more and higher-quality patents are more likely to obtain first mover advantages and to establish more competitive positions in the market, while patents also raise the entry barriers for potential entrants (Hsu et al., 2015). This points towards a negative relationship between the quantity of innovative output in terms of patents and bankruptcy risk, which is also what previous research is showing. As an example, Eisdorfer and Hsu (2011) find that a firm's patent count scaled by the total industry patent count has a negative relationship with the probability of bankruptcy, concluding that firms that fall behind their competitors in patent competition are more likely to go bankrupt. Bai and Tian (2020) present similar findings, suggesting a negative relationship between patent count and bankruptcy risk. Though previous studies are more aligned on the relationship between patents and bankruptcy risk compared to R&D spending, there are also some results that point in a different direction. For example, Buddelmeyer et al. (2010) show that the effect of patent applications on firm survival is negative. However, it is important to distinguish between different types of patent measures in this context. As previously mentioned, patent applications is more of an input measure in comparison to granted patents, and hence contain more uncertainty - which may explain the negative relationship to firm survival. In fact, Buddelmeyer et al. (2010) also conclude that the positive effects of innovation most clearly relate to realized measures of innovation, such as patent stocks, rather than input measures of innovation. Given that theory as well as previous studies point towards a negative relationship between realized patents and bankruptcy, our second hypothesis expresses expectations of a negative association between the quantity of innovate output and the probability of bankruptcy:

H2: There is a negative association between the quantity of innovative output and the probability of bankruptcy.

Innovative effectiveness or efficiency is not at all as commonly studied in existing literature as R&D spending or patents are. Hence, there are also much fewer conclusions drawn regarding its relation to the probability of bankruptcy. Bai and Tian (2020) find a negative relationship between bankruptcy and their measure of innovative effectiveness, namely R&D productivity. Moreover, as previously mentioned in section 2.2.2, Hirshleifer et al. (2013) find a positive effect of innovative efficiency on future earnings as well as stock price. This, in combination with efficiency being considered as something positive by nature, suggests a negative effect on bankruptcy risk. Hence, our third hypothesis proposes that innovative efficiency is negatively associated with the probability of bankruptcy:

H3: There is a negative association between the innovative efficiency and the probability of bankruptcy.

2.4.2 Bankruptcy risk in technology-intensive industries

Looking at firms in technology-intensive industries, their level of innovative work has been shown to be higher compared to other industries (Bai and Tian, 2020; Eisdorfer and Hsu, 2011). Eisdorfer and Hsu (2011) present results implying that higher technological and innovative competition leads to a higher frequency of bankruptcy, and in particular, firms that lag behind in the competition are more likely to go bankrupt. Similar findings are presented by Agarwal and Gort (2002), who show that the hazard rate is higher for firms in technology-intensive industries because technology becomes obsolete faster in such industries. These findings suggest that bankruptcies are more common among technology-intensive firms. However, important to note is that this reasoning considers the highly competitive climate in such industries. It does not necessarily mean that innovative performance has a positive association to bankruptcy in technology-intensive industries. For instance, Bai and Tian (2020) find a positive relationship to bankruptcy for R&D intensity while a negative one for patent count as well as R&D productivity when studying technology-intensive firms specifically.

3. Methodology

3.1 Research design

We have designed a prediction study with the aim of investigating if parameters measuring innovative performance can predict bankruptcy. We look at the statistical and economic significance to bankruptcy prediction among three independent variables all measuring innovative performance in different ways. Since the prediction model has a dichotomous outcome, in other words a firm either goes bankrupt or not, both probit and logit models could be used. However, within the research area of bankruptcy prediction in general as well as in relation to innovation specifically, the logit model seems to be popular to use (Ohlson, 1980; Campbell et al., 2008; Altman and Sabato, 2007; Bai and Tian, 2020; Eisdorfer and Hsu, 2011). Hence, we perform quantitative research using logistic regressions to predict the probability of bankruptcy. By doing a logit transformation of the odds of going bankrupt, the logistic regression enables us to model a linear relationship between the outcome variable and the independent variables. In the fitted logistic regression, the method of maximum likelihood is used, which means that the values that the estimates of the independent variables take on are the ones that maximize the likelihood of the observed data.

$$logit(P(BANKR_{i,t} = 1) = log\left(\frac{P(BANKR_{i,t} = 1)}{1-P(BANKR_{i,t} = 1)}\right) = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

The logit model excludes some of the key assumptions that apply for general linear regression models, namely linearity, normal distribution, and homoscedasticity. Instead, some of the underlying assumptions are that the data is collected from a random sample of observations and that the independent variables do not have multicollinearity (Christensen, 1990).

In terms of determining which variables are statistically significant to predict bankruptcy in the logistic regression, both z-statistics, p-values and t-statistics are prevalent in existing literature. In our study, we choose to look at z-statistics when testing our hypotheses, which is in line with numerous previous studies (Bai and Tian, 2020; Campbell et al., 2008), and we test the statistical significance of the coefficients at the 1%, 5% as well as 10% level.

To expand and strengthen our analysis, we also conduct two-tailed independent t-tests to find out if there are any differences in variable means between bankrupt and non-bankrupt firm-years that are statistically significantly different from zero. This provides a good basis for preliminary analysis, before running the actual regressions. In addition, we distinguish which firms belong to technology-intensive industries specifically to compare the statistics as well as the results for those firms compared to firms in general, which is in line with Bai and Tian (2020). We believe such comparison is of interest since, as mentioned in section 2.4.2, firms belonging to technology-intensive industries are shown to be more innovative.

3.2 Research models

We use four different multivariable logistic regression models to test our hypotheses. Model R1, R2 and R3 test $H_0 1$, H2 and H3 respectively, and for that purpose they include the independent variable that has been generated from the specific hypothesis that aims to be tested. This means model R1 includes the R&D intensity variable ($RDS\_TS_{i,t-3}$), model R2 includes the patent count variable ($PAT\_HC_{i,t-3}$) and model R3 includes the patent-to-R&D ratio variable ($PAT\_RDS_{i,t-3}$). In the stated regression below for model R1-R3, *Innovation Variable* is used to denote any of these three. The fourth regression model, R4, includes all three independent variables in order to further test the hypotheses and to see if and how the results may change when the variables interplay. To control for other factors that might explain a firm's bankruptcy risk or that could possibly affect the independent variables, several control variables are included in all regression models. Subscript i refers to the firm, t refers to the year and j refers to the industry. Moreover, β_0 represents the constant term while $\varepsilon_{i,t}$ represents the error term. Further definitions and motivations of the variables can be found in section 3.3.

Model R1-R3: $logit(P(BANKR_{i,t} = 1) = \beta_0 + \beta_1 Innovation\ Variable_{i,t-3} + \beta_2 TL\_TA_{i,t-1} + \beta_3 CL\_CA_{i,t-1} + \beta_4 INTWO_{i,t-1} + \beta_5 SIZE_{i,t-1} + \beta_6 WC\_TA_{i,t-1} + \beta_7 NI\_TA_{i,t-1} + \beta_8 FU\_TL_{i,t-1} + \beta_9 CHIN_{i,t-1} + \beta_{10} OENEG_{i,t-1} + \beta_{11} AGE_{i,t-1} + Year_t + Industry_j + \varepsilon_{i,t}$

Model R4: $logit(P(BANKR_{i,t} = 1) = \beta_0 + \beta_1 RDS\_TS_{i,t-3} + \beta_2 PAT\_HC_{i,t-3} + \beta_3 PAT\_RDS_{i,t-3} + \beta_4 TL\_TA_{i,t-1} + \beta_5 CL\_CA_{i,t-1} + \beta_6 INTWO_{i,t-1} + \beta_7 SIZE_{i,t-1} + \beta_8 WC\_TA_{i,t-1} + \beta_9 NI\_TA_{i,t-1} + \beta_{10} FU\_TL_{i,t-1} + \beta_{11} CHIN_{i,t-1} + \beta_{12} OENEG_{i,t-1} + \beta_{13} \mathrm{AGE}_{i,t-1} + Year_t + Industry_j + \varepsilon_{i,t}$

3.3 Variables

3.3.1 Dependent variable

The dependent variable, $BANKR_{i,t}$, is a dummy variable indicating bankruptcy. It equals one if a firm has filed for bankruptcy in a given year *t*, and zero if not. The indicator takes on a value of zero for the years up and until the year *t-1*, a value of one in year *t* and then the firm disappears from the dataset. The indicator equals zero if a firm disappears from the dataset for some other reason than bankruptcy, such as liquidation, merger, or acquisition. Apparently, a time gap between the last observable annual report and the bankruptcy filing is common, but we have restricted it to equaling a maximum of three years to get as accurate results as possible. For regression purposes, the bankrupt firms with a time gap exceeding one year are treated as if they filed for bankruptcy the year following the last observable annual report.

The reasons for choosing the filing year for bankruptcy, and not for example the bankruptcy completion date, are numerous. Firstly, it can be considered literature standard to use the filing date as an indicator of bankruptcy (Campbell et al., 2008; Bai and Tian, 2020; Altman, 1968; Franzen et al., 2007). Moreover, though firms may manage to survive as a going concern after a bankruptcy filing, the process of filing for bankruptcy as such indicates that a firm is unable to meet its financial obligations and faces financial distress. Therefore, we believe it is reasonable to use the filing date as an indicator when separating bankrupt firms from non-bankrupt ones. Since including data for the firm-years following a bankruptcy filing may disturb the analysis due to lengthy bankruptcy processes and abnormal operations distorting the figures, we have chosen to delete all firm-year observations following a bankruptcy, which is in line with the study by Campbell et al. (2008).

3.3.2 Key independent variables

Since the purpose of this study is to examine the relationship between innovative performance and bankruptcy, the key independent variables are different measures of a firm's innovative performance.

R&D intensity ($RDS\_TS_{i,t-3}$) - As a measure of the intensity of innovative input, we choose to use the commonly used measure R&D intensity, defined as R&D spending divided by total sales in line with Bai and Tian (2020) and Ugur et al. (2016). As mentioned in section 2.2.1, Swedish firms, unless restricted by complementary regulations, have the possibility to capitalize development costs. Hence, we calculate the total annual R&D spending as the sum of the expensed R&D and the current year's investment in capitalized R&D. Since all firms in the sample apply the cost of sales method, the depreciation of the capitalized R&D is included in their R&D expenditures in the income statement, and consequently also reported as expensed R&D in the database. To distinguish the actual expensed R&D for the year from the depreciation part, we need to adjust the reported expensed R&D by excluding depreciation. Because of lacking data on the R&D depreciation figure, we choose to estimate the annual depreciation on R&D by assuming an annual depreciation rate of 20% on the opening balance, which is in line with previous studies (Zhang, 2015; Franzen et al., 2007; Hirshleifer et al., 2013). This approximated depreciation figure is also used when calculating the annual investment in capitalized R&D, by adjusting for it when calculating the difference in closing and opening balance. Using the formula presented below, expressed in absolute terms, we obtain a proxy for firms' annual R&D intensity. By scaling R&D spending by total sales, we capture the innovative performance relative to the firm size. We choose to lag the variable with three years since investments in R&D are long-term, strategic investments as mentioned in section 2.2.2. Doing this, we aim to capture also the potential long-term positive effects beyond the immediate increased costs involved and thus mitigate the risk of getting a distorted view of the association to bankruptcy.

$$R\&D\ intensity = \frac{(Reported\ expensed\ R\&D\ -\ R\&D\ Depreciation)\ +\ (CB\ Capitalized\ R\&D\ +\ R\&D\ Depreciation\ -\ OB\ Capitalized\ R\&D)}{Total\ sales}$$

Patent count ($PAT\_HC_{i,t-3}$) - To measure the quantity of innovative output, we follow the literature standard and look at the patent count. Specifically, we define it as the annual successful patent applications for a firm dated by application year, following Eisdorfer and Hsu (2011). The yearly patent flow can be considered a better measure than a firm's total patent stock in a given year, since it is more informative to a firm’s economic value according to Hall (1993). By looking at successful applications, we make sure that the measure captures realized output of innovation. When calculating a firm's number of patent applications, we choose to include all types of patents (patents of invention, utility models, design patents), which is also in line with Eisdorfer and Hsu (2011). We look at patents registered by The Swedish Patent and Registration Office (Sw: Patent- och registreringsverket, PRV), both in Sweden as well as in foreign markets. We scale the innovative output by the staff headcount to take firm size into account. For the firms in the sample that do not have any data about registered patents, we let the patent count variable equal zero, which is in line with Hsu et al. (2015). We choose to lag this variable by three years too, which is reasonable since there generally is a proper time gap between the patent application and the actual grant of the patent, though seldom longer than three years (PRV, 2021).

Patent-to-R&D ratio ($PAT\_RDS_{i,t-3}$) - We choose to define our third independent variable, that aims to measure a firm's innovative efficiency, as the innovative output divided by the innovative input, which also follows the definition of innovative efficiency by Hirshleifer et al. (2013). In detail, the figure is derived by dividing annual successful patent applications, dated by application year, by annual R&D spending. This independent variable, just like the previous two described, is lagged with three years.

3.3.3 Control variables

Financial variables from the O-score model - As mentioned in section 2.3, traditional bankruptcy prediction models largely emphasize financial figures as predictive indicators. Several such figures have been shown to predict bankruptcy well according to various studies, which is comprehensible due to the nature of bankruptcy and its direct relation to financial performance. To control for such potentially explanatory factors, we include all nine accounting variables included in Ohlson's O-score model as control variables in all four regression models. This is one of the most well-known bankruptcy prediction models, and it manages to capture the essence of many of the prevalent control variables used in similar studies as ours. There are several reasons for choosing Ohlson's O-score model specifically over other prediction models. To start with, it is one of few models solely consisting of non-market-based ratios, which suits our study well since it only concerns private firms. Furthermore, Ohlson's study is conducted using a sample of considerable size compared to other famous models, and the model also presents a high prediction accuracy of 96%. More specifically, we choose to use Ohlson's model 1 which uses a one-year prediction horizon, since this model has the strongest goodness-of-fit statistics compared to model 2 and 3 (Ohlson, 1980). Thus, all the variables from the O-score model will be lagged by one year, and we expect the same signs on the variables as Ohlson himself does in his study since his expectations are very reasonable. The definitions as well as the expected signs of all nine variables are listed in Table 1 in section 3.3.4. For the detailed definitions of the variables $SIZE_{i,t-1}$ and $FU\_TL_{i,t-1}$, we choose to define them in the same way as Franzen et al. (2007) do in their study when applying the O-score model.

Firm age ($AGE_{i,t-1}$) - We also include firm age as a control variable, since the impact of firm age on firm survival is emphasized in several studies (Zhang, 2015; Ugur et al., 2016). For instance, new firms tend to be of a smaller size which make them more sensitive to macroeconomic factors (Watson and Everett, 1996; Phillips and Kirchhoff, 1989). In addition, younger firms have less assets that can be considered possible collateral, which also affects firm survival negatively (Zhang, 2015). Thus, we expect a negative relationship between firm age and a firm's tendency to go bankrupt. We calculate firm age by subtracting the registration year from the observation year, and define the control variable $AGE_{i,t-1}$ as the natural logarithm of firm age in line with the study by Zhang (2015). To increase the applicability of our results, we choose to also lag this control variable by one year to ensure that the information would be available for prediction purposes.

Industry and year fixed effects - As common in previous studies, we also control for industry and year fixed effects to account for variations across industries as well as macroeconomic factors (Zhang, 2015; Eisdorfer and Hsu, 2011; Hsu et al., 2015). As mentioned in section 2.2.2, the level of innovation as well as the innovative competition considerably vary between industries which composes a rationale for controlling for industry fixed effects. Furthermore, the occurrence of bankruptcies tends to correlate with the state of the economy and the business cycle, where a macroeconomic downturn often results in an increasing number of firms going bankrupt (Ugur et al., 2016; Bhattacharjee et al., 2009). This supports the choice of including year fixed effects as well. By including fixed effects, we also correct for potential problems with endogeneity in our sample. The fixed effects are controlled for in terms of dummies.

3.3.4 Summary of expected signs and definitions

Table 1. Expected signs and definitions of all variables

Variable	Definition	Expected sign
$RDS\_TS_{i,t-3}$	R&D spending divided by total sales	Indeterminate
$PAT\_HC_{i,t-3}$	Total successful patent applications per application year divided by total headcount	-
$PAT\_RDS_{i,t-3}$	Total successful patent applications per application year divided by R&D spending	-
$TL\_TA_{i,t-1}$	Total liabilities divided by total assets	+
$CL\_CA_{i,t-1}$	Current liabilities divided by current assets	+
$INTWO_{i,t-1}$	A dummy variable which equals one if net income was negative for the last two years, zero otherwise	+
$SIZE_{i,t-1}$	Log(total assets)	-
$WC\_TA_{i,t-1}$	Working capital, defined as current assets less current liabilities, divided by total assets	-
$NI\_TA_{i,t-1}$	Net income divided by total assets	-
$FU\_TL_{i,t-1}$	Funds provided by operations, defined as pretax income plus depreciation, divided by total liabilities	-
$CHIN_{i,t-1}$	Change in net income defined as $(NI_t – NI_{t-1})/(\|NI_t\| + \|NI_{t-1}\|)$	-
$OENEG_{i,t-1}$	A dummy variable which equals one if total liabilities exceeds total assets, zero otherwise	Indeterminate
$AGE_{i,t-1}$	Log(observation year less registration year)	-

This table presents a summary of all variables used in the regressions, both independent and control, and their respective definitions and expected signs. In order to not lose observations where any denominator or value that is logarithmized equals zero, we add a +1 to all such values used in the variables given that the data requires it. Hence, $PAT\_RDS_{i,t-3}$, $CHIN_{i,t-1}$ *and* $AGE_{i,t-1}$ *all have one unit added to the denominator or the value being logarithmized.*

3.4 Empirical data

3.4.1 Data collection

To test our hypotheses and run the required regressions, we combine five different datasets derived from several different sources. Three of the datasets are sourced from the Serrano Database, produced by the Swedish House of Finance, and these constitute the main data source for our study. The reason for choosing the Serrano Database is that it contains data for all Swedish firms for an adequate period of time, including both financial statement data as well as bankruptcy data from the Swedish Companies Registration Office (Sw: Bolagsverket), and also other general company data from Statistics Sweden (Sw: Statistiska Centralbyrån). Hence, the database suits our study well and all together, these three datasets provide all the inputs needed for constructing the regression variables as well as restricting our sample, except for the patent count and the list of Swedish listed firms. From the first Serrano dataset, which contains general company data, we derive information about firms' legal form, group situation, activity, registration date and industry code. From the second Serrano dataset including financial statement data, we extract data regarding the firms' financial accounts and their types, fiscal year, staff headcount and all financial data required for the regression variables. The third and last Serrano dataset provide information about bankruptcy filings, including the date of the filings.

The patent data, on the other hand, is derived from the fourth dataset sourced from PAtLink, which is also produced by The Swedish House of Finance. PAtLink has information about all patents belonging to Swedish firms in the last decades, and we extract the number of patents per firm and application year. Lastly, the fifth dataset is sourced from Retriever Business and contains a list of Swedish firms that are, or have recently been, listed, which is needed to be able to exclude those from our sample.

Serrano has all required input data for the years 1998-2018, as well as some bankruptcy data for 2019, while PAtLink has data for 1990-2018. Due to the use of lagged variables, where the innovation variables are lagged with as much as three years and also include calculations using opening balance figures, we are able to study bankruptcies over the time period 2002-2019. We have done some thorough research and to the best of our knowledge, there have not been any significant changes in the filing or registration process of bankruptcies or patents during this time period, hence verifying that the data is comparable over the years.

3.4.2 Sample selection

We commenced the process of creating our sample by merging the two Serrano files containing general company data as well as financial statement data, and then also included data on patent count per firm and application year from the PAtLink dataset. Since we use panel data, we linked the different datasets using both organizational number as well as observation year, leading to observations missing this data dropping out. In this process, we carefully matched the fiscal year data into calendar year data looking at the calendar year when the fiscal year ended. Subsequently, we excluded all observations of firms that are not limited liability companies as well as firms that currently are, or have recently been, listed. In that way, we restricted our sample to Swedish private limited liability companies only. After careful consideration, we also excluded all firms that are subsidiaries, both in Swedish and foreign groups, since we concluded that their financial data was not very informative and would distort our sample and disturb our analysis.

In the next step, we restricted our sample to only include firms of small and medium size, using the European Commission's definition of small and medium-sized enterprises (SMEs) (European Commission, 2012). They determine size based on staff headcount as well as either turnover or balance sheet total in EUR. In this study, the monetary figures are translated into SEK from EUR based on the assumption that EUR 1 = SEK 10, which makes them approximations. This restriction was applied on individual firm-years, meaning that a firm that has been a SME anytime since 1998 is included in our sample during those specific years. However, for all such firms, we also allowed for their size to be lower than the requirements for individual firm-years after the SME requirements have been fulfilled, since firms tend to decrease in size when approaching a bankruptcy. Hence, this was done to not lose important observations prior to a bankruptcy. The overall reason for excluding firms of both micro and large size is to make the sample more comparable, since both the possibility for as well as the importance of innovation differ depending on size. In this step, all observations where headcount information was missing or zero were also excluded.

Further, we excluded all firms that do not use the cost of sales method, as the data for those firms is insufficient, as described in section 1.5, which prevents us from thoroughly calculating some of the key independent variables. After that, we also removed firms that are operating within the Finance & Real estate branch sector as defined per the Serrano Database. This exclusion is in line with the approach in previous studies, where such firms are often excluded due to structural differences as well as differences in regulatory oversight (Ohlson, 1980; Zhang, 2015). In this step, firms missing an industry classification were also excluded. This was followed by excluding firm-years for parent firms' individual accounts as well as independent firms' wrongfully reported consolidated accounts in order to get more accurate data and to get rid of duplicates.

To further enhance the accuracy of the data in our sample, we also removed all the firm-years where a firm is not both registered and active as well as the firm-years where some required financial data is missing, abnormal or seems incorrect. The financial information required is turnover, tax expense, net income, current assets, total assets, book value of equity, current liabilities as well as total value of equity and liabilities combined. We chose to not require data for any of the R&D related figures, namely expensed and capitalized R&D, nor for any type of depreciation. The reason for this is that some firms do not have these items in their financial statements, which can show up as missing values in the database. Hence, all the missing values for the figures mentioned were coded into zeros instead. We also excluded firm-year observations where total sales or assets equalized zero, since such firm-years can be considered abnormal. Furthermore, we carefully looked through the data to see if any of the reported numbers in the sample seemed incorrect. In this process we noticed that depreciation was incorrectly reported as positive in some cases when it should not have been so, and therefore all observations with reported positive depreciation of any kind were excluded from the sample. Lastly, we excluded all firm-years where the fiscal year falls short of or exceeds 12 months, as well as made sure that there were no duplicates left in the sample. After all these restrictions, we ended up with a raw data sample containing 21,668 firm-year observations.

In the next step, we merged the raw data sample with the bankruptcy filings from the third Serrano dataset based on organization number and year. In this merge, we only included bankruptcy observations for the firms we had financial data for. Thereafter, we excluded all firm-years following a firm-year where a bankruptcy filing occurred, as per the definition of the dependent variable in section 3.3.1. In this step, no observations from the raw data sample were deleted, but a few of the bankruptcy observations dropped out. Afterwards, we excluded all bankruptcy filings with a time gap to the last observable annual report exceeding three years. Since the vast majority of the bankruptcy filings had a time gap of at least a year, the inclusion of bankruptcies in the raw data sample increased the sample size by 150.

Last of all, we excluded all firm-year observations that had missing data for any of the lagged variables generated. This gave us a final sample of 10,335 firm-year observations, representing 1,659 individual firms in total of which 67 bankrupt ones, meaning that 4.04% of the firms in our sample file for bankruptcy. The distribution of the final sample across industry groups, as sectioned by the Serrano Database, as well as the distribution of bankruptcies over the years are presented in Appendix 1 and 2, respectively.

Table 2. Sample selection process

Category	Number of observations
Total number of firm-years for all Swedish firms for the period 1998-2018	7,560,985
Firms that are not limited liability companies	-24,348
Firms listed on any Swedish exchange	-18,132
Firms that are subsidiaries	-1,747,102
Firms that are not SMEs or have headcount data that is missing or zero	-5,462,858
Firms that do not use the cost of sales method	-274,009
Firms that are in the Finance & Real estate branch sector or missing industry code	-4,185
Firm-years for parent firms' individual accounts, as well as for independent firms' consolidated accounts	-7,435
Firm-years where a firm is not both registered and active	-73
Firm-years where required financial data is missing, abnormal or seems incorrectly reported	-201
Firm-years where the fiscal year exceeds or falls short of 12 months	-974
Raw data sample	**21,668**
Additional bankruptcy observations (incl. 2019)	+150
Raw data sample including all bankruptcy observations	**21,818**
Firm-years with missing data for the lagged regression variables	-11,483
Final sample	**10,335**

3.4.3 Data quality check

When working with databases containing large amounts of data, such as the Serrano Database, there is always some uncertainty regarding the accuracy of the data. Striving for a sample of as high quality as possible, we have carefully cross-checked financial data for randomly selected firms by comparing it to the actual numbers in annual reports derived from Retriever Business. This included checking what the different database figures actually measure, as well as that the data aligns with actual reported numbers. Moreover, to further enhance the quality of the data in our sample all continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level, hence suppressing the effect of extreme outliers.

4. Empirical results and analysis

4.1 Descriptive statistics

Examining the descriptive statistics for the final sample in Table 3, we see that the mean of the dependent variable $BANKR_{i,t}$ is 0.006, which means that 0.6% of the firm-year observations in our sample are bankrupt ones. Moreover, both the dependent variable $BANKR_{i,t}$ as well as the three independent variables have a highly positive skewness, where the 75th percentile presents a value of zero. This indicates that both bankruptcy and innovation are unusual among the firms in our sample. When comparing the means of the innovation variables to those presented in previous studies, we note that our reported means are substantially lower. For example the mean that we present for the variable $RDS\_TS_{i,t-3}$ of 0.019 can be compared to those in studies by Bai and Tian (2020) as well as Ugur et al. (2016), who report higher means of 0.09 and 0.20 respectively. When looking at the variable $PAT\_RDS_{i,t-3}$, we can see that also this mean is lower compared to what Hirshleifer et al. (2013) present for the same measure. This suggests that Swedish private firms might innovate less than firms, and especially listed firms, in non-Swedish markets that have been examined in previous studies - at least looking at our specific measures of innovation. However, no broader conclusions on this can be drawn. Note that when interpreting the statistics, it is important to bear in mind the +1 that is added to some denominators and values being logarithmized, as described in the note to Table 1. This might for example explain the high maximum value of 51.000 for $PAT\_RDS_{i,t-3}$.

Table 3. Summary of descriptive statistics

N=10,335	**Mean**	**Std**	**Min**	**25^{th}**	**Median**	**75^{th}**	**Max**
$BANKR_{i,t}$	0.006	0.080	0.000	0.000	0.000	0.000	1.000
$RDS\_TS_{i,t-3}$	0.019	0.348	-0.019	0.000	0.000	0.000	13.349
$PAT\_HC_{i,t-3}$	0.005	0.082	0.000	0.000	0.000	0.000	4.353
$PAT\_RDS_{i,t-3}$	0.044	0.998	-0.001	0.000	0.000	0.000	51.000
$TL\_TA_{i,t-1}$	0.644	0.224	0.023	0.496	0.662	0.797	3.051
$CL\_CA_{i,t-1}$	0.704	0.420	0.011	0.443	0.649	0.882	9.057
$INTWO_{i,t-1}$	0.060	0.238	0.000	0.000	0.000	0.000	1.000
$SIZE_{i,t-1}$	17.125	1.313	14.533	16.099	16.967	18.051	20.574
$WC\_TA_{i,t-1}$	0.226	0.242	-1.866	0.059	0.217	0.382	1.078
$NI\_TA_{i,t-1}$	0.064	0.106	-1.899	0.016	0.060	0.112	0.924
$FU\_TL_{i,t-1}$	0.276	0.431	-4.255	0.099	0.210	0.382	22.171
$CHIN_{i,t-1}$	0.020	0.548	-1.000	-0.263	0.033	0.325	1.000
$OENEG_{i,t-1}$	0.022	0.146	0.000	0.000	0.000	0.000	1.000
$AGE_{i,t-1}$	2.898	0.637	1.386	2.398	2.833	3.332	4.762

This table presents a summary of the descriptive statistics for all the variables for the final sample including 10,335 firm-year observations. All continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4.

In Table 4, we can see the descriptive statistics for the final sample divided by bankrupt and non-bankrupt firm-year observations. We conduct two-tailed independent t-tests to examine how the means of the variables differ between the two groups, and whether the differences are statistically significantly different from zero. As presented in Table 4, the means of the variables $RDS\_TS_{i,t-3}$ and $PAT\_HC_{i,t-3}$ are higher for bankrupt firm-years, though we cannot draw any further conclusions since the mean differences are not statistically significant. However, there is a significant mean difference for the variable $PAT\_RDS_{i,t-3}$, which measures innovative efficiency in terms of the patent-to-R&D ratio, where the mean is statistically significantly higher for non-bankrupt firm-years at the 1% level. Although the t-test does not reveal significant explanatory power, this is in line with our third hypothesis that innovative efficiency is negatively associated with bankruptcy. The association will be tested further in terms of statistical significance in section 4.3.

Furthermore, we note that for the control variables derived from Ohlson's O-score model, there is a highly statistically significant difference in means between the two groups for all the variables. This suggests that the financial performance differs between bankrupt and non-bankrupt firms, which is reasonable. For the variables $TL\_TA_{i,t-1}$, $CL\_CA_{i,t-1}$, and $INTWO_{i,t-1}$, those present a statistically significantly higher mean among the bankrupt firm-years, which is in line with our expectation that those are positively associated with bankruptcy. Moreover, the variables $SIZE_{i,t-1}$, $WC\_TA_{i,t-1}$, $NI\_TA_{i,t-1}$, $FU\_TL_{i,t-1}$ and $CHIN_{i,t-1}$ present statistically significantly lower means among the bankrupt firm-years, which is also in line with our expectation that those are negatively associated with bankruptcy probability. The remaining financial variable, $OENEG_{i,t-1}$, which had no determinate expectation, shows a statistically significantly higher mean among bankrupt firm-years, which is plausible due to the nature of the measure. However, there is no significant mean difference for $AGE_{i,t-1}$.

Table 4. Bankrupt vs Non-bankrupt firm-year observations

	Bankrupt firm-years (N=67)	Non-bankrupt firm-years (N=10,268)
Variable	**Mean**	**Mean**
$BANKR_{i,t}$	1.000	0.000
$RDS\_TS_{i,t-3}$	0.066	0.018
$PAT\_HC_{i,t-3}$	0.006	0.005
$PAT\_RDS_{i,t-3}$	0.000***	0.044***
$TL\_TA_{i,t-1}$	0.924***	0.642***
$CL\_CA_{i,t-1}$	1.145***	0.701***
$INTWO_{i,t-1}$	0.269***	0.059***
$SIZE_{i,t-1}$	16.497***	17.130***
$WC\_TA_{i,t-1}$	-0.025***	0.228***
$NI\_TA_{i,t-1}$	-0.101***	0.065***
$FU\_TL_{i,t-1}$	-0.046***	0.278***
$CHIN_{i,t-1}$	-0.279***	0.022***
$OENEG_{i,t-1}$	0.224***	0.020***
$AGE_{i,t-1}$	2.787	2.899

*This table presents the mean for all variables for bankrupt and non-bankrupt firm-year observations in the final sample respectively. All continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4. *,** and *** denote significance of the mean difference at the 0.1, 0.05 and 0.01 level, respectively and refer to the two-tailed t-test.*

As mentioned in section 3.1, we also distinguish firms in technology-intensive industries from firms in general. We follow the same approach as when separating between bankrupt and non-bankrupt firm-years and conduct two-tailed independent t-tests, where we compare how the means for technology-intensive industries differ from the means for all industries combined.

In terms of defining which industries are technology-intensive, we use OECD's definition of such industries, which is based on ISIC codes (OECD, 2011). Since we have data on the SNI 2007 codes, we look at the first two digits of those as well as their respective definitions and find corresponding matches to the industries that OECD classify as either high-technology or medium-high-technology industries. Please see Appendix 3 for the major industry groups, given by the two-digit SNI 2007 codes, that we define as technology-intensive.

The descriptive statistics divided by all industries and technology-intensive industries respectively are presented in Appendix 4. By examining the table, we see that the mean of the dependent variable $BANKR_{i,t}$ is statistically significantly higher for the firms in technology-intensive industries, which supports the results that such industries have higher hazard rates as presented by Eisdorfer and Hsu (2011) as well as Agarwal and Gort (2002). Also, the means of all three innovation variables are statistically significantly higher for technology-intensive firms, which implies that innovative performance among firms in such industries is more common, as expected.

4.2 Correlation matrix

In Table 5.A and Table 5.B, we present the Pearson correlation matrix for all variables for the final sample. The results show that regarding the three independent variables, $PAT\_HC_{i,t-3}$ is statistically significantly positively correlated with both $RDS\_TS_{i,t-3}$ and $PAT\_RDS_{i,t-3}$ at the 1% level. This implies that a firm with a larger patent count is more likely to have higher R&D intensity as well as a higher patent-to-R&D ratio. These results seem very logical. R&D intensity measures the intensity of innovative input in terms of R&D spending, and investments in R&D are necessary to receive any innovative output in terms of patents. Furthermore, it is comprehensible that the efficiency of which a firm transforms their R&D input into patents correlates with the patent count. Though some of the pairs among the three independent variables are significantly correlated, the correlations can be considered rather low, which suggest that all three variables are of value on their own to include to better reflect innovative performance. No significant correlation between the independent variables and $BANKR_{i,t}$ can be derived from the Pearson correlation matrix. However, those relationships will be further tested through the logit regressions in section 4.3.

Moreover, the tables display highly statistically significant correlations between all the financial control variables derived from the O-score model and the dependent variable $BANKR_{i,t}$, which once again suggest that financial performance and bankruptcy are related. Also, there are statistically significant correlations between almost all the financial control variables. However, correlations between those can be expected since the inputs for these variables are often shared, such as total assets and total liabilities. In addition, many of the financial control variables are naturally connected since they measure a firm's financial position in different ways, and thus often point in similar directions. For the control variable $AGE_{i,t-1}$, there is no statistically significant correlation to $BANKR_{i,t}$, but to many of the other control variables, suggesting that experience and financial performance are related.

Since some variable pairs, especially among the financial control variables, showed rather high correlations, we tested for multicollinearity by looking at the variance inflation factor (VIF) for all variables, which is a measure to detect multicollinearity. The results are presented in Appendix 5, and we can see that all VIF values are considerably lower than 10, which generally can be considered a limit for when the correlation is acceptable (Woolridge, 2018). More specifically, no VIF value exceeded 4, and the values for our three independent variables were around 1 - which implies low correlation to the other variables. This means that the correlations between our variables do not imply any statistical problem.

Table 5.A. Correlation matrix

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $BANKR_{i,t}$	1.000						
(2) $RDS\_TS_{i,t-3}$	0.011	1.000					
(3) $PAT\_HC_{i,t-3}$	0.002	0.351***	1.000				
(4) $PAT\_RDS_{i,t-3}$	-0.004	-0.002	0.156***	1.000			
(5) $TL\_TA_{i,t-1}$	0.101***	-0.065***	-0.059***	-0.029***	1.000		
(6) $CL\_CA_{i,t-1}$	0.085***	-0.012	-0.025***	-0.020**	0.561***	1.000	
(7) $INTWO_{i,t-1}$	0.071***	0.165***	0.143***	0.032***	0.153***	0.139***	1.000
(8) $SIZE_{i,t-1}$	-0.039***	0.039***	0.035***	0.057***	-0.241***	-0.216***	0.050***
(9) $WC\_TA_{i,t-1}$	-0.084***	-0.004	0.030***	0.010	-0.622***	-0.824***	-0.113***
(10) $NI\_TA_{i,t-1}$	-0.127***	-0.209***	-0.188***	-0.011	-0.305***	-0.185***	-0.291***
(11) $FU\_TL_{i,t-1}$	-0.060***	-0.162***	-0.146***	0.002	-0.392***	-0.196***	-0.173***
(12) $CHIN_{i,t-1}$	-0.044***	-0.003	-0.001	0.014	-0.017*	-0.028***	0.200***
(13) $OENEG_{i,t-1}$	0.112***	0.002	-0.005	-0.007	0.395***	0.239***	0.228***
(14) $AGE_{i,t-1}$	-0.014	-0.017*	-0.024**	0.000	-0.124***	-0.075***	-0.021**

Table 5.B. Correlation matrix continued

Variable	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(8) $SIZE_{i,t-1}$	1.000						
(9) $WC\_TA_{i,t-1}$	0.217***	1.000					
(10) $NI\_TA_{i,t-1}$	-0.035***	0.182***	1.000				
(11) $FU\_TL_{i,t-1}$	0.020**	0.223***	0.645***	1.000			
(12) $CHIN_{i,t-1}$	0.014	0.034***	0.325***	0.175***	1.000		
(13) $OENEG_{i,t-1}$	-0.095***	-0.248***	-0.262***	-0.106***	-0.027***	1.000	
(14) $AGE_{i,t-1}$	0.104***	0.085***	-0.021**	0.005	-0.009	-0.060***	1.000

*These tables present the Pearson correlation matrix for all variables calculated on the final sample including 10,335 firm-year observations. All continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4. *,** and *** denote significance of the correlations at the 0.1, 0.05 and 0.01 level, respectively.*

4.3 Regression results

Using the logistic regression models specified in section 3.2, we test the association between the three different measures of innovative performance, that is R&D intensity, patent count and the patent-to-R&D ratio, and the probability of bankruptcy. The first three regression models, R1, R2 and R3, each includes one of the three independent variables $RDS\_TS_{i,t-3}$, $PAT\_HC_{i,t-3}$ and $PAT\_RDS_{i,t-3}$ and hence fulfills the purpose of testing $H_0 1$, H2 and H3 respectively, while R4 tests all three variables in combination. The results of all four regressions are presented in separate columns in Table 6.

The dependent variable for the regressions is $BANKR_{i,t}$, and the control variables, which are included in all four regressions, are $TL\_TA_{i,t-1}$, $CL\_CA_{i,t-1}$, $INTWO_{i,t-1}$, $SIZE_{i,t-1}$, $WC\_TA_{i,t-1}$, $NI\_TA_{i,t-1}$, $FU\_TL_{i,t-1}$, $CHIN_{i,t-1}$, $OENEG_{i,t-1}$ and $AGE_{i,t-1}$. In all regressions, industry dummies and year dummies are included to control for fixed effects, and standard errors are clustered by firm, which also correct for potential issues with serial correlation. The industry fixed effects are accounted for based on the industry groups presented in Appendix 1. Industry and year fixed effects are included in all regressions to consistently account for the differences in innovation and bankruptcy patterns between both sectors and years. The number of firm-year observations run in the regressions equals 10,239, which is slightly lower than the final sample of 10,335 because of certain industries and years being omitted due to perfect prediction or collinearity.

The first regression, R1, leads to the conclusion that the independent variable measuring innovative input in terms of R&D intensity, $RDS\_TS_{i,t-3}$, is not statistically significant, and this fact remains also in the combined regression R4. Consequently, our first null hypothesis cannot be rejected. Noteworthy still is that the sign of the coefficient is negative in both regressions, but due to the lack of significance it is difficult to draw any broader conclusions regarding the relationship to bankruptcy risk. However, in the second regression named R2, the independent variable measuring innovative output in terms of patents, $PAT\_HC_{i,t-3}$, shows a negative coefficient, in line with our expectations, and is also statistically significant at the 1% level. This is true still in the combined regression R4 and implies that we find rather solid support for our second hypothesis. The coefficient of the third independent variable, $PAT\_RDS_{i,t-3}$, which measures innovative efficiency in terms of the patent-to-R&D ratio, is also negative in accordance with our expectations and statistically significant at the 5% level on a stand-alone basis in R3. In the fourth, combined regression, R4, the statistical significance is enhanced to the 1% level. Thus, this indicates that we also find support for our third hypothesis.

The increase in significance of $PAT\_RDS_{i,t-3}$ in the combined model is plausible, since the explanatory power of the efficiency measure might strengthen when the input and output measures on their own are controlled for. In fact, by looking at the log likelihood as well as Mc Fadden's Pseudo R^2 for the four regressions, which are both useful tools in terms of comparing the goodness-of-fit of multiple logistic regressions, we note that the overall fit of the combined model, R4, is superior to the others since it presents the highest values. It seems that when taking both innovative input, output as well as the efficiency into account, bankruptcies are better predicted. Further discussion and interpretation of our main findings will be elaborated on in section 5.

Besides analyzing the statistical significance of the independent variables $PAT\_HC_{i,t-3}$ and $PAT\_RDS_{i,t-3}$, we also look at the economic significance to get an understanding of the economic implications in terms of the magnitude of the effect on the odds of bankruptcy. As the outcome variable in our logistic regressions is the natural logarithm of the bankruptcy odds, we exponentiate the coefficients to get the non-logarithmized impact of the variables on the odds of bankruptcy, also called their odds ratios. We choose to look at the coefficients in the combined model, R4, since that is the model that best reflects reality in terms of the three innovation variables interplaying, and it also seems to have the best fit. When exponentiating the coefficient for $PAT\_HC_{i,t-3}$ of -1.870, we obtain an odds ratio of 0.154, meaning that a one unit increase in $PAT\_HC_{i,t-3}$ decreases the bankruptcy odds with 84.6%. The same procedure for $PAT\_RDS_{i,t-3}$, using the coefficient of -4.891, provides an odds ratio of 0.008, implying that a one unit increase decreases the bankruptcy odds with 99.2%. Thus, the economic significance of both these variables can be considered substantial, which further strengthens the support of our second and third hypotheses. However, it is important to keep in mind that it is uncommon that the patent count scaled by headcount increases by as much as one unit, while the patent-to-R&D ratio could more easily do so given that the R&D spending is zero, interpreted as one Swedish krona in our variable definition, as further explained in Table 1.

As for the control variables, five out of ten show statistically significant results at meaningful significance levels. $INTWO_{i,t-1}$ consistently shows a positive coefficient, aligned with Ohlson's own results (Ohlson, 1980) as well as our expectations, that is significant on a level ranging between 5% and 10% throughout the four regressions. $WC\_TA_{i,t-1}$ displays a negative coefficient, as expected, that is significant on the 5% level in all four regressions. Moreover, $SIZE_{i,t-1}$, $FU\_TL_{i,t-1}$ and $CHIN_{i,t-1}$ all display negative coefficients, in line with our expectations, that are highly statistically significant on the 1% level throughout all four regressions. The remaining five control variables that do not show statistically significant results are $TL\_TA_{i,t-1}$, $CL\_CA_{i,t-1}$, $NI\_TA_{i,t-1}$, $OENEG_{i,t-1}$ and $AGE_{i,t-1}$. $TL\_TA_{i,t-1}$, $NI\_TA_{i,t-1}$ and $AGE_{i,t-1}$ all show coefficients as expected, while $CL\_CA_{i,t-1}$ shows a slightly negative coefficient in contrast to our expectations. The expected sign for $OENEG_{i,t-1}$ was indeterminate, and the coefficient turned out to be positive. However, the lack of significance prevents us from drawing any further conclusions based on this.

Table 6. Regression results

Variable	R1	R2	R3	R4
$RDS\_TS_{i,t-3}$	-0.144 (-0.70)			-0.079 (-0.46)
$PAT\_HC_{i,t-3}$		-2.041*** (-2.81)		-1.870*** (-2.81)
$PAT\_RDS_{i,t-3}$			-6.451** (-2.09)	-4.891*** (-11.08)
$TL\_TA_{i,t-1}$	0.800 (1.12)	0.740 (1.06)	0.831 (1.18)	0.727 (1.02)
$CL\_CA_{i,t-1}$	-0.024 (-0.07)	-0.041 (-0.11)	-0.005 (-0.01)	-0.051 (-0.14)
$INTWO_{i,t-1}$	0.896* (1.92)	0.931** (2.03)	0.883* (1.88)	0.937** (2.05)
$SIZE_{i,t-1}$	-0.503*** (-3.25)	-0.496*** (-3.24)	-0.506*** (-3.26)	-0.489*** (-3.17)
$WC\_TA_{i,t-1}$	-2.198** (-2.29)	-2.261** (-2.38)	-2.149** (-2.29)	-2.284** (-2.35)
$NI\_TA_{i,t-1}$	-0.365 (-0.31)	-0.348 (-0.30)	-0.291 (-0.27)	-0.371 (-0.31)
$FU\_TL_{i,t-1}$	-1.243*** (-3.68)	-1.410*** (-4.04)	-1.217*** (-4.05)	-1.422*** (-3.82)
$CHIN_{i,t-1}$	-0.691*** (-3.00)	-0.676*** (-2.95)	-0.698*** (-3.03)	-0.673*** (-2.94)
$OENEG_{i,t-1}$	0.225 (0.35)	0.211 (0.32)	0.230 (0.35)	0.210 (0.32)
$AGE_{i,t-1}$	-0.112 (-0.48)	-0.128 (-0.55)	-0.108 (-0.47)	-0.127 (-0.55)
Constant	5.075* (1.65)	5.193* (1.70)	5.104* (1.66)	5.081* (1.66)
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	10,239	10,239	10,239	10,239
Log likelihood	-311.730	-310.569	-311.666	-310.316
Pseudo R^2	0.218	0.221	0.218	0.222

*This table presents the results for the four logistic regressions R1, R2, R3 and R4, as described in section 3.2. All regressions include industry and year fixed effects and consist of 10,239 firm-year observations, of which 66 are bankrupt ones. Standard errors are clustered by firm, and all continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4. The z-statistics are presented in the parentheses below the coefficients. *,** and *** denote significance of the coefficients at the 0.1, 0.05 and 0.01 level, respectively.*

In line with the study by Bai and Tian (2020), we also test the relationship between the three different measures of innovative performance and the probability of bankruptcy for technology-intensive industries specifically, which are defined in Appendix 3. This is done by using the same four regressions as previously used, with the only difference that industry fixed effects are now excluded. The regression results of these regressions are tabulated in Appendix 6. Though there were 334 firm-year observations in the technology-intensive sample to start with, as reported in Appendix 4, several years drop out in the regressions due to perfectly predicting $BANKR_{i,t}$ being 0. Hence, the sample included in the regression results consists of 138 firm-year observations, of which 10 are bankruptcies.

For R&D intensity, we find similar results as in the main analysis presented in Table 6, where $RDS\_TS_{i,t-3}$ is statistically insignificant both on a stand-alone basis as well as in the combined model. Nonetheless, the results for $PAT\_HC_{i,t-3}$ differ compared to the all industry sample, as the coefficient no longer is statistically significant at any meaningful level in either R2 or R4. Noteworthy, though, is that we find similar but even more significant results for the third independent variable, $PAT\_RDS_{i,t-3}$, for the technology-intensive industries. The coefficient is all the while negative and consistently statistically significant at the 1% level. Further elaborations on these results and potential explanations to why they differ will be presented in section 5.

Notable is also that only three of the control variables are statistically significant at any meaningful level of significance in any of the regressions this time, compared to the five control variables with significant coefficients in Table 6. These variables are $TL\_TA_{i,t-1}$, $CL\_CA_{i,t-1}$ and $WC\_TA_{i,t-1}$, where $WC\_TA_{i,t-1}$ is the only one that is consistently statistically significant at a level of at least 5% in all four regressions. This implies that there are other factors that play a more important role for technology-intensive firms in terms of predicting bankruptcy.

4.4 Robustness tests and additional analyses

4.4.1 Robustness tests

To check the robustness of our results, we test the relationship between our three measures of innovative performance and the probability of bankruptcy using different time lags for the independent variables specifically. We choose to do this using the fourth regression model, R4, since it includes all three independent variables in combination and also seems to predict bankruptcy the best according to the performed goodness-of-fit tests.

The robustness tests include four alterations of the combined regression model R4. In the first two altered models, we calculate the average as well as the cumulative measures over three years for the three independent variables. Similar robustness tests are performed in previous studies, such as Hsu et al. (2015). As all three independent variables consist of ratios, we define the average measure for each as the mean of the ratios over the past three years. The calculations of the cumulative measures are somewhat more complicated. As for $RDS\_TS_{i,x}$ and $PAT\_RDS_{i,x}$, we define the cumulative measure as the sum of the numerators over the past three years divided by the sum of the denominators for the same period. For $PAT\_HC_{i,x}$, on the other hand, we define it as the sum of the numerators over the past three years divided by the average number of the denominator for the same period. The reason for this is that the patent count is scaled by headcount for the sole reason of taking size into account, and size is not an additive measure in the same way as total sales and R&D spending are. The two other alterations of the combined model use simple time lags for the independent variables, of two years and one year respectively. These are prevalent time lags in previous studies, and hence of interest to test.

Apart from altering the time lag of the independent variables in R4, the regression is identical to the original one used in the main analysis in section 4.3. However, in order to include as many observations as possible in the regressions with a shorter time lag than three years, the sample used in the regressions is not the same. Instead, it is the raw data sample including all bankruptcy observations, as described in Table 2 in section 3.4.2, that is being used. For consistency, we continue to study bankruptcies over the years 2002-2019. The results of the altered regressions are presented in Table 7.

The regression outcomes of the regression models using three-year average as well as three-year cumulative measures show rather similar results compared to the ones in our main analysis, presenting statistically significant coefficients with negative signs for both $PAT\_HC_{i,x}$ and $PAT\_RDS_{i,x}$, but no statistically significant coefficients for $RDS\_TS_{i,x}$. In both regressions, the coefficient of $PAT\_HC_{i,x}$ is statistically significant at the 5% level, while the coefficient of $PAT\_RDS_{i,x}$ is highly statistically significant at the 1% level. Also for the altered regression using two-year lagged variables the results are similar, with statistically significant negative coefficients for the same two variables, now both at the 5% level. Examining the results for the three previously mentioned altered regression models, we see no sign of any substantial changes in the outcome. On the contrary, the results are very similar, suggesting that the original combined model R4 used in the main analysis is appropriate and provides for robust results.

For the regression where the independent variables are only lagged with one year, none of the coefficients are statistically significant at any meaningful level. Potential explanations and interpretations of this will be further elaborated on in section 5. However, worth noting is that that specific regression has lower values for both the log likelihood as well as Mc Fadden's

Pseudo R^2, indicating a worse goodness-of-fit compared to the other models. Another noteworthy remark, though insignificant, is that $RDS\_TS_{i,t-1}$ shows a positive coefficient. Still, due to the lack of significance, we cannot draw any broader conclusions regarding the one-year lagged variable's association to the probability of bankruptcy.

Table 7. Regression results for different time lags of innovation variables

Variable	3-year average	3-year cumulative	2-year lag	1-year lag
$RDS\_TS_{i,x}$	0.018 (0.10)	0.022 (0.13)	-0.034 (-0.31)	0.111 (0.93)
$PAT\_HC_{i,x}$	-3.343** (-2.16)	-1.074** (-2.19)	-1.028** (-2.04)	-1.278 (-1.12)
$PAT\_RDS_{i,x}$	-15.533*** (-5.83)	-5.410*** (-7.41)	-6.487** (-2.11)	-759.639 (-0.43)
$TL\_TA_{i,t-1}$	0.717 (1.01)	0.722 (1.02)	1.134 (1.58)	1.024 (1.51)
$CL\_CA_{i,t-1}$	-0.044 (-0.12)	-0.042 (-0.12)	-0.149 (-0.42)	-0.263 (-0.66)
$INTWO_{i,t-1}$	0.939** (2.05)	0.937** (2.05)	0.592 (1.38)	0.750** (2.02)
$SIZE_{i,t-1}$	-0.492*** (-3.17)	-0.494*** (-3.19)	-0.413*** (-3.16)	-0.354*** (-2.84)
$WC\_TA_{i,t-1}$	-2.285** (-2.38)	-2.276** (-2.37)	-1.876** (-2.09)	-2.302** (-2.53)
$NI\_TA_{i,t-1}$	-0.316 (-0.28)	-0.318 (-0.28)	-0.529 (-0.67)	-0.458 (-0.71)
$FU\_TL_{i,t-1}$	-1.424*** (-4.01)	-1.415*** (-4.02)	-0.989*** (-4.06)	-0.601*** (-3.90)
$CHIN_{i,t-1}$	-0.679*** (-2.97)	-0.679*** (-2.97)	-0.788*** (-3.71)	-0.900*** (-4.29)
$OENEG_{i,t-1}$	0.209 (0.32)	0.209 (0.32)	0.591 (1.09)	0.565 (1.12)

(Continued)

Table 7. *Continued*

$AGE_{i,t-1}$	-0.126 (-0.55)	-0.126 (-0.55)	0.005 (0.03)	-0.030 (-0.19)
Constant	5.152* (1.68)	5.182* (1.69)	2.829 (1.07)	2.201 (0.88)
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	10,239	10,239	11,549	13,004
Log likelihood	-310.000	-310.115	-390.848	-449.207
Pseudo R^2	0.222	0.222	0.222	0.208

*This table presents the results of four logistic regressions using alternate time lags for the three independent variables, where all regressions include all three independent variables. Column 1 presents a regression using three-year average measures, and column 2 presents a regression using three-year cumulative measures. These two regressions both consist of 10,239 firm-year observations, of which 66 are bankrupt ones. Column 3 presents a regression using a two-year time lag for the independent variables and consists of 11,549 firm-year observations, of which 85 bankrupt ones. Column 4 presents a regression using a one-year time lag and consists of 13,004 firm-year observations, of which 96 bankrupt ones. All regressions include industry and year fixed effects. Standard errors are clustered by firm, and all continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The original definitions of all variables are listed in Table 1 in section 3.3.4. The z-statistics are presented in the parentheses below the coefficients. *,** and *** denote significance of the coefficients at the 0.1, 0.05 and 0.01 level, respectively.*

4.4.2 Additional analysis

As mentioned in section 1.5, this thesis is delimited to studying the relationship between innovative performance and the probability of bankruptcy in firms applying the cost of sales method only, meaning that the implications of our results are of relevance for those firms specifically. However, to provide some insight into the potential applicability of our results also for firms applying the alternative accounting method, the nature of expense method, we do some additional analysis to investigate whether there are any structural differences in innovative performance between the two types of firms. Since the level of R&D spending cannot be established for firms applying the nature of expense method, the measure of innovative performance that we look at is the patent count.

We conduct a two-tailed paired t-test to examine whether the two types of firms' patent count means differ, and if the mean difference is statistically significantly different from zero. In order to analyze as many data points as possible, we use the raw data sample of 21,668 as described in Table 2 in section 3.4.2. We then match it with nearest-neighbor firms in a corresponding raw data sample for firms applying the nature of expense method based on total assets and industry. Since we match on size in terms of total assets, the patent count is no longer scaled by staff headcount. The results of the paired t-test are presented in Appendix 7 and show that firms applying the cost of sales method have a higher patent count mean of 0.235, compared to 0.084 for firms applying the nature of expense method, and that the difference between the two is statistically significantly different from zero at the 1% level. These results suggest that the level of innovation among firms applying the nature of expense method, which are excluded from our study, is lower than for firms applying the cost of sales method. Alternatively, their innovative work is not captured by the patent measure specifically, but by other measures of innovative output. In any case, the results indicate that the innovative performance in the two types of firms structurally differs, which implies that our results are not applicable to firms applying the nature of expense method.

To further test the applicability of our results for firms applying the nature of expense method, we run the regression R2, which only includes the independent variable $PAT\_HC_{i,t-3}$, using a sample corresponding to our previously used final sample but for firms applying the nature of expense method. This adjusted final sample includes 89,442 firm-year observations, of which 1,153 bankrupt ones. The regression results are untabulated, but present a coefficient for $PAT\_HC_{i,t-3}$ that is not statistically significant at any meaningful level. This does not align with our main results as presented in Table 6, and further supports that our results are not applicable for firms applying the nature of expense method.

5. Discussion

The aim of this study is to examine the association between innovative performance and the probability of bankruptcy, where innovative performance is measured in terms of R&D intensity, patent count as well as the patent-to-R&D ratio. When examining our results, we make several interesting findings. Firstly, we cannot reject our first null hypothesis that the intensity of innovative input, measured as R&D intensity, is not associated with the probability of bankruptcy. When conducting the robustness tests, our results hold. This suggests that a firm's tendency to go bankrupt is not related to how much money it spends on R&D.

As discussed in section 2.4.1, previous studies present conflicting results on the association between innovative input in terms of R&D spending and bankruptcy risk. Moreover, as emphasized in section 2.2.2, innovative work entails a clear trade-off for firms, where it on the one hand is of importance to be able to keep a competitive advantage, but on the other hand can be considered risky. Due to this dichotomy, our results might not be so surprising after all. On the contrary, the contradictory characteristics of innovation could possibly explain the insignificant results that we present. R&D intensity could be considered containing the most uncertainty among our three measures of innovative performance, and hence best reflects the trade-off that innovative work implies. It might be that the positive effects net the negative ones for the Swedish private firms we have studied, and thus R&D intensity neither increases nor decreases the probability of bankruptcy significantly. In support of our insignificant results, there are also previous studies that do not find any significant results for innovative input in terms of R&D spending and its association to bankruptcy (Mahmood, 2000). Due to the lack of significance in our results, no broad conclusion can be drawn regarding R&D intensity and its association to bankruptcy risk.

Furthermore, our main results show statistically significant evidence that the quantity of innovative output in terms of patent count is negatively associated with the probability of bankruptcy. This is in line with our second hypothesis as well as previous studies' results (Eisdorfer and Hsu, 2011; Bai and Tian, 2020). These results proved to be robust in all the alternations of the model, except for when the time lag is one year, which will be discussed further down in this section. Our findings indicate that it is the quantity of innovative output rather than input that is of importance for firms to mitigate the risk of filing for bankruptcy, implying that firms should focus more on the actual output of innovation rather than setting targets for R&D investment. One potential explanation to why we find support for a negative association for the innovative output, but none for the input, is that patents are realized inventions, in contrast to R&D spending, as discussed in section 2.2.1. For that reason, the patent count measure can be considered less uncertain and could thus be more significant for firm survival.

When looking at firms in technology-intensive industries only, the patent count variable no longer shows any statistically significant association with the probability of bankruptcy. A potential explanation to this might be that patents are more common among technology-intensive firms than in general and might therefore not contain as valuable information about a firm's competitive position. However, the insignificant results might as well be due to the rather small number of technology-intensive firms in our final sample, which makes the results for technology-intensive firms in particular less robust.

Our findings also support our third hypothesis that innovative efficiency is negatively associated with the probability of bankruptcy through statistically significant results for the patent-to-R&D ratio. Also in this case, the results were robust for all alterations of the model except for when the time lag is one year. This implies that firms which are efficient at transforming their investments in R&D into innovative output in terms of patents are less likely to go bankrupt. Our findings suggest that the share of innovative input that actually leads to innovative output matters more than the size of the innovative input as such for firm survival. Looking at the results for technology-intensive industries specifically, the patent-to-R&D ratio is the only measure of innovative performance that is statistically significant to bankruptcy prediction, where the results indicate a negative association. This suggests that innovative efficiency is of extra importance in comparison to the innovative input and output for firms in such industries.

To the best of our knowledge, our exact measure of innovative efficiency has not previously been studied in relation to bankruptcy risk, which makes it more difficult to compare our results to existing literature. However, as mentioned in section 2.4.1, Hirshleifer et al. (2013) find that innovative efficiency, defined in the exact same way as in our study, results in higher future earnings and stock price. This could be argued to be in line with our findings, since good financial performance normally does not lead to bankruptcy.

A possible explanation for the insignificant results for both innovative output and innovative efficiency when using the one-year time lag specifically is that innovation investments can be considered long-term strategic investments. Thus, a one-year time lag might be a too short time horizon to capture the potential long-term positive effects of innovation. Hence, the insignificant results using a time lag of one year might say more about what a reasonable prediction horizon is, rather than the robustness of the results as such. This is further supported by the worse goodness-of-fit statistics that the model using a one-year time lag presents in comparison to the other time lags.

6. Conclusion

6.1 Summary of findings

The purpose of the performed study is to investigate how firms' innovative performance impact the probability of bankruptcy. The measures of innovative performance that we look at are the intensity of innovative input in the form of R&D intensity, the quantity of innovative output in the form of patent count as well as innovative efficiency in the form of the patent-to-R&D ratio. In line with what the literature suggests, we expect an association in any direction between the intensity of innovative input and the probability of bankruptcy, and negative associations with the probability of bankruptcy for the quantity of innovative output as well as innovative efficiency.

To address the research question, we test the relationship between the three measures of innovative performance and the probability of bankruptcy using logistic regressions. Using a sample consisting of 10,335 firm-year observations, we find that the quantity of innovative output and innovative efficiency have statistically significant negative associations with the probability of bankruptcy, while we find no statistically significant results for the intensity of innovative input. These results also held for the majority of the robustness tests.

6.2 Practical contributions

Besides the contributions to the research field, as presented in section 1.4, our findings could also be of interest to various company stakeholders, such as managers, creditors, equity investors, suppliers, and customers - especially those engaged in Swedish private firms. Managers might gain a better understanding of where to allocate their resources to better avoid bankruptcy, where our results indicate that it is the actual output as well as the efficiency of innovative work that should be prioritized in the first place, rather than focusing on the size of the investments made. Creditors, equity investors, suppliers and customers might improve their judgement regarding which firms that could potentially be in the risk-zone of bankruptcy or not, where our findings suggest that innovatively efficient firms as well as firms with relatively many patents face a lower risk of going bankrupt. In addition to this, the results could be of interest to everyone that quests to better understand bankruptcy prediction in general.

6.3 Limitations

There are some limitations to our study to be aware of when interpreting our results. The single most important limitation is that data on R&D expenditure was not accessible for firms applying the nature of expense method, which forced us to exclude all such firms from the sample. This means the results may not be robust when looking at firms applying the nature of expense method, which is further supported by our additional analysis in section 4.4.2. Since the vast majority of private SMEs in Sweden chooses to apply the nature of expense method, this exclusion was of considerable size and largely affected our sample size in a negative way, which can be considered a shortcoming of the study. However, according to the paired t-test conducted in section 4.4.2, the firms applying the cost of sales method that constitute our final sample seem to be the ones where innovation is more prevalent, which implies an appropriate target group for the purpose of the study. Moreover, another limitation to the study is that it only concerns SMEs. This was done to better achieve comparability across firms in the sample, but it is important to bear this in mind when analyzing our results.

An additional limitation to the study to be aware of considers our choice of measures of innovative performance. It is difficult to capture all different aspects of innovation in just a few measures, and as discussed earlier there are several measures to choose from. Thus, our results apply to our three specific measures used, but might not apply to other measures of innovative performance. We aimed at increasing the relevance of our measures by looking at what is commonly used in previous studies, but also by choosing measures that capture different aspects of the innovation process. Some further limitations to our study concern the exact definitions of the innovation variables used. For example, due to the deficiency of data on depreciation of capitalized R&D, we had to come up with approximate figures. Of course, this affects the precision of our results. Moreover, due to unavailability of grant year data for the patents in the PAtLink dataset, we had no other choice than looking at the application year, even though the grant date is prevalent in previous studies since that is when the information becomes public. Similarly, data on patent citations was not available, which could have been a good complement to take the quality of patents into account.

Another limitation is that we recoded several bankruptcies as if they occurred the year after the last observable annual report, even though they occurred one or two years later than so in reality, in order to not miss out on too many bankruptcy observations. Though we tried to mitigate the risk of distorting the data by limiting the time gap to a maximum of three years, it is important to bear in mind that some bankruptcy observations had longer time lags in reality than what is being captured by the lagged variables used in the regressions.

6.4 Suggestions for future research

In the process of producing this thesis, several ideas on interesting topics for future research have emerged. Our study focuses on the relationship between innovative performance and bankruptcy prediction in firms applying the cost of sales method specifically, and as our additional analysis under section 4.4.2 indicates, the results might look differently for firms applying the nature of expense method. Hence, there is a great opportunity to further research the relationship for those firms in particular. If not using the exact same measures as we do, as the data collection process for those measures for such firms might be complicated, one could use alternative measures of innovative performance as well.

Furthermore, our findings in the additional analysis in section 4.4.2 imply that firms applying the two different accounting methods have different characteristics in terms of innovative performance. This suggests that when firms choose a certain type of accounting method, their choice entails information about their characteristics. Since no broader conclusions regarding this can be drawn from our results, further research regarding firms' choices of accounting method and what it depends on composes a possible topic for future research.

Some other suggestions for future research could be to examine the relationship between patent count and bankruptcy risk in Swedish private firms looking at the grant date instead of the application date, or patent citations instead of patent count. This could help improve the predictability of bankruptcy in terms of understanding which measure that is the best to use for prediction purposes. Of course, there is also an opportunity to investigate what the relationship to bankruptcy looks like for alternative measures of innovative performance on the Swedish market for private firms. To further improve the accuracy of bankruptcy prediction models, it could also be of interest to research underlying reasons for the common time gap between the last observable annual report and the bankruptcy filing. A better understanding of that phenomenon could enhance the exactness of the lagged variables used.

References

Agarwal, R. & Gort, M. 2002, "Firm and Product Life Cycles and Firm Survival", *The American Economic Review,* vol. 92, no. 2, pp. 184-190.

Altman, E.I. 1984, "A Further Empirical Investigation of the Bankruptcy Cost Question", *The Journal of Finance,* vol. 39, no. 4, pp. 1067-1089.

Altman, E.I. 1968, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *The Journal of Finance,* vol. 23, no. 4, pp. 589-609.

Altman, E.I. & Sabato, G. 2007, "Modelling Credit Risk for SMEs: Evidence from the U.S. Market", *Abacus,* vol. 43, no. 3, pp. 332-357.

Bai, Q. & Tian, S. 2020, "Innovate or die: Corporate innovation and bankruptcy forecasts", *Journal of Empirical Finance,* vol. 59, pp. 88-108.

Bhattacharjee, A., Higson, C., Holly, S. & Kattuman, P. 2009, "Macroeconomic Instability and Corporate Failure: The Role of the Legal System", *Review of Law & Economics,* vol. 5, no. 1, pp. 1-32.

Buddelmeyer, H., Jensen, P.H. & Webster, E. 2010, "Innovation and the determinants of company survival", *Oxford Economic Papers,* vol. 62, no. 2, pp. 261-285.

Campbell, J.Y., Hilscher, J. & Szilagyi, J. 2008, "In Search of Distress Risk", *The Journal of Finance,* vol. 63, no. 6, pp. 2899-2939.

Christensen, C.M., Kaufman, S.P. & Shih, W.C. 2008, "Innovation Killers", *Harvard Business Review,* vol. 86, no. 1, pp. 98-105.

Christensen, R. 1990, *Log-Linear Models,* Springer New York, New York, NY.

Eberhart, A., Maxwell, W. & Siddique, A. 2008, "A Reexamination of the Tradeoff between the Future Benefit and Riskiness of R&D Increases", *Journal of Accounting Research,* vol. 46, no. 1, pp. 27-52.

Eisdorfer, A. & Po-Hsuan Hsu 2011, "Innovate to Survive: The Effect of Technology Competition on Corporate Bankruptcy", *Financial Management,* vol. 40, no. 4, pp. 1087-1117.

Ericson, R. & Pakes, A. 1995, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *The Review of Economic Studies,* vol. 62, no. 1, pp. 53-82.

Franzen, L.A., Rodgers, K.J. & Simin, T.T. 2007, "Measuring Distress Risk: The Effect of R&D Intensity", *The Journal of Finance,* vol. 62, no. 6, pp. 2931-2967.

Hagedoorn, J. & Cloodt, M. 2003, "Measuring innovative performance: is there an advantage in using multiple indicators?", *Research Policy,* vol. 32, no. 8, pp. 1365-1379.

Hall, B.H. 1993, "The Stock Market's Valuation of R&D Investment During the 1980's", *The American Economic Review,* vol. 83, no. 2, pp. 259-264.

Hirshleifer, D., Hsu, P. & Li, D. 2013, "Innovative efficiency and stock returns", *Journal of Financial Economics,* vol. 107, no. 3, pp. 632-654.

Hsu, P., Lee, H., Liu, A.Z. & Zhang, Z. 2015, "Corporate innovation, default risk, and bond pricing", *Journal of Corporate Finance,* vol. 35, pp. 329-344.

Li, S., Shang, J. & Slaughter, S.A. 2010, "Why Do Software Firms Fail? Capabilities, Competitive Actions, and Firm Survival in the Software Industry from 1995 to 2007", *Information Systems Research,* vol. 21, no. 3, pp. 631-654.

Mahmood, T. 2000, "Survival of newly founded businesses: A log-logistic model approach", *Small Business Economics,* vol. 14, no. 3, pp. 223-237.

Merton, R.C. 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *The Journal of Finance,* vol. 29, no. 2, pp. 449-470.

Miao, H., Ramchander, S., Ryan, P. & Wang, T. 2018, "Default prediction models: The role of forward-looking measures of returns and volatility", *Journal of Empirical Finance,* vol. 46, pp. 146-162.

Ohlson, J.A. 1980, "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research,* vol. 18, no. 1, pp. 109-131.

Pandit, S., Wasley, C.E. & Zach, T. 2011, "The Effect of Research and Development (R&D) Inputs and Outputs on the Relation between the Uncertainty of Future Operating Performance and R&D Expenditures", *Journal of Accounting, Auditing & Finance,* vol. 26, no. 1, pp. 121-144.

Phillips, B.D. & Kirchhoff, B.A. 1989, "Formation, Growth and Survival; Small Firm Dynamics in the U.S. Economy", *Small Business Economics,* vol. 1, no. 1, pp. 65-74.

Thorburn, K.S. 2000, "Bankruptcy auctions: costs, debt recovery, and firm survival", *Journal of Financial Economics,* vol. 58, no. 3, pp. 337-368.

Ugur, M., Trushin, E. & Solomon, E. 2016, "Inverted-U relationship between R&D intensity and survival: Evidence on scale and complementarity effects in UK data", *Research Policy,* vol. 45, no. 7, pp. 1474-1492.

Watson, J. & Everett, J.E. 1996, "Do small businesses have high failure rates? Evidence from Australian retailers", *Journal of Small Business Management,* vol. 34, no. 4, pp. 45.

Wooldridge, J.M. 2018, *Introductory econometrics : a modern approach,* Cengage, Boston, MA.

Zhang, W. 2015, "R&D investment and distress risk", *Journal of Empirical Finance,* vol. 32, pp. 94-114.

Online sources

Bokföringsnämnden (2020) *Kategori 2.* Available at: https://www.bfn.se/om-bokforingsnamnden/k-projektet/kategori-2/ (Accessed: 5 March 2021).

Bolagsverket (2019) *Resultaträkning i årsredovisningen*. Available at: https://bolagsverket.se/ff/foretagsformer/aktiebolag/arsredovisning/delar/resultatrakning/resultatrakning-i-arsredovisningen-1.17932 (Accessed: 5 March 2021).

European Commission (2012) *SME definition.* Available at: https://ec.europa.eu/growth/smes/sme-definition_en (Accessed: 4 March 2021).

KPMG (2021) *IFRS vs. US GAAP: R&D costs*. Available at: https://advisory.kpmg.us/articles/2017/ifrs-vs-us-gaap-rd-costs.html (Accessed: 4 May 2021).

OECD (2011) *ISIC REV. 3 TECHNOLOGY INTENSITY DEFINITION.* Available at: https://www.oecd.org/sti/ind/48350231.pdf (Accessed: 26 April 2021).

PRV (2021) *Tid till besked för svensk ansökan*. Available at: https://www.prv.se/sv/patent/ansoka-om-patent/svensk-patentansokan/sa-handlaggs-din-ansokan/tid-till-besked/ (Accessed: 5 March 2021).

Riksdagen (2020) *Årsredovisningslag (1995:1554)*. Available at: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/arsredovisningslag-19951554_sfs-1995-1554 (Accessed: 5 March 2021).

Skatteverket (2021) *Egenupparbetade immateriella tillgångar.* Available at: https://www4.skatteverket.se/rattsligvagledning/edition/2021.6/324718.html (Accessed: 10 April 2021).

Appendix

Appendix 1. Distribution of final sample across industry groups

Industry sector	Observations	Proportion of sample
Energy & Environment	52	0.5%
Materials	121	1.2%
Industrial goods	1,516	14.7%
Construction industry	689	6.7%
Shopping goods	1,160	11.2%
Convenience goods	5,424	52.5%
Health & Education	151	1.5%
IT & Electronics	110	1.1%
Telecom & Media	43	0.4%
Corporate services	744	7.2%
Other	325	3.1%
All industries	**10,335**	**100%**

Appendix 2. Number of bankruptcy filings per year

Year	Active firms	Bankruptcies	Proportion of bankruptcy cases	Cumulative bankruptcies
2002	617	10	1.6%	10
2003	653	7	1.1%	17
2004	705	7	1.0%	24
2005	741	5	0.7%	29
2006	737	2	0.3%	31
2007	721	5	0.7%	36
2008	707	6	0.8%	42
2009	708	4	0.6%	46
2010	689	2	0.3%	48
2011	663	2	0.3%	50
2012	622	6	1.0%	56
2013	592	3	0.5%	59
2014	505	1	0.2%	60
2015	432	1	0.2%	61
2016	425	1	0.2%	62
2017	413	1	0.2%	63
2018	404	3	0.7%	66
2019	1	1	100%	67

Appendix 3. Industries defined as technology-intensive

SNI 2007 major group
20 - Manufacture of chemicals and chemical products
21 - Manufacture of basic pharmaceutical products and pharmaceutical preparations
26 - Manufacture of computer, electronic and optical products
27 - Manufacture of electrical equipment
28 - Manufacture of machinery and equipment n.e.c.
29 - Manufacture of motor vehicles, trailers and semi-trailers
30 - Manufacture of other transport equipment
51 - Air transport
61 - Telecommunications

Appendix 4. All industries vs Technology-intensive industries

	All industries (N=10,335)	Technology-intensive industries (N=334)
Variable	**Mean**	**Mean**
$BANKR_{i,t}$	0.006**	0.030**
$RDS\_TS_{i,t-3}$	0.019**	0.100**
$PAT\_HC_{i,t-3}$	0.005***	0.034***
$PAT\_RDS_{i,t-3}$	0.044***	0.928***
$TL\_TA_{i,t-1}$	0.644**	0.615**
$CL\_CA_{i,t-1}$	0.704***	0.514***
$INTWO_{i,t-1}$	0.060***	0.105***
$SIZE_{i,t-1}$	17.125***	17.733***
$WC\_TA_{i,t-1}$	0.226***	0.344***
$NI\_TA_{i,t-1}$	0.064***	0.028***
$FU\_TL_{i,t-1}$	0.276***	0.197***
$CHIN_{i,t-1}$	0.020	0.024
$OENEG_{i,t-1}$	0.022	0.027
$AGE_{i,t-1}$	2.898***	3.134***

*This table presents the mean for all variables for all industries and technology-intensive industries in the final sample respectively. All continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4. *,** and *** denote significance of the mean differences at the 0.1, 0.05 and 0.01 level, respectively and refer to the two-tailed t-test.*

Appendix 5. Multicollinearity test

	VIF	1/VIF
$RDS\_TS_{i,t-3}$	1.21	0.83
$PAT\_HC_{i,t-3}$	1.21	0.82
$PAT\_RDS_{i,t-3}$	1.03	0.97
$TL\_TA_{i,t-1}$	2.20	0.46
$CL\_CA_{i,t-1}$	3.20	0.31
$INTWO_{i,t-1}$	1.28	0.78
$SIZE_{i,t-1}$	1.10	0.91
$WC\_TA_{i,t-1}$	3.56	0.28
$NI\_TA_{i,t-1}$	2.25	0.44
$FU\_TL_{i,t-1}$	1.94	0.51
$CHIN_{i,t-1}$	1.27	0.79
$OENEG_{i,t-1}$	1.29	0.77
$AGE_{i,t-1}$	1.03	0.97
Mean VIF	**1.74**	

This table presents the VIF values for all variables for the final sample. All continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4.

Appendix 6. Regression results for technology-intensive industries

Variable	R1	R2	R3	R4
$RDS\_TS_{i,t-3}$	-2.735 (-1.30)			-2.200 (-1.03)
$PAT\_HC_{i,t-3}$		-18.817 (-1.43)		-6.901 (-1.29)
$PAT\_RDS_{i,t-3}$			-6.328*** (-4.55)	-6.482*** (-5.06)
$TL\_TA_{i,t-1}$	3.158 (1.52)	3.635* (1.93)	5.336*** (2.75)	3.361 (1.59)
$CL\_CA_{i,t-1}$	-5.272* (-1.82)	-4.839 (-1.61)	-5.509 (-1.58)	-5.125* (-1.77)
$INTWO_{i,t-1}$	1.553 (1.23)	1.543 (1.29)	1.056 (0.75)	1.752 (1.38)
$SIZE_{i,t-1}$	0.014 (0.03)	0.095 (0.16)	0.107 (0.18)	0.105 (0.18)
$WC\_TA_{i,t-1}$	-9.360*** (-2.89)	-8.586** (-2.56)	-7.829** (-2.22)	-8.557*** (-2.63)
$NI\_TA_{i,t-1}$	-9.357 (-1.61)	-8.778 (-1.25)	-6.595 (-0.89)	-8.784 (-1.51)
$FU\_TL_{i,t-1}$	0.952 (0.57)	1.673 (0.74)	1.378 (0.50)	0.936 (0.67)
$CHIN_{i,t-1}$	0.106 (0.21)	0.055 (0.10)	0.067 (0.12)	0.148 (0.25)
$OENEG_{i,t-1}$	-1.862 (-0.87)	-2.046 (-0.96)	-1.818 (-0.86)	-2.091 (-0.96)
$AGE_{i,t-1}$	-0.634 (-0.74)	-0.618 (-0.75)	-0.591 (-0.68)	-0.667 (-0.80)
Constant	2.698 (0.23)	0.465 (0.04)	-0.965 (-0.07)	0.863 (0.07)
Industry FE	NO	NO	NO	NO
Year FE	YES	YES	YES	YES
Observations	138	138	138	138
Log likelihood	-24.142	-23.915	-24.462	-23.382
Pseudo R^2	0.327	0.333	0.318	0.348

*This table presents the results for the four logistic regressions R1, R2, R3 and R4, as described in section 3.2, for technology-intensive industries specifically. All regressions include year fixed effects and consist of 138 firm-year observations, of which 10 are bankrupt ones. Standard errors are clustered by firm, and all continuous accounting input data required for any of the variables is winsorized at 1st percentile and 99th percentile level. The definitions of all variables are listed in Table 1 in section 3.3.4. The z-statistics are presented in the parentheses below the coefficients. *,** and *** denote significance of the coefficients at the 0.1, 0.05 and 0.01 level, respectively.*

Appendix 7. Results from paired t-test

Variable	Observations	Mean	Std. Err.	Std. Dev.
Patent count for firms applying the cost of sales method	21,668	0.235	0.020	2.881
Patent count for firms applying the nature of expense method	21,668	0.084	0.010	1.417
Difference	21,668	0.152	0.021	3.099
t-statistic = 7.2023***				
P(|T| > |t|) = 0.0000				

*This table presents the results for the paired t-test comparing the patent count mean for firms applying the two different types of accounting methods. *,** and *** denote significance of the mean difference at the 0.1, 0.05 and 0.01 level, respectively and refer to the two-tailed t-test.*