COPENHAGEN BUSINESS SCHOOL

MASTER THESIS

Long-Run Performance of IPOs

How to determinate the performance of IPOs in the long-run within the European Market

Master of Science in Economics and Business Administration
Finance and Investment

Author:
Kristoffer Hassinggaard JENSEN

Supervisor:
Christian RIX-NIELSEN

Number of pages: 79
Characters: 161.184
September 15, 2017

Abstract

The thesis examines the long-run performance of Initial Public Offerings (IPO) within the European Union and test if it is possible to predict long-run performance of IPOs.

The research is based on a data sample of firms going public in the period from 2007 to 2013. Based on a unique data set specifically constructed for this thesis an analysis is conducted of firms going public in a hot and cold market period. Furthermore, the data gives the potential to investigate the significance of different market conditions, industry, and period effects. Also, the specific impact of all predictors probability of obtaining abnormal returns is analysed.

Four main variables; Age, $\log(Size)$, ROA, and $Underpricing$ were tested in this research. The empirical results from the statistical test demonstrated that IPO performance is possible to predict, but also that some parameters have a better prediction power, and that some variables have a negative correlation while others have a positive correlation. This master thesis also indicates that firm- and market-specific variables have a significant influence on IPOs long-run performance.

Acknowledgements

I would like to thank Barbara Jeszenszky for supporting me throughout the process of finalising this master thesis. I would also like to thank Kristian Schmidt, who have given me sparring on the statistical aspect of the research and more specifically in the construction of the R-coding.

Finally, I would like to thank my supervisor Christian Rix-Nielsen for supporting me in the whole process of the construction of this master thesis.

Contents

Chapter 1

Introduction

1.1 Introduction

The first Initial Public Offering (IPO) can be detected back to 1602 when the Dutch East India Company established The Amsterdam Stock Exchange for the enterprise to deal with its printed stocks and bonds (Basti, Kuzey, and Delen, 2015). Since the 17th century, the equity markets have developed in close relation to the economic growth and the same is the case for the numbers of firms going public. Today IPOs is a large part of the financial world. The general idea of an IPO is to obtain additional financing for the firm to continue growing, but research indicates that there is no clear reason why firms go public (Brau, 2012). If there is no clear reason why companies are go public, is it still possible for investors to obtain a financial gain when investing in IPOs? If a financial gain is possible can any common characteristics between the IPOs obtaining an abnormal return be concluded?

1.2 Problem statement

As stated in the introduction IPOs have been a part of the financial world almost from its beginning meaning there is a significant amount of research and literature about the topic. Most of the literature is focused on underperformance and underpricing leaving a gap in the academic literature with a focus on over-performance.

The purpose of this master thesis is therefore, to fill out this gap by analysing the performance of IPOs in the long-run. The analysis is fundamentally built upon current research, previously published articles and a data sample from 28 European countries in the period 2007-2013. Based on the method used and the empirical data sampled the purpose is to exterminate if there is any correlation between fundamental data and IPOs obtaining abnormal returns and whether it is possible to forecast which IPOs would outperform the market.

Current research is mostly focused on the American IPO market, while a few studies have been conducted in the European market, but with a narrow focus on a single market or a single region leaving evidence for a lack in existing research within this field.

1.2.1 Question for research

Based on the problem formulation and the gap in the current academic research upon this subject gives the opportunity for this study to fill in this missing research by answering the following question:

"Is it possible to forecast which IPOs is outperforming the market based on available public information?"

To answer the question of research, four hypotheses are constructed. Yet, this research is also thought to give investors a simple set of parameters, which have a significant effect to predict abnormal returns. Each hypothesis examines different aspects of a firm, which could have an impact on the performance when being traded in the public market. These parameters cover firm specific information, market information, and issuer information. The parameters include basic company information such as size, age, and industry, but also more in depth financial information. Market information is covering two specific areas, which are the period of the IPO, where this research is covering the financial crises in 2008 to 2009, and which region the firm goes public in. Lastly, an analysis dependent on type will be conducted, so if the IPO firm is backed by Venture Capital or Private Equity, but also underpricing will be studied in correlation with long-run performance.

The parameters in this research are all based on pre-IPO information except for the underpricing, which is available after closing on the first day of trading. It is a fundamental aspect of this research that the information needed for each criterion is accessible to the public and easy to access.

The research question is tested on a data sample, which is manually collected in the period from 2007 to 2013 within the European market. The method of the study and the choice of parameters for determining which IPOs have an abnormal return is based on literature covering the different aspect of IPO theory. Both method and existing theory will be discussed in the next chapters.

1.3 Definition

In this section, the definition of the main aspects of the problem statement and question of research will be defined to ensure an equal understanding of the primary research parameters.

1.3.1 Definition of public information and outperforming

Public information is defined as information available for all potential investors before the IPO date and easily accessible. In general, the information is presented in the prospects (Jenkinson and Ljungqvist, 2001).

outperforming is measured in this research as obtaining a larger return by investing in an IPO firm than investing in a benchmark portfolio with a Buy-and-Hold Abnormal Returns (BHAR) strategy (Loughran and Ritter, 1995). A more detailed description of the BHAR model will be made in the next chapters.

1.4 Limitations

The data for this research is collected in the period from 2007 to 2013 with in the 28 European countries. This particular time is chosen due to several parameters. The first year of the period is covering a very optimistic market, which then turns directly into one of the largest global financial crises from 2008 to 2009 with almost illiquid markets and a Eurozone crisis in the following years after 2009. The cap of 2013 is because the research is using a BHAR strategy with a horizon up to three years and for firms to obtain this period the cap is set to 2013.

The purpose of this research is to evaluate if specific parameters can predict outperformance. This research heavily builds on literature on IPO underperformance and underpricing, where fundamental parameters are identified as measures of market return. Current research covers a broader aspect of performance with a higher focus on the correlation between performance and fundamentals. This research will build upon this by examining if it is possible to develop a model, which can predict outperformance based on these basic parameters.

1.5 Overview

This research paper covers six sections and has the following structure.

Section 1 ***Introduction***

The first section contains a general introduction to IPO history as well as modern theory. After the introduction, the problem statement will describe the overall purpose of the research and the thoughts behind it. The problem statement leads to the question of research where the core of the paper is introduced. Following, the definition of the core parts of the question of research is defined followed by limitations and overview.

Section 2 ***Existing theory and Empirical framework***

This section contains a thorough description of an IPO process and an analysis of the relevant literature within the subjects of IPO performance and prediction of outperformance. The first two parts of this section will provide the theoretical basis for empirical research. In the last part of the section, several hypotheses are formulated based on existing theory and empirical framework.

Section 3 ***Methodology***

The methodology section presents the research and the structure of the study. The statistical models and the theory behind them are specified, and the regression correlated to the hypothesis is constructed. A description of the parameters used to forecast outperformance is part of this section as well.

Section 4 ***Data***

The first part of the section describes the framework of the data, the data collection process, and analysis of the data sample. In the end, cross-sectional statistics is performed, and a summary of the general patterns is conducted for further analysis.

Section 5 ***Empirical results***

The results from the regressions carried out in the previous section are presented in this section. The results are analysed and discuss the hypothesis constructed and the underlying theory.

Section 6 ***Conclusion***

The last section contains the conclusion, which summarises the main results, and ensure that the question for research is answered and finally provides a suggestion for further studies.

Chapter 2

Existing theory and Empirical framework

2.1 Initial public offering

As described in the introduction IPOs are a large part of the financial markets, and from the firm's point of view, the decision to engage in an IPO process is argued to be one of the biggest milestones in a life cycle of a company. The process of an IPO can be divided into five steps, the choice of market, producing a prospectus, marketing, pricing and allocation and after the IPO (Jenkinson and Ljungqvist, 2001).

2.1.1 Why IPO

The subject of why firms are going public has received little attention in the academic world, and most entrepreneurs and CEO's are often responding to the statement that the purpose of the IPO is to obtain additional capital (Brau, 2012). Overall, an IPO process can either be a process of selling existing shares or issue additional shares. When the current owners decide to sell their existing stocks, they will give up some of the control with the firm against receiving additional equity. In the second process, the current owners retain the control of the companies, but the margin of the control is decreasing since the buyers of the new shares will enter the ownership circle. Moreover, it is common that an IPO process consists of both selling existing shares and creating new shares.

An IPO process is providing a firm with new or additional financing, which supports potential new and further growth opportunities (Huyghebaert, Van De Gucht, and Van Hulle, 2007). This supports the general assumption that the essential reason for a company to go public is to raise additional capital (Ritter and Welch, 2002). In recent years, the assumption about the main reason for why firms are going public, has been analysed, but without any clear conclusion of why companies went public (Brau, 2012) and therefore capital increase is still assumed as the main reason for firms to go public.

For existing company owners, there are multiple exit strategies such as a merger, IPO and sale. Another primary reason for an IPO, aside from raising capital, is an exit strategy, where existing shares are sold in the process (Zingales, 1995). Using IPOs as an exit strategy is common when venture capitalist (VC) and private equity investors (PE) has a significant stake of shares in a firm.

For PE investors a clear exit strategy is one of the early steps in the process, when investment opportunities are evaluated so that a potential financial gain can be obtained quickly (Schmidt, Steffen, and Szabó, 2010). When a firm is going public and backed by a PE investor, it tends to send a strong signal to the market. PE investors prefer to exit a nonperforming business as soon as possible rather than keep them in their portfolio. A firm in the process of going public, backed by a PE investor, thus sends a clear signal to the market about the PE investor confidence in the enterprise's performance and future growth opportunities (Schmidt, Steffen, and Szabó, 2010). Several studies have been conducted exploring the impact a PE investor has when firms go public. Both in the general IPO process as well as the performance of IPOs in aspects of underpricing theory and long-run performance theory (Bergström et al., 2006), (Katz, 2009). Existing literature, analysing long-run performance, concludes that firms backed by a PE investor indicate higher performance compared with both other IPOs and the market as a whole (Levis, 2011).

VC investors have a high exit turnover rate for firms within its portfolio companies, and the most common way to leave an investment is through an IPO (Admati et al., 1998). VC backed IPOs has been the main subject of several studies for an extended period, and at a greater extent than PE backed IPOs. The research conducted implies that IPOs backed by VC investors is performing better than IPOs, where the stake holders in the enterprise do not include a VC investor (Brav and Gompers, 1997). Research indicates that the reason for a higher performance of VC backed IPOs is due to increased management performance and corporate governance (Bessler and Kurth, 2007). The reputations of the VC investors also have an impact on the IPO where VC investors with a low-level of reputation in the market could offset the common positive effect a VC investors have. Therefore, it is expected that an enterprise backed by a VC investor with a high-level of reputations is sending a strong signal to the market, and the IPO firm will outperform both the market and none VC backed IPOs (Krishnan et al., 2011).

Despite the most common reasons for firms to go public, financing and exit strategy, as mention above, there are several other factors determining the decisions of going public. In the academic world studies and research has led to an agreement in which theories describe why firms are going public (Brau, 2012). The next section will touch upon some of the common theories of why firms go public.

The statement that firms go public to obtain additional financing, is further analysed by Pagano, Panetta, and Zingales (1998), which argues that by going public, companies are overcoming borrowing constraints as well as increasing the bargaining power to the primary provider of funding. By overcoming financing limitations and increasing the bargaining power, the costs of debt is decreasing and therefore the argument that enterprises are going public to obtain additional financing could have a more specific reason (Pagano, Panetta, and Zingales, 1998), this is also supported by Rajan (1992). The results from the studies confirm that firms seeking to overcome the borrowing constraints have a higher probability to conduct an IPO (Pagano, Panetta, and Zingales, 1998). There was no clear relationship between bargaining power and the possibility to carry out an IPO, but the data from the research indicated that firms did increase their bargaining power after they went public (Pagano, Panetta, and Zingales, 1998).

When firms go public, it tends to attract a significant amount of attention. This recognition is to some extent used to create a first-mover advantage, especially in the technological industries (Maksimovic and Pichler, 2001). The first-mover advantage established by going public can by accelerated with the combination of using underpricing to generate more awareness of the IPO (Demers and Lewellen, 2003). Therefore, an IPO process can be used as one part of a strategic management plan by the firm. The hypothesis that IPOs attract more attention to the enterprise is tested by Demers and Lewellen (2003), where traffic on the IPO firm web page as well as media coverage is measured pre IPO date and post IPO date. The results show a significant increase in web page visits, as well as media coverage both for a regular IPO and in the case with underpricing. The increase in awareness of the IPOs has the possibility to achieve a first-mover advantage.

Using an IPO as a part of Mergers and Acquisitions (M&A) strategy is another motivation for starting an IPO process. The IPO can be utilised for multiple parts of a M&A strategy. When a firm goes public, the valuation of the company is no longer uncertain, which compared to M&A makes the process faster, easier and less indeterminate. Valuation is normally a large part of a M&A process and the current trading price gives a good indication of the value of the enterprise (Lyandres, Sun, and Zhang, 2005). Furthermore, an IPO results in both stocks and shares which can be used to either acquire other firms or be a part of a stock deal (Brau, Francis, and Kohers, 2003). Brau and Fawcett (2006) shows that in a data sample from 2000 to 2002 of IPO transactions IPO firms become acquirers to a greater extent than compared to non-IPO firms.

There are clearly different factors, which could have an impact on the IPO process and why firms go public, as described in the sections above. Each factor represents different approaches and results in both disadvantage and advantage depending on the background of the firms and which approach is optimal for the IPO company. When analysing and researching performance of IPOs, it is important that each factor is taken into account, when cross-sectional data is analysed since each variables have a different impact on the performance across countries, markets, industries and period.

2.1.2 Where to go public

After the firm makes the decision to go public two additional new issues occurs in the process of the IPO; finding investors to purchase the stocks, and in which market the shares should be traded in. Normally, this subject has attracted fewer studies, since firms previously tended to go public on their domestic market. With the globalised world, firms have almost infinite markets they can pick upon as target market. There is no clear trend for companies to choose a domestic market or an international market. There are multiple reasons for why a particular market is preferred. It could be that the local market is relatively small, and in order to create a liquid share a large international stock exchange is preferred. Other responses could be regulatory requirements such as listing requirements, trading requirements, and accounting requirements (Jenkinson and Ljungqvist, 2001).

A well-developed and driven legal system is important when firms are considering which market to enter. In general, the legal systems consist of the Common law and the Civil law systems (Dainow, 1966). In the financial world, the most common understanding is that investors are better protected in the Common law systems than in the Civil law systems (La Porta et al., 1997). Recent research is not consistent with this assumption and argues that the protection of investors are the same, but they are handled differently (Brau and Fawcett, 2006). Within the European region, both law systems are present, but the Civil law is the most common and only the United Kingdom, Ireland, Malta, and Cyprus are using the Common law (Brau and Fawcett, 2006). IPO activity has a positive correlation between the legal origin, investor rights and law enforcement (La Porta et al., 1997), and countries using the Common law system tend to have higher IPO activity than the Civil law systems countries.

The legal systems have been proved to have a significant impact on firms not only in the IPO process, but pre-and post-IPO as well (Claessens and Laeven, 2003). Some jurisdictions favour some stakeholders of a company. La Porta et al. (2000) argues that the French Civil law system leaves room for management and majority shareholders to transfer profit out of the firm, as well as assets, leaving the minority investors with a significant disadvantage. This could result in a more illiquid share, or a lower share price to be necessary before trading is conducted. IPO regulations gives firms a set of tools to offset a weak legal system by allowing firms to go public in a different market but also make a cross-listing where the second market is characterised by having a robust legal system giving a higher protection for investors (Engelen and Essen, 2010).

Regulations of firms when going public needs to compile with both listing requirements from country specific official regulator, such as the Financial Service Authority (FSA), and specific requirements from the stock exchange. Usually, the stock exchange were related to a particular country, but in the modern financial world, this trend is not universal anymore. Most of the stock exchanges are private firms with a broad portfolio of stock markets in different countries. NASDAQ OMX is one of these firms, which own the stock exchanges in Armenia, Denmark, Estonia, Finland, Latvia, Lithuania, Sweden and Iceland. This results in a generalised regulation of stock exchange between countries (Jenkinson and Ljungqvist, 2001). The same is the case with the official regulation requirements in the various countries, especially in Europe where most of the regulation is determined by the European Union and not by each country (Lastra, 2006).

For some firms, meeting the requirements of both the FSA and the particular stock exchange is a big hurdle and requires a large cost or is in some cases it is not even possible to meet the requirements. However, it is still possible for firms to go public on a stock exchange. A listing, where the requirements from the FSA is not obtained, is possible if the stocks are traded on "unofficial" stock exchanges (grey market), where firms only have to fulfil the stock markets demands. The grey market has an increasing role in the financial world, and some of the grey markets have a high-level of reputation such as the Alternative investment market of The London Stock Exchange. The stocks traded in this market are just as attractive as stocks on the usual stock exchanges. Normally, firms which go public by an “unofficial” listing, are less mature businesses such as start-ups and growth companies. But also companies with a lack of a solid track record, or firms which are simply too small for the regular market going public in the grey market (Stringham and Chen, 2012).

Despite the fact that there are different requirements and legal systems from country to country, globalisation of the world changes the impact on the importance of picking which market to use for an IPO firm. This gives the management of the enterprise several additional options, and makes an IPO not the only option for large and old firms. The most important aspect is to choose a market where there is enough trading to ensure a liquid stock in order to have a successful IPO.

2.1.3 How to go public

After the preliminary steps, the essential parts of the IPO process start. The next steps of the process are crucial to ensure a successful IPO, where the firm has to produce a prospectus, decide the share price, and allocation of the share.

Most companies are lacking the knowledge and have no or little experience with taking firms public. Management of the enterprise has to require support for the IPO process and build a team of participants such as lawyers, underwriters, and accounting advisers and thus achieve the knowledge base able to drive the whole IPO process. Building the right team has been proved to have a significant impact on the performance of IPOs (Su and Bangassa, 2011). Typically, the team consists of a lead manager, or co-lead managers, who then construct the full team based on the specific knowledge requirements of the IPO. A lead manager is often from an investment bank, but in recent years the market for consulting IPO processes is getting more competitive. Corporate finance boutiques and accounting firms are focusing more on this market and investing heavily in building special teams who can drive the IPO process. The leading manager is then in charge of the full process, and it is essential for firms to consider experience, expertise and fess of the IPO team, but the team will also make an internal due-diligence of the IPO firm to ensure they can conduct a successful IPO.

The IPO team has several initiated tasks, which need to be executed before the final steps can be conducted. The first key aspect is to ensure that the enterprise is allowed to go public. In order to do this, the team collects key financial information as well as legal information, which is presented to the financial regulators, who has to allow the firm to go public. After the clearance with the regulators, the final part of the IPO process begins, where a prospectus is constructed and presented to potential investors and later to the public market.

An essential part of the IPO process is the setting of the share prices as desired by the company. The price setting can be estimated based on regular valuation models, such as the Discounted Cash Flow (DCF). In the prise setting process a pre prospectus is produced, where the overall analysis of the firm and the valuation is presented. The pre prospectus is then presented to potential investors as pre marketing material, and the investors will give their view of the company to the IPO team. In general, there are two different approaches to finding the share price after the valuation and the pre marketing phase; fixed price basis and book-building (Jenkinson and Ljungqvist, 2001). In the fixed price model, investors know at which price the stock is offered, but the overall demand for shares is unknown until listing. In book-building investors informing the IPO team of how many shares they would acquire at a particular price, and based on this information the IPO team can set the price. There has been several discussions and research about the different approaches to pricing the shares. Book-building has become the most popular method in recent years and according to Sherman (2005) there are several reasons for the increased popularity. The method reduces the risk for both investors and issuers since it gives more control to the IPO team. Book-building demands more work for the IPO team, and it has proved that the direct costs are twice the costs of an auction process (Ljungqvist, Jenkinson, and Wilhelm Jr., 2003).

With an enormous cost difference between book-building and fixed-price, it is important to use the best method for the IPO firm. Research shows that book-building results in less underpricing (Ljungqvist, Jenkinson, and Wilhelm Jr., 2003) but there is no clear indication of long-run performance. Degeorge, Derrien, and Womack (2007) argues that more hyped IPOs are more suited for using the book-building method. With no clear signs of which method is the best, it is important for the firm to consider the hype surrounding the IPO and the costs of going public when the price of the shares shall be established.

2.1.4 When to go public

In theory, the value a firm can obtain in the process of going public is estimated based on a fair valuation. Therefore, the period of the IPO is insignificant meaning that the total number of companies going public is independent of time (Jenkinson and Ljungqvist, 2001). But as Ibbotson and Jaffe (1975) documents, there is a cluster in the number of IPOs and Ritter (1984) supports the conclusions regarding clusters provided in previous research. The global economy is a key driver of the IPO cycle. In a period of high growth and a positive trend in the economy, increases the change of adverse selection, and more unhealthy firms are going public since they have a larger chance of a successful IPO due to the optimistic market (Yung, Çolak, and Wang, 2008). This is also known as the hot and cold market theorem and is one of the most researched topics in the field of IPOs.

A period with high numbers of IPOs is the typical indication of a hot market (Jenkinson and Ljungqvist, 2001). Yet, underpricing and over-subscription are also common, but can vary from industry to industry (Ritter, 1984). Despite a significant amount of research about the hot issue market, there are different conclusions about the firms going public in this period. Some scholars argue that in a period of hot market more high-quality companies go public (Welch, 1989), (Allen and Faulhaber, 1989), (Grinblatt and Hwang, 1989). This is because the offer price is less affected by adverse selection and findings of significant first day returns are supporting this argument (Ritter, 1984). When looking at the long-run performance, the picture is turned upside down and firms who went public in a hot market period are underperforming compared to the market, which indicates the firms had a low-quality when going public (Loughran and Ritter, 1995). The reason for an increase in low-quality enterprises in a hot market period is argued to be due to managers taking advantage of a "window of opportunity" (Loughran and Ritter, 1995). New studies are indicating that smaller and riskier companies from specific industries is a general characteristic in the hot markets. This trend is catalysed by innovation in technology and increasing productivity (Stoughton, Wong, and Zechner, 2001), (Benveniste, Busaba, and Wilhelm, 2002), (Maksimovic and Pichler, 2001). From the results of the different studies, there is no clear sign of what is driving a hot market, but firms going public in a hot market period has a higher underpricing and are under performing in long-run periods. This could indicate that periods of hot markets are driven by investors sentiment, rather than fundamental factors (Helwege and Liang, 2004).

Long periods with a high level of underpricing could shift the fundamental relationship between demand and supply within the market for IPOs. As a result of this change in the fundamentals, the market will automatically go from a hot market period to a cold market period. As mentioned above some studies show that in a hot market period a large number of low-quality firms are being listed. Eventually, the market will begin to notice the growing number of low-quality businesses, which will result in a decrease in valuation of IPOs. Therefore, high-quality companies will discard from an IPO process, since they cannot obtain a fair share value. Other studies suggests that enterprise should start an IPO process in a period with a cold market since the company will obtain a higher price if it is a high-quality firm (Ibbotson and Jaffe, 1975). New studies show that firms operating in the hot market period are generating a higher share price in the IPO process (Helwege and Liang, 2004).

As the previous section states, there is no precise alignment in the studies regarding the hot and cold market periods affect the IPO process. Some studies conclude that this is the best period to obtaining the highest price is in the hot market, while others have proven the opposite. The same is the case of the impact of the hot and cold market where some sectors have a high correlation while others have almost no correlation. Newest research indicates that a firm going public in the hot market will have a high underpricing and low long-run performance.

2.2 IPO performance

Post IPO performance has been a subject of several studies, and especially short-run and long-run performance has conducted an enormous amount of attention. Studies of short-run performance have led to one of the most familiar concepts in IPO theory known as underpricing or the ex-ante premium. Long-run performance has conducted a various number of studies as well, and is mostly captioned in the concept long-run underperformance. Jenkinson and Ljungqvist (2001) argue that underpricing and long-run underperformance are so common in research they can be considered as equilibrium outcomes. For individual companies, long-run performance is measured in most research as capital gains as well as the firm ability to stay profitable, which is aligned with shareholder theory (Freeman, Wicks, and Parmar, 2004).

2.2.1 Underpricing

Underpricing or the ex-ante premium (underpricing is the word of preference in this paper) is an indirect cost or an opportunity cost for the original owners when going public (Jenkinson and Ljungqvist, 2001). Underpricing is the difference between IPO offer price and the first day of trading closing price (Guo, 2011).

Several studies investigate the range of underpricing both in different periods and among different countries. Guo (2011) estimates that in a period from 1960 to 2006 IPOs in the American market had an equal-weighted average return of 17%. Underpricing is relatively volatile between countries, and in the period from 2000 to 2006 Norway had an average first-day return of 4%, and in the same period, China had a stunning 57% first-day return average (Banerjee, Dai, and Shrestha, 2011). The gap between underpricing in different countries is due to differences in regulations, contractual mechanisms and overall characteristics of the particular market (Fauzi, Hewa-Wellalage, and Locke, 2012).

Several studies have analysed the reason why underpricing occurs. One of the main subjects is signalling theory, where underpricing is used by the IPO firm in order to handle information asymmetry. Despite the obvious timing problem of using underpricing as a signal, since it is not available before the IPO and therefore the investor cannot use the signal proactively it is still commonly used. One of the first studies conducted in this area argues that underpricing is used as an indication of a high-quality firm (Ibbotson, 1975). Later studies suggest that companies with no information asymmetry did not experience underpricing, indicating that the company did not have any reasons for sending signal to the market (Allen and Faulhaber, 1989).

Another branch of studies has focused on underpricing as part of the adverse selection and more particularly the winner's curse theorem. Underpricing is used to compensate investors, who have limited or even no information about the real value of the IPO firm (Ritter, 1984). Studies prove that there is a correlation between limited information pre-IPO and the size of the underpricing (Beatty and Ritter, 1986). Other studies argue that the underwriters have better information about the price than the issue, since they have a larger understanding of the market behaviour and thus knowledge of the market demand for the new share (Baron, 1982). The issuing firm, therefore needs to rely on the recommended price from the investment bank supporting the company in the IPO process. Often the recommended price is below the actual market price since the investment bank has an incentive to secure they do not end up with an IPO with a negative first-day return or shares, which they are unable to sell (Ruud, 1993), (Ritter and Welch, 2002). It has been shown that in larger IPOs the underpricing is higher than average, which is argued to be because in larger transactions there is a higher level of interaction between the issuer and the underwriters (Michaely and Shaw, 1994).

Underpricing is seemingly inevitable, but the fundamentals behind underpricing seem less clear. However underpricing is only a short-run indication of an IPO performance. Since underpricing is based on the first day of trading this means both liquidity and systematic risk is irrelevant and therefore the underwriters must have the biggest impact on the underpricing, since they are setting the price for the new share (Ritter and Welch, 2002). Underpricing is, therefore, more a performance measure for the underwriters since they are the only ones who can benefit from a high underpricing (Jenkinson and Ljungqvist, 2001). Even though underpricing is a short-run phenomenon, some studies have suggested that underpricing could have a correlation with the performance of an IPO in the long-run (Ritter, 1991).

2.2.2 Long-run performance

The performance of IPOs in a longer time horizon, from one to five years, has been a subject for research in several years and numerous studies have focused on the performance issue. Generally, in long-run IPOs underperform the market, and it is so common that underperformance is the third anomaly in IPO theory according to (Ritter, 1991). Long-run performance is today known as long-run underperformance since investors, who buy shares in the first period of trading after the IPO, have a lower return compared to a benchmark portfolio or index (Fama and French, 2004a).

Several hypotheses have been constructed in the research of long-run underperformance. One of the first studies suggests that underperformance is related to the underwriters and the book-building process. As suggested by Miller (1977) the price of the new shares heavily depends on the expectation from the underwriters, where the most optimistic underwriters are setting the offer price. After going public, more information is available for the market and a share price set by an overconfident investor will be adjusted by the market mechanism, leading to a decrease in the share price. In a similar study by Miller (1977) and Shiller (1990) it is argued that investment banks are setting the offer price lower than the market value so potential investors can get a high first-day return, which will create an excess demand for the new share to ensure that the shares can be sold pre-IPO.

In the research conducted by Ritter (1991) several parameters were investigated as an explanation of the long-run underperformance, risk mismeasurement, bad luck or fads of over-optimism. The results from the studies predict that the age of the firm has a significant impact on the performance of the share; younger companies have a larger underperformance than its benchmark. Ritter (1991) also finds that companies going public in a peak of industry specific fads of over-optimism have a higher underperformance.

Profitability has in several studies been linked as one of the most fundamental facts to indicate the long-run performance of IPOs. In a research from Goergen, Khurshed, and Mudambi (2007) evidence for a relationship between profitability pre-IPO and the long-run performance as well change in ownership and size of the firm were founded. This argument is supported by Fama and French (2004a) where a study of survival rates show a clear relation between profitability and performance. It is even concluded that profitability is the most important fundamental factor. In the same research, another conclusion stated that there is an increasing trend of fewer profitability firms going public, but the growth rates of the companies have increased compared to older studies. The trend is due to the fact that there is an increase in smaller companies going public, but there is no correlation to any particular industry (Fama and French, 2004a).

In relevant theory regarding long-run performance there is a universal consensus that firms which have undertaken an IPO are underperforming the market compared to the market matched in size and industry (Ritter, 1991). Some studies have, however, shown that the particular time horizon has an impact on the performance (Ibbotson, 1975). Why IPO tends to underperform and which fundamental parameters are the underlying reason is a question with no clear answer from existing research regarding this matter.

2.3 Prediction of a long-run performance

Prediction of share performance has obtained a significant amount of academic focus. To measure the performance of companies investors can use endless parameters as well benchmarks and models. First, it is necessary to understand how investors estimate what a fair price for a share is and the relation to firm specific parameters and market parameters. In addition, the match between stock performance for an individual share and the market needs to be clear. Furthermore, the fundamental difference between models used to measure performance between the market and an accurate stock has to be distinguished.

2.3.1 Performance measurement

As stated in the section 1.1 in the introduction, stock markets has a long history, and the same is applicable for the most common valuation method today, the Discounted Cash Flow model (DCF). In the first academic discussion, research and establishment of a fundamental valuation model was established in the early 1900s (Fisher, 1907). After the stock market crash in 1929 further research was done and a final model for valuation was constructed by Fisher (1930), Williams (1938). In the same period, researchers began to investigate why companies went bankrupt. The first relevant academic literature suggests that fundamental accounting parameters could predict bankruptcy (Smith and Winakor, 1935). Their argument is based on an empirical investigation, where there is a significant difference in financial ratios between bankrupt companies and non-bankrupt companies.

After the initial research, some of the most popular models in the field of asset pricing was developed which has a closer relation between market parameters and firm specific parameters. The Capital Asset Pricing Model (CAPM), constructed by Treynor (1961), Sharp (1964), Lintner (1965), and Mossin (1966) is used to estimate the required rate of return of an asset in relation to its sensitivity. Later empirical studies indicate a lack of significance in statistical test, but the model is still used due to its simplicity (Ang and Chen, 2007), (Fama and French, 2004b). An additional two parameters were added to the CAPM model; Small Minus Big (SMB) and High Minus Low (HML), to optimise the model and make it more efficient. This is known as the Fama-French three-factor model (Fama and French, 1992). Several studies confirm the new model is more reliable in the prediction of performance than the CAPM (Gaunt, 2004). In recent years further two parameters were added to the original three-factor model due to evidence from Titman, Wei, and Xie (2004), Novy-Marx (2013). The two additional parameters, Robust Minus Weak (RMW) and Conservative Minus Aggressive (CMA) were added to construct a five-factor model and tested against the three-factor model. The five-factor model had a higher probability to predict performance (Fama and French, 2015).

In addition to the models described above there are the market-based models and the accounting-based models. Where the most common market-based models are the Black-Scholes model (Black and Scholes, 1973) and The Merton model (Merton, 1974) and the most common accounting-based models is the Altman z-score (Altman, 1968) and the Ohlson o-score (Ohlson, 1980).

Measuring of IPO performance might sound straightforward but in relevant literature several different models are used, and the performance of a particular IPO fluctuates between models (Lyon, Barber, and Tsai, 1999). The performance of IPOs is determined as a return, which is different in comparison to a benchmark such as a stock index or portfolio of a firm matching the IPO firm on various parameters (Schultz, 2003). The difference between the IPO return and a market return is known as Abnormal Return (AR) (Ritter, 1991) and is the simplest model to estimate underperformance and over performance.

Based on the simple AR calculation the Cumulative Abnormal Return (CAR), and the Buy-and-Hold Abnormal Return (BHAR) is constructed to measure performance. In prior studies, the two models are the most common to use when estimating long-run performance. According to Fama (1998), the main difference between CAR and BHAR is that an investor acquire a stock at the IPO date and holds it for a longer period after the acquisition. The return of the investment is better described by BHAR. BHAR is only looking at the performance on the two times the investors interact with the market and therefore if wealth gain is the main subject for the research, BHAR is the argued model to use. On the other hand, BHAR tends to magnify underperformance as a result of the compounding of returns (Brav, Geczy, and Gompers, 2000). Brav, Geczy, and Gompers (2000) also claim that BHAR has a larger likelihood to make a type I error, which is a rejection of the null hypothesis of no abnormal return when the hypothesis is true. Therefore CAR is a better model to use from a statistic point of view (Fama, 1998). Despite the statistical disadvantage of BHAR, the model is still the preferred estimator of abnormal returns in recent studies.

2.3.2 Financial parameters

Although the first models of performance measurement were established more than a century ago, there is still no final result as to which parameters should be used in prediction of return. Though there is a displacement of variables and some variables are more rapidly used in empirical research than others. With a set of financial parameters, performance can be measured in a quantitative way by comparing the financial parameters with other firms' parameters and previous performance of the firms to estimate improvements. A company can use one set of parameters to track performance and investors can use a significantly different set of parameters to determine if a firm is performing as the investors had expected. In a sample of 190 papers analysing bankruptcy prediction, 93% used financial ratios as the fundamental parameters in their models (Jardin, 2009). In the case of IPOs, it is hard to get the necessary financial information for the investors since IPOs do not necessarily provide a long historical record of financial reports. Some misleading information could occur in the financial presentation provided by the firms going public pre-IPO date, by using cost-cutting techniques and manipulating accounting methods in order for the firm to look better than what is actually the case (Teoh, Welch, and Wong, 1998).

Performance measured by financial ratios are generally divided into profitability, liquidity, efficiency and debt (Ak et al., 2013). Investors can use a broad range of parameters to evaluate firm performance and B. Chin, D. Gun, and Wai (2010) argue that there are 48 ratios which can be used to analyse a company. Although a significant focus from investors is on financial ratios and financial fundamentals Savsar and Karaca (2012) have estimated that only 20% of a change in an enterprise value is correlated with financial ratios.

When investors analyse firm performance two financial ratios are primarily used: Return on Equity (ROE) and Return on Asset (ROA). Both ratios are characterised as profitability financial ratios. ROE is a measure of profit in relation to shareholder equity and is used by investors to measure how well firms use investments to generate growth for its shareholders and the efficiency of profit generation to shareholder equity. ROA is a measure of profit in relation to asset and is an indication of how the asset is generating profit for the firm. ROE and ROA were constructed by Donaldson Brown in the early 1900s and has been a measure of performance ever since (Kaplan, 1984). Several studies have focused on the liability of ROE and ROA as a sufficient indicator of efficiency for firms. ROA as an indication of bankruptcy was introduced by Beaver (1966) who argues that ROA is the second best financial ratio to predict financial failure and ROA could be used as an early indication of bankruptcy up to five years prior to the bankruptcy date. This is supported by several later research (Zmijewski, 1984), (Beaver, McNichols, and Rhie, 2005). Both ROE and ROA has shown positive correlation in predicting bankruptcy as well as predict stock prices and performance (AlOmoush and Al-Shubiri, 2013), (Skogsvik and Skogsvik, 2010).

Other significant parameters to touch upon is current assets to current liabilities ratio, and total liabilities to total assets. Current assets to current liabilities ratio are useful for investors to check a firms capability to fulfil its short- and long-run obligations, and it is one of the most common parameters in bankruptcy prediction known as liquidity ratio (Altman, 1968), (Beaver, 1966). Total liabilities to total assets is a ratio which investors use to detect less healthy enterprises, and the ratio is increasing when firms are approaching a bankruptcy (Altman, 2002).

2.3.3 Firm and market parameters

Despite the fact that financial ratios are setting the stage for a firm to be successful it is still necessary for investors to consider some basic company and market parameters, when analysing a business. In recent years firm and market parameters have gained more interest not only from investors, but also in academic research, where studies are indicating a trend that this set of parameters has a more significant prediction power of firm performance than previously. Some of the most common fundamentals cover industry, market volatility, size and age of the firm.

Firm age is one of the parameters most research has concentrated on. The average age of firms when going public has since the beginning of the 20th-century changed dramatically from an average of ten years before the 1990s to an average of three years since 2000 (Jovanovic and Rousseau, 2001). A way to understand the decrease in age is to look at the phase before the IPO as a trade-off between internal learning for the firm and opportunity cost related to postponing the date for the firm to go public (Jovanovic and Rousseau, 2001). Therefore, Jovanovic and Rousseau (2001) argue that the lower average age of firms entering the market could indicate a cluster of IPO firms. This is unusually promising since they do not need the internal learning before going public. It comes to be the total opposite of what several studies from pre and early 1990s shows. In a paper by Muscarella and Vetsuypens (1989) it is shown that IPOs with a higher risk require larger average initial returns, and age has a significant correlation with the risk of a firm. This is supported by Ritter (1991) who finds evidence of different performance between industries, where those industries with the lowest average of age was underperforming significantly more, compared to industries with a normal or high average of age. Newer studies confirm that younger firms are underperforming even when the data sample is based on the period from 1991 to 1997 (Clark, 2002) where Jovanovic and Rousseau (2001) argue younger companies would perform better than the historical average. Despite the fact that there are different views on correlation between age and performance all studies show that there is a correlation nonetheless (Clark, 2002).

In close relation to the age of an enterprise is the size of the enterprise which also has undertaken an extensive amount of research. In an empirical research based on data from 1970 to 1990 in the US, Loughran and Ritter (1995) find that size has a significant impact on IPO performance. This is supported by several bankruptcy models which all find evidence for a correlation between size and bankruptcy (Altman, 2002), (Shumway, 2001). The same results are confirmed in newer research with evidence from the European market and from the American market (Cogliati, Paleari, and Vismara, 2011), (Ritter, 2011). All the studies mentioned above agree on the fact that larger companies are indicating better performance in comparison to smaller enterprises.

Different performance among industries is a common assumption in the financial world (Ritter, 1991). Several studies have continued investigating industry difference and how industries are performing amongst each other. One of the characteristics of the industry that have proved to have an impact on performance is if the industry is diversified. Industries with a high level of competition tend to have a higher bankruptcy rate (Peristiani and Hong, 2004). Another study focus on industries with a high degree of regulation and found a high level of correlation between regulation and IPO performance (Akhigbe, Borde, and Whyte, 2003). Also, the life cycle has been proved to have an impact on performance. A study showed that there is a correlation between industrial age and performance, resulting in a significantly better performance of IPOs, which went public in a more mature sector than IPOs going public in younger industry (Ang and Boyer, 2009).

The industry in which a firm is operating is difficult for the firm to change, but which market they decide to go public on is optional. Therefore, it is interesting for a researcher to see if they can find any correlation between countries and IPO performance. Some of the largest markets for IPOs are the European market and the American. Several comparisons have been made between the two markets. In a research conducted by Ritter (2003), it is touched upon how the local financial crises can have a tormented impact on IPO performance from market to market such as the Euro crisis post the global financial crisis. As mentioned earlier in this paper, country specific law systems and regulations can also have an impact on where firms decide to go public. In similar studies research have shown that companies going public on a forging stock exchange are performing better, but if the home market has a weak legal system there is an advantage of going public on a foreign market offset (Bell, Moore, and Filatotchev, 2012). Empirical research indicates that even between the 28 European countries there are significant differences in performance (Berk and Peterle, 2015). Research, in general, argues that country or region is closely correlated with IPO performance.

2.4 Hypotheses

Based on the theoretical framework presented above and the gaps in the current literature, the following section will frame several hypotheses. The hypotheses will set the foundation for this paper to answer the question for research.

Several empirical studies have shown a correlation between firm age and long-run performance. Most of the studies conclude that younger firms are underperforming in comparison to older firms (Ritter, 1991). Younger firms have a higher risk, as older companies have better financial records, access to better financing, higher reputation, and a stable foundation (Ritter, 1991). This is supported by Peristiani and Hong (2004) arguing that firm age is one of the best pre-IPO key ratios to predict long-run performance. When IPO performance is evaluated in the bankruptcy setting, older companies have a lower risk of failing within the first period after an IPO than younger enterprises (Demers and Joos, 2007). Despite this almost all empirical studies show younger firms are underperforming, but still firms going public are getting younger and younger. Age of companies entering the market via an IPO is therefore still interesting to look into and explore whether age is still a contributing factor to IPO performance. The following hypothesis has thus been constructed.

Hypothesis 1:
There is a positive relationship between firm age and long-run abnormal returns.

The size of a company is in theory correlated with how successful and efficient it is, and its power related to competitors and bargaining power to financial providers. Larger firms are less correlated with country and industry volatility since they have a more international and cross sectional operation. Among the first to show a significant power of firm size and performance are Loughran and Ritter (1995). When measuring size, most empirical studies are either using the total asset or revenue (Shalit and Sankar, 1977). But with an increase in asset light and small revenue companies going public, it is interesting to see if size is still correlated with long-run performance.

Hypothesis 2:
There is a positive relationship between firm size and long-run abnormal returns.

Profitability is one of the main financial ratios in bankruptcy prediction and stock performance in general and is broadly used in the financial world when analysing stocks. Several studies have confirmed that profitability has a significant indication of stock performance (Skogsvik and Skogsvik, 2010). As mention above ROE and ROA was constructed almost 100 years ago and is still widely used in the financial world. It is therefore interesting to investigate if profitability still has a strong correlation to performance in the stock market.

Hypothesis 3:
There is a positive relationship between firm profitability and long-run abnormal returns.

Underpricing is associated with several different underlying parameters of why underpricing occurs: differences in regulations, market characteristics, firm quality, and information asymmetry (Ibbotson, 1975), (Fauzi, Hewa-Wellalage, and Locke, 2012), and (Allen and Faulhaber, 1989). Underpricing is related to short-run performance, but studies have indicated a correlation between underpricing and long-run performance as well (Ritter, 1991).

Hypothesis 4:
There is a negative relationship between underpricing and long-run abnormal returns.

2.4.1 Dummy variables

In addition to the four primary hypotheses, several dummy variables are constructed and will be tested in relation to the four primary hypotheses in order to obtain a more significant result from the primary hypotheses. The dummy variables consist of the following:

Dummy variable 1:

Indication of which specific industry the company is operating in.

Dummy variable 2:

Region indication of where the IPO is listed.

Dummy variable 3:

Indication of whether or not the companies going public is backed either by a venture capitalist or a private equity found.

Dummy variable 4:

The period the firm went public was characterised as a time of a "Hot" or "Cold" market period.

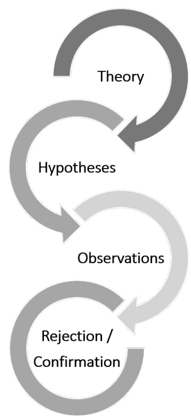
Chapter 3

Methodology

One of the fundamental assumptions in finance argues that it is not possible for investors to obtain an abnormal long-run return since the market automatically will adjust and that there is no such thing as a free lunch (Friedman, 1975). In contrast to this assumption, empirical research indicates that investors are able to gain an abnormal return in the long-run. The long-run abnormal returns have a high sensitivity to the method used to measure the return (Brav, Geczy, and Gompers, 2000) and by appointing different methods, the long-run abnormal performance are no longer present (Fama, 1998). The most essential part of measuring abnormal return is to provide a definition of normal returns (Fama, 1970). Since the method applied for measuring abnormal returns has such a large impact on the results, and as abnormal returns results are a part of the fundamental design of the research, an in-depth analysis and description of the measurement method is conducted. The following chapter will analyse and describe the method used to answer the constructed hypotheses from the section above.

3.1 Design of the research

FIGURE 3.1: The deductive research approach[2]



The research in this paper uses two fundamental methodologies; a deductive approach and the explanatory research. The deductive method is used to build the overall lines through the study, while the explanatory method is applied in the test of the hypotheses constructed in section 2.4. The two approaches complement each other, which increases the quality of the research conducted.

The deductive framework consists of four steps as illustrated in figure 3.1. According to Bryman and Bell (2015) a hypothesis or multiple hypotheses arise from existing theory and context. The first step is to build a fundamental knowledge base by studying and analysing previous papers and articles. After the hypotheses are stated, an empirical research

[2]The figure shows the four main stages in the deductive approach (Bryman and Bell, 2015)

of the hypothesis is conducted. In this step, data related to the subject is collected and analysed by using existing and relevant theory. When the data is interpreted, the explanatory method is used to ensure the right conclusions are obtained. Bryman and Bell (2015) argues that the last step in the deductive approach is more inductive, since the results from the empirical research are set up against the existing theory used in the process of building the hypothesis.

The explanatory method is designed to find a relationship between factors (cause) and performance (effect) (De Vaus, 2001). A crucial part of the process is to understand the difference between causal explanatory relationship and correlation between the two variables: explanatory and dependent, where both methods can indicate a significant effect (De Vaus, 2001). Lacking the capability to identify and distinguish between the two reasons for a given effect and use it in the research will result in a conclusion, which is only observing rather than explaining the effects (De Vaus, 2001).

The research conducted in this paper is designed in line with the deductive approach (Bryman and Bell, 2015) supported by the explanatory research (De Vaus, 2001). First, a comprehensive review and in-depth analysis of theory, empirical research, and fundamentals in relation to IPOs are outlined. Based on the findings in the theoretical framework several hypotheses are constructed to answer the question of research in an accurate matter. The hypotheses are tested based on a data set collected specifically to this paper. The main trends in the data are described and the data is subject to an empirical research, where the results are inferring with existing research and theory. Based on the findings from the study the constructed hypotheses are either rejected or confirmed, and the result is connected to the research question.

3.2 IPO performance parameters

The variables used in this research are based on existing research showing that the variables have a significant impact on IPO performance. Despite the fact that existing literature already confirms the parameters' correlation to performance, the financial world is changing from year to year. Research with the newest data available could change some of the fundamental assumptions about the correlation between some specific parameters and IPO performance. The following section presents an overview of different parameters related to either the firm going public or the market in that period. In this paper, a firm parameter is defined as a variable that is only related to the company, for example region and industry are categorised as firm parameters.

3.2.1 Firm parameters

According to the first hypothesis constructed previously, *Age* of the firm has a significant impact on long-run IPO performance, which is in line with several studies mentioned in the previous section.

Age is defined as the number of years from the foundation of the firm until the day of the IPO $Date_{Listing} - Date_{Established}$ and is measured in numbers of years. In general, age of a company indicates how successful an enterprise is and has been historically, since only the strongest and healthiest can stay competitive over a longer period. In line with this Ritter (1991) have proved that younger IPOs are outperformed by firms with a higher age when going public. One thing is how age is related to performance within the general performance measurement of IPOs, but age has a strong indication of IPO performance over all as well (Peristiani and Hong, 2004). Age is expected to have a positive relation to long-run performance, but performance between younger and more mature enterprises could point in the direction of higher performance by more mature companies.

The parameter $\log(Size)$ is calculated as the natural logarithm of the total asset of the firm measured pre-IPO, in thousand €, *log(total asset)*. Size can be estimated in several different ways and unfortunately, studies have proved fluctuation between each method (Al-Khazali and Zoubi, 2005). Despite the difference there is more consistency between the two most common methods, which are based on total asset or revenue, and several studies indicate a clear correlation between size and IPO performance (Loughran and Ritter, 1995). $\log(Size)$ is expected to have a positive effect on IPO performance.

A firm's financial performance is closely related to the performance in the stock market and profitability is assumed as one of the parameters indicating stock market performance. Profitability is in this paper indicated as ROA. ROA is used due to its unique track record related to the subject of IPO performance, both in a negative context, such as IPO failure (Beaver, McNichols, and Rhie, 2005) and in long-run performance (Skogsvik and Skogsvik, 2010). ROA is defined as $ROA = \frac{EBIT^3}{\textit{total asset}}$ and measures a firms ability to generate profit based on their assets. Several studies indicate that companies with greater efficiency and earning ability, which ROA is a proxy for, has higher performance than firms less efficient and a lower earning ability (Chemmanur and Paeglis, 2005). It is expected that ROA has a positive relation to long-run performance.

Underpricing is a parameter, which is present due to numerous underlying parameters such as information asymmetry, the period of the IPO, firm quality, market characteristics, and regulations. As mentioned in the previous section underpricing is a short-run phenomenon and its occurrence can be due to several factors. Despite this underpricing has shown to have a significant correlation with long-run performance (Ritter, 1991). Underpricing can be estimated based on the first day of trading and up to three months of trading. In this study underpricing is measured as the initial return from the first day of trading and is calculated as the difference between the first day of stock return minus the first day of market return $\textit{First day stock return} = \frac{\textit{closing price - offer price}}{\textit{offer price}}$

[3]Earnings before interest and tax

- *First day market return* $= \frac{\textit{closing price - opening price}}{\textit{opening price}}$. Existing literature indicates a higher degree of correlation between underperformance and a negative long-run performance. The area of correlation between underpricing and long-run performance has received a relatively small amount of focus from the academic world. Despite the lack of studies combining underpricing and long-run performance it is still expected to see a negative relationship since the underlying fundamentals, which are the reason behind underpricing, normally have a negative correlation with long-run performance.

In addition to the variables listed which are used in testing the hypothesis, a set of dummy variables is included in the regression. By including a set of dummy variables it is possible to use the same regression on multiple groups. The first three variables are firm specific, and the last one is related to the market.

The first dummy variable is an $Industry$ dummy, which is covering 15 different industries. The variable is taking the value of 1 if the company going public is operating in the particular industry and 0 for all other industries. A more detailed description of the $Industry$ variable is made in section 4.3.2. From existing literature there is a mixed assumption of the effect industries have on IPO performance (Ritter, 1991), (Fama and French, 2004a). However, newer studies have indicated that more mature industries have a higher performance (Ang and Boyer, 2009). It is expected to see a difference in performance between industries, but industries with the same characteristics would have a smaller difference.

The second dummy variable is a $Region$ dummy, which contains a set of four different regions within the European union. A more detailed description of the $Region$ variable is obtained in section 4.3.3. When a firm from the data sample belongs to one of the four regions, it will be assigned the value NE, SE, EE, or WE, which is an indication of each of the regions, constructed to cover the European market. The reason to include a dummy variable is mainly due to two factors. The number of IPOs conducted in the period varies significantly from region to region. To take this into account, a regional variable is added to the regression. The second reason is due to the fact that there is lack in existing research, where typically only one country or one region is used when analysing long-run performance. This paper analyses several regions and the dummy variable indicates which region is analysed.

The final firm specific parameter dummy variable is a $VC/PE\ backed$ dummy variable. This variable is only consisting of one group and will take the value of 1 if the company going public is supported by either a Venture capitalist or a Private equity firm. Otherwise, the dummy variable takes the value of 0. Several studies have confirmed that businesses underpinned by a venture capitalist are performing better than none venture capitalist supported IPOs (Bessler and Kurth, 2007). The same is the case when an enterprise is backed by a private equity investor and studies have even indicated that PE backed companies perform better than VC backed firms (Levis, 2011). Both in the studies of VC and PE backed IPOs the is conclusion that having a professional investor as a significant stakeholder sends a strong signal to the market and is used by the market as a guarantee of the quality of the company going public.

3.2.2 Market parameters

The definition of a firm and a market parameter as interpreted in section 3.2 leaves the dummy variable $Hot\ market$ as the only parameter characterised as a market variable. It is possible to argue that the $Hot\ market$ variable is under the control of the firm as well, but as described in section 2 there are multiple steps and decisions which need to be taken before the actual date of the IPO. From the decision of going public is taken to the IPO date there is a period which could be up to several years, and as we saw with the financial crises in 2008 to 2009, the outlook of the financial markets can change quickly. Withdrawing or postponing an IPO is possible in order to avoid the down time in the market. But firms using these opportunities and later going public have a significant underperformance compared to firms who went through the IPO despite the difficult period (Dunbar and Foerster, 2008). The data employed in this paper covers a period with significant economic differences due to both the global financial crises in 2008 and 2009 and the Euro crisis in the following years (Ritter, Signori, and Vismara, 2012). The Hot market variable takes the value of 1 if the IPO is obtained in a hot market period and 0 if the firm goes public in a period of a cold market. The number of IPOs conducted in each year determines if the market is defined as hot or cold based on the data collected. The time of the financial crises 2008 and 2009 is a clear cold market, but the Euro crisis is a bit more blurry since the global economy had a positive trend after 2009 and the European Union was trying to eliminate the effect of the Euro crisis resulting in only 2012 as a cold market year. More specific overview of the numbers of IPOs in each year are described in section 4.3.1 as well as which years are defined as hot or cold market periods. Loughran and Ritter (1995) argues that enterprises going public in a hot market period has a high short-run performance, but a negative long-run performance. Therefore, it is expected that the hot market dummy has a negative correlation with long-run performance.

3.2.3 Event time and Calendar time

In IPO performance evaluation occurs with two different approaches; event time or calendar time. The event time method is used when the long-run performance of individual IPOs in the data sample is analysed, and each event (IPO) is considered individually and weighted separately. The calendar time method is used when a period of IPOs is explained (Schultz, 2003). The majority of existing literature on long-run performance is based on the event time approach (Levis, 2011), (Ritter, 1991), (Goergen, Khurshed, and Mudambi, 2007). Despite event time being the most common method it still has some drawbacks. Underperformance is higher when estimated with event time according to Schultz (2003). This is in line with Espenlaub, Gregory, and Tonks (2000) who argue underperformance is less evident when calendar time is used compared to event time. Also statistically, there are some drawbacks. Test in the event time environment has more power than in calendar time according to Loughran and Ritter (1995). Therefore, the event time approach is the most suitable for this research and is thus the method of preference in this paper.

Also, as mentioned throughout the previous sections in this paper there is a large cluster of IPOs in specific periods (Hot and Cold markets). This is a natural effect in the market and not a market inefficiency argued by Schultz (2003) and the event time approach disregards clustering of data in time.

The purpose of this paper is to investigate characteristics of individual IPOs and measure the effect on long-run performance, and event time helps facilitate this.

3.2.4 Benchmark

When measuring performance, several different models can be used as described in section 2.3.1. All the models are characterised by the capability of measuring abnormal returns. In the process of measuring abnormal returns a benchmark is constructed, which will return the normal return (Fama, 1970), (Loughran and Ritter, 2000). Both Ritter (1991) and Brav, Geczy, and Gompers (2000) indicate IPO performance is significantly correlated with the underlying benchmark.

In existing literature, both a broad market index or a constructed portfolio of companies matched by size and industry is used to adjust IPO returns. If a broad market index is used as the benchmark, the index will include the company going public which will make the statistical test bias towards a higher explanatory power (Loughran and Ritter, 2000). If the index is cleaned, it is still argued to be inappropriate since there would be a mismatch of systematic risk of the firm and the index (Jenkinson and Ljungqvist, 2001).

The benchmark applied should have as high as possible correlation with the IPO firm according to Schultz (2003). Therefore, the characteristics and risk of the benchmark and the company should be aligned when constructing the portfolio used to measure the standard return. This is confirmed by Ritter (1991) who uses size and industry, which also tends to be the trend in general empirical research of IPO performance.

On the other hand, if the benchmark used is matched by size it is not possible to test if the size has a significant impact on IPO performance. Abnormal returns are widely used when the performance of investments managers' performance is evaluated, and bonuses and fees are calculated. Therefore, investors will construct or argue to use a benchmark, which they are 100% sure they will be able to outperform.

Taking all this into consideration the STOXX 600[4] index is used as the benchmark in this paper. The reason is that this index covers 600 firms within the European region and consists of

[4] Full description of the STOXX 600 index can be find by using the following link `https://www.stoxx.com/index-details?symbol=SXXP`

small-, mid-, and large cap enterprises. Therefore, it is a well-diversified portfolio, where all parameters mentioned above can be estimated up against without any major correlation between the parameters or presence of IPOs firms in the index.

3.2.5 Long-run abnormal returns

As mention earlier, the performance of an enterprise and how to measure the performance can be obtained in numerous ways. In general, IPO performance is measured either by operating as presented by Jain and Kini (1994) or as the share price traded in the stock market (Ritter, 1991). Obviously, the performance of the stock is correlated with the operation, but the two approaches have a significant difference. Operating performance is also known as accounting based performance and is evaluating past performance of the company and is therefore backwards looking. Meanwhile, the share price performance is forward looking since the price is determined by the expectations for future earnings as mentioned in section 2.3.1. With the focus of this research on IPO performance in respect, to a pre-defined benchmark as explained in section 3.2.4 is the share price as a measure of the performance is used as the parameter.

After determining how performance is measured (operating or stock) and which benchmark used to estimate the normal return, abnormal returns are calculated. As mentioned in section 2.3.1 several methods can determine the abnormal return, in this paper, both the CAR and BHAR are used. Stock and market returns are mathematical calculated as:

$$r_{it} = \frac{I_t}{I_{t-1}} - 1 \tag{3.1}$$

$$r_{mt} = \frac{M_t}{M_{t-1}} - 1 \tag{3.2}$$

Where I and M is the stock and market price beginning of each trading month with t determine the period $t = 1, ..., T$ and $T \in \{12, 24, 36\}$ is the trading horizon. i and m is indicating the individual IPO and benchmark.

CAR is estimated in three steps: first, the abnormal return for each individual IPO is estimated, ar_{it}. The return of each individual stock, r_{it}, and the corresponding benchmarks return, r_{mt}, is calculated. The two returns are then subtracted from each other.

$$ar_{it} = r_{it} - r_{mt}$$

Subsequently, all the estimated abnormal returns are summed up and the average return is calculated with respect to the numbers of IPOs in the event period.

$$AR_t = \frac{1}{n}\sum_{i=1}^{n} ar_{it}$$

The average benchmark-adjusted returns, AR_t are summarised for each event month, which results in the cumulative abnormal return.

$$CAR_{q,s} = \sum_{t=q}^{s} AR_t \tag{3.3}$$

Furthermore, if an IPO firm is delisted before the period of one, two, or three years the IPO and the benchmark is the time from the start of the month until the delisting day in the month (Ritter, 1991). The CAR model described above is the general model, but in this paper the model will be adjusted. The adjustment consists of estimating AR_t as the summed average for the individual IPO over the period of trading. The adjusted CAR model is expressed as:

$$CAR_i^T = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - r_{mt}) \tag{3.4}$$

To test the statistical significance of CAR a t-statistics is based on the Crude Dependence Adjustment Test presented by Brown and Warner (1980) and as applied by Goergen, Khurshed, and Mudambi (2007) for corrections of cross-sectional dependence in the data sample applied. Cross-sectional dependence could affect the number of rejection due to clustering of IPOs in periods and lower the number of individual IPO event month behaviour (Brown and Warner, 1980).

$$t - stat = \frac{CAR_i^T}{\sqrt{t\big[\sum_{t=1}^{T}(\overline{AR_t} - (CAR_{36}/36))^2\big]/(35)}} \tag{3.5}$$

With CAR_i^T representing the cumulative adjusted return in the period of the event t and $\overline{AR_t}$ the adjusted return in the same period. CAR_{36} corresponds to the same as CAR_i^T but defines the long-run CAR. The t-statistics above corresponds to a period of three years and will be adjusted to match both one and two years.

BHAR is estimated as the return from the day purchases of the stock is conducted until the stock is sold again, the firm is delisted, or end of the period performance is evaluated, which in this paper is up to three years (Cogliati, Paleari, and Vismara, 2011).

$$BHAR = \left(\prod_{t=1}^{T} (1 + R_{it}) \right) - \left(\prod_{t=1}^{T} (1 + R_{mt}) \right) \tag{3.6}$$

Again, as in the CAR model, delisting of firms before the period of either one, two, or three years has to be taken into account in the BHAR model. BHAR handles delisting with cutting its total return on the date of the delisting and the BHAR model with delisting is (Loughran and Ritter, 1995).

$$BHAR = \left(\prod_{t=1}^{min[T,delist]} (1 + R_{it}) \right) - \left(\prod_{t=1}^{min[T,delist]} (1 + R_{mt}) \right) \tag{3.7}$$

R_{it} is the estimated return of stock i in the period t and R_{mt} corresponds to the return of the benchmark in the same t period.

BHAR is expected to have a positive skew leading to distribution sample of t being negatively skewed, which results in biased significance levels in the lower tail test (Lyon, Barber, and Tsai, 1999). BHARs compounding of returns could catalyse the measured performance in particular years (Jelic, Saadouni, and Wright, 2005) which leads to the positive skew of BHAR. According to Lyon, Barber, and Tsai (1999), this will result in a positive bias of the probability of making a type I error and reject the null hypothesis on a false basis. To assess the bias and increase the statistical significance of BHAR a skewness adjusted t-statistic is suggested by Lyon, Barber, and Tsai (1999) and applied in the same fashion as Goergen, Khurshed, and Mudambi (2007).

$$t - stat = \sqrt{n}\left(S + \frac{1}{3}\hat{\gamma}S^2 + \frac{1}{6n}\hat{\gamma} \right) \tag{3.8}$$

Where

$$S = \frac{\overline{BHAR}}{\sigma(BHAR)}$$

and

$$\hat{\gamma} = \frac{\sum_{i=1}^{n}(BHAR_i - \overline{BHAR})^3}{n\sigma(BHAR)^3}$$

Where $\overline{BHAR}$ is the average of the Buy-and-Hold Abnormal Return in the sample, and $\sigma(BHAR)$ is the cross-sectional standard deviation of abnormal returns in the data sample, with n as the number of firms in the data sample.

3.3 Statistical models

In current research, the decision of which statistical regression to use has received almost no attention despite mismatch between the regression model, and the data and the purpose of the research can lead to insignificant results and conclusion. The next section will briefly outline some basic statistical models and decision behind using a logistic regression model in this paper.

3.3.1 Specification

When specifying which model to use, the data and the setting of the research are the determining factors. As in this paper the focus is on if the IPOs are performing better than the market rater than how high the overperformance is.

The purpose of this research is to predict if a set of parameters can indicate if IPOs are performing better than the market, i.e. the dependent value is above zero or below zero. This can be written in terms of indicator function:

$$1_{\{x>0\}} = \begin{cases} 1 \text{ if } x > 0, \\ 0 \text{ if } x \leq 0. \end{cases} \tag{3.9}$$

Based on the purpose of this paper the logistic regression model is applied and more specific the logit model. The logit regression model is mathematically expressed as:

$$\begin{aligned} Pr(Y = 1|X_1.X_2, ..., X_k &= F(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \upsilon_i) \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki})}} \end{aligned} \tag{3.10}$$

Where F is the cumulative standard logistic distribution function, which has a specific form expressed as the final part of equation 3.3. The coefficients in the model are estimated by maximum likelihood. With the use of the maximum likelihood the statistics and the confidence intervals are constructed as in a regular regression since the maximum likelihood coefficients are consistent and approximately normally distributed (H. Stock and W. Watson, 2010).

In this paper CAR and BHAR are variables, which can takes value in the interval $(-\infty, \infty)$, the logit model will therefore in this article assign any CAR and BHAR values below zero with 0 and all values above zero with 1.

3.3.2 Selection

The stepwise method is used, when the logistic regression is applied to test the constructed hypotheses. The stepwise procedure can consist of forward selection, backward elimination, and bidirectional elimination. Both the forward selection and backward elimination will be used in different parts of the regressions constructed in section 3.3.3. The forward selection regression starts with only one variable in the first regression and tests the significance level of that variable. In the backward elimination, all variables are tested at the same time and then one variable is eliminated at the time, depending on the significance level. The reason for using both methods is due to the advantage and disadvantage of each method. The advantage of using backward elimination is that the joint predictive capability of all variables is estimated. This is not possible when only the forward selection method is used, since this method will overlook the joint predictability. By using the forward elimination and backward elimination, a more detailed knowledge of each variable can be obtained, since a variable is estimated to have significance in the forward selection regression, but not in the backward elimination.

When constructing the model with the highest prediction power the starting point is to include all the regressors, and then apply the Akaike's Information Criterion (AIC) approach. The AIC approach considers both the fit of the model and the number of parameters used in the model. This approach makes it possible to construct the simplest model while still considering the explanatory power (Akaike, 1973).

$$AIC(p) = -2l(\hat{\beta}; x) + 2p \tag{3.11}$$

Where l is the log likelihood, $\hat{\beta}$ is the estimated coefficients, x is the the indicator function, and p is the p-dimensional parameter vector. The estimated AIC value roughly equals the number of parameters minus the likelihood of the overall model and therefore, the smaller the value, the better the model.

3.3.3 Regression formulation

In the following section, a number of regressions are developed for the purpose of analysing the relationship between some independent variables and long-run IPO abnormal returns. The independent variables include financial, firm, and market parameters, which are defined and discussed in section 2.3.2 and 2.3.3. The regression is specified to ensure the capability of testing the hypotheses established in section 2.4.

The variables in the regressions presented below have previously been tested in several studies, but in a setting with multiple regions and industries and several other variables the effects of the parameters capability to indicate performance are more unknown. Therefore, it is assumed that there is no bias related to existing studies when constructing the regressions and analysing the results from the regressions.

A set of dummy variables defined in section 2.4.1 is included in the regression to test for fixed effects such as time, industry, region, and financial backing. The dummy variables are tested with backward elimination and by examining the dummy variables where the difference across the fixed parameters can be controlled.

Since the variables tested and specified in the hypothesis cannot be assigned to a specified category a null hypothesis will be constructed to test the hypotheses.

The dependent value is the prediction probability of abnormal returns, and is either estimated with CAR or BHAR. If the value is below zero the IPO firm is underperforming and vice versa for a value above zero. Where i is the first IPO and T is the period of the analysis. Each stock are assigned with a set of dummy variables; industry, region, market, and financial backed. A general mathematically expression of the dependent value estimated with CAR is:

$$E[1_{\{CAR_i^T > 0\}}] = 1 * P(CAR_i^T > 0) + 0 * P(CAR_i^T \leq 0) = 1 * P(CAR_i^T > 0) \tag{3.12}$$

The expression is similar when estimated based on BHAR.

The first regression is a test of 2.4, and it includes Age. The regression is set up in the following way:

$$\log\left(\frac{P(CAR_i^T > 0|Age)}{1 - P(CAR_i^T > 0|Age)}\right) = \alpha + \beta_{Age} * Age_i \tag{3.13}$$

Where the dependent variable is the probability of achieving abnormal returns, estimated with both CAR and BHAR, and α is determined as the intercept, β_{Age} are the coefficient estimates for the variable Age. The regression is testing if Age has a significant impact on long-run abnormal returns and in line with existing theory as mentioned in section 2.3.3, it is expected Age has a positive relationship with IPO performance. The purpose of the first sub regression, 3.4 is to test the results from regression 3.3 when a set of dummy variables are added to the regression, industry, region, VC/PE backed, and "hot" market a more detailed conclusion can be made based on the results from the regression.

$$\begin{gathered}\log\left(\frac{P(CAR_i^T > 0|Age, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}{1 - P(CAR_i^T > 0|Age, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}\right) = \\ \alpha + \beta_{Age} * Age_i + \delta_{Industry} * D_{Industry} + \delta_{Region} * D_{Region} + \\ \delta_{VC/PE} * D_{VC/PE} + \delta_{Hot} * D_{Hot}\end{gathered} \tag{3.14}$$

Where $\delta_{Industry}$, δ_{Region}, $\delta_{VC/PE}$ and δ_{Hot} are the coefficients estimated for each dummy variable. It is expected that Age will still have a positive relation to abnormal returns, but the significance level will fluctuate between each dummy variable.

To test the second hypothesis, 2.4 a similar regression as 3.3 is formulated. The second regression includes $\log Size$, where $\log(Size)$ is determined as the natural logarithm of the total asset as described in section 3.2.1. It is expected that $\log(Size)$ is a significant parameter to predict long-run abnormal returns for IPOs, which is in line with several studies 2.3.3.

$$\log\left(\frac{P(CAR_i^T > 0|\log(Size))}{1 - P(CAR_i^T > 0|\log(Size))}\right) = \alpha + \beta_{Size} * \log(Size)_i \quad (3.15)$$

Again the dependent variable is the probability of achieving abnormal return, estimated with both CAR and BHAR, and α the intercept, β_{Size} are the coefficient estimates for the variable $\log(Size)$.

A set of dummy variables are added to the second regression, industry, region, VC/PE backed, and "hot" market to construct a second sub regression. The second sub regression shall test $\log(Size)$ within multiple parameters to verify the results from the simple regression 3.5.

$$\begin{gathered}\log\left(\frac{P(CAR_i^T > 0|\log(Size), D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}{1 - P(CAR_i^T > 0|\log(Size), D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}\right) = \\ \alpha + \beta_{\log(Size)} * \log(Size)_i + \delta_{Industry} * D_{Industry} + \delta_{Region} * D_{Region} + \\ \delta_{VC/PE} * D_{VC/PE} + \delta_{Hot} * D_{Hot}\end{gathered} \quad (3.16)$$

The dummy variables added are in line with the first regression with estimated coefficients of $\delta_{Industry}$, δ_{Region}, $\delta_{VC/PE}$ and δ_{Hot}. The $\log(Size)$ is expected to have different effects between the variables.

Hypothesis three 2.4 argues that there is a positive relationship between profitability and long-run abnormal returns. To test this hypothesis a regression with the parameter ROA is developed. ROA as a prediction of profitability is based on previous theory, 2.3.2, arguing that ROA is a suitable proxy for profitability.

$$\log\left(\frac{P(CAR_i^T > 0|ROA)}{1 - P(CAR_i^T > 0|ROA)}\right) = \alpha + \beta_{ROA} * ROA_i \quad (3.17)$$

Where the independent variable ROA is defined as $ROA = \frac{EBIT}{\text{total asset}}$, β_{ROA} is the coefficient estimate for the profitability. The dependent variable y_{it} is the abnormal return calculated with both CAR and BHAR, and α is determined as the intercept.

As in the previous two test of hypotheses one and two a set of dummy variables is added, $\delta_{Industry}$, δ_{Region}, $\delta_{VC/PE}$ and δ_{Hot}.

$$\begin{gathered}\log\left(\frac{P(CAR_i^T > 0|ROA, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}{1 - P(CAR_i^T > 0|ROA, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}\right) = \\ \alpha + \beta_{ROA} * ROA_i + \delta_{Industry} * D_{Industry} + \delta_{Region} * D_{Region} + \\ \delta_{VC/PE} * D_{VC/PE} + \delta_{Hot} * D_{Hot}\end{gathered} \quad (3.18)$$

It is expected that ROA has a significant effect on long-run abnormal return both when it is tested separately and with the inclusion of the four dummy variables.

The last regression directly corresponds to one of the stated hypothesis, testing the effect of underpricing on long-run abnormal returns. Despite underpricing being a short-run phenomenon empirical research has proved a connection to long-run performance as well, and it is expected the following regression will confirm this assumption. Opposite the last three hypothesis and regression where the measured parameters are assumed to have a positive correlation to abnormal returns it is expected that underpricing has an adverse impact on long-run abnormal returns.

$$\log\left(\frac{P(CAR_i^T > 0|Underpricing)}{1 - P(CAR_i^T > 0|Underpricing)}\right) = \alpha + \beta_{Underpricing} * Underpricing_i \tag{3.19}$$

Where $\log\left(\frac{P}{1-P}\right)$ is estimated with both CAR and BHAR and α is determined as the intercept, $\beta_{Underpricing}$ are the coefficient estimates for the variable $Underpricing$. The same set of dummy variables is added to the regression, and the last sub regression is constructed. As mentioned in 2.2.1 several parameters have a significant impact on underpricing, and some of these parameters are represented in the dummy variables such as δ_{Hot} and δ_{Region}.

$$\begin{gathered}\log\left(\frac{P(CAR_i^T > 0|Underpricing, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}{1 - P(CAR_i^T > 0|Underpricing, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}\right) = \\ \alpha + \beta_{Underpricing} * Underpricing_i + \delta_{Industry} * D_{Industry} + \delta_{Region} * D_{Region} + \\ \delta_{VC/PE} * D_{VC/PE} + \delta_{Hot} * D_{Hot}\end{gathered} \tag{3.20}$$

In existing theory and empirical research of long-run abnormal returns the parameters, which are tested in this paper, have been tested in several different markets and over different periods, but it is uncommon to check the parameters together, resulting in a lack of testing of imperfect multicollinearity. In this paper, a regression is set up with all the parameters mentioned above including the dummy variables as well, with the purpose of testing the correlation between the parameters. First a simple is model constructed where only the four parameters are tested and then a second model will be constructed including all the dummy variables.

$$\begin{gathered}\log\left(\frac{P(CAR_i^T > 0|Age, \log(Size), ROA, Underpricing)}{1 - P(CAR_i^T > 0|Age, \log(Size), ROA, Underpricing)}\right) \\ = \alpha + \beta_{Age} * Age_i + \beta_{Size} * \log(Size)_i + \beta_{ROA} * ROA_i + \\ \beta_{Underpricing} * Underpricing_i\end{gathered} \tag{3.21}$$

$$\log\left(\frac{P(CAR_i^T > 0|Age, \log(Size), ROA, Underpricing, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}{1 - P(CAR_i^T > 0|Age, \log(Size), ROA, Underpricing, D_{Industry}, D_{Region}, D_{VC/PE}, D_{Hot})}\right)$$
$$= \alpha + \beta_{Age} * Age_i + \beta_{Size} * \log(Size)_i + \beta_{ROA} * ROA_i +$$
$$\beta_{Underpricing} * Underpricing_i + \delta_{Industry} * D_{Industry} + \delta_{Region} * D_{Region} +$$
$$\delta_{VC/PE} * D_{VC/PE} + \delta_{Hot} * D_{Hot} \tag{3.22}$$

Where the independent and dependent variable is specified in the same manner as in the previous regressions. The result from the last regression is expected to show a significant correlation between Age and $\log(Size)$ as it must be assumed that older firms have a larger base of asset than younger firms.

3.3.4 Test

The next section will specify the statistical test of the parameters from the regression above to obtain the right results and conclusions.

When testing the hypothesis, it is crucial to understand the underlying assumptions and statistic. To ensure this a four step approach is used: Null hypothesis construction (Step 1), Significant level (Step 2), choose a test statistic (Step 3), and compute the test statistic (Step 4). The four phases sets the necessary fundamentals, and depending on the outcome of the four phases a more accurate approach is used.

Step 1: Based on the formulation of the hypothesis the null hypothesis is tested against the two-sided alternative hypothesis.

$$H_0 : \beta_1 = 0 \; vs. \; H_1 = \beta_1 \neq 0 \tag{3.23}$$

Rejection H_0 means that the regression has a significant impact on long-run abnormal returns.

Step 2: The second steps are to determine which significance level we would like to test. The most common in statistic related is on a 95% level. Whereas the 99% level is the most common degree in the pharmaceutical industry where precision is more crucial. In the test of the hypothesis in this research, a 95% significance level is used as the baseline.

Step 3: The third steps are to analyse which test statistic to use for the test. Based on the paragraphs it is concluded a t-statistic, which is best suitable.

Step 4: The last steps are correlated with the actual test of the null hypothesis H_0 in the setting of this paper. The test of H_0 requires three steps.

First is the standard error of $\hat{\beta}_1$ computed. The $SE(\hat{\beta}_1)$ is an estimator of $\sigma_{\hat{\beta}_1}$, the standard deviation of the sampling distribution of $\hat{\beta}_1$.

$$SE(\hat{\beta}_1) = \sqrt{\sigma^2_{\hat{\beta}_1}} \tag{3.24}$$

Where

$$\sigma^2_{\hat{\beta}_1} = \frac{1}{n} * \frac{\frac{1}{n-2}\sum_{i=1}^{n}(X_i - \overline{x})^2 \hat{u}_i^2}{\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x})^2\right]^2} \tag{3.25}$$

Secondly the t-statistic is computed.

$$t = \frac{\hat{\beta}_1 - \beta_0}{SE(\hat{\beta}_1)} \tag{3.26}$$

The third step is to compute the p-value. In large sample such as the assumed data of this paper, the t-statistic are under the H_0 approximately distributed as a standard normal random variable and therefore, the p-value is estimated.

$$p - value = Pr(|Z| > |t^{act}|) = 2\Phi(-|t^{act}|) \tag{3.27}$$

Where a p-value is less than 5%, when tested on a 95% significance level, provides evidence against the null hypothesis. Meaning the hypothesis is rejected at a 5% significance level.

All the following steps are automatically estimated in a statistical program but to analyse the outputs and make the right conclusions an understanding of the underlying calculation is necessary.

In addition to the test statistic described above the Receiver Operating Characteristic curve (ROC) will be used to evaluate the quality of the model estimated from the regressions as well as the models type one and two errors will be analysed. The ROC curve is a graphical way of analysing the performance of a binary model. The ROC curve is plotting the true positive rate (TPR) against the false positive rate (FPR), also known as the sensitivity and specificity.

$$\begin{gathered} sensitivity = P(positive\, result|condition\, positive) \\ sensitivity = P(Y \leq j|x = 0) \end{gathered} \tag{3.28}$$

$$\begin{gathered} specificity = P(negative\, result|condition\, negative) \\ specificity = P(Y \leq j|x = 1) \end{gathered} \tag{3.29}$$

Where the curve is plots the points for j=0,1,....,c. ROC connects the point from (0,0), which is j=c, to the point (1,1) which is j=0. If a point falls below the ROC line a positive prediction is better estimated by guessing randomly. Therefore, the better the model the more space between the ROC line and the estimated curve. This is known as the AUC and is estimated by:

$$A = \int_{\infty}^{-\infty} TPR(T)(-FPR'(T))dT \tag{3.30}$$

Chapter 4

Data

The data used in this study is collected manually and is uniquely created to investigate the hypothesis described in section 2.4. The following chapter will describe the process used to create the data set and give a statistical overview of the data. In empirical research, the quality of the data has the largest impact on the quality of the conclusion. Therefore the following chapter does not only provide a description of the collecting process but will also provide a discussion of assumptions, methods and limitations through the process.

4.1 Collection criteria

In the following collecting process, an IPO is defined as the event of shares being introduced to the public stocks exchange for the first time(Jenkinson and Ljungqvist, 2001). If a firm has been listed previously or traded on another public stock exchange it is defined as a secondary listing and is therefore not included.

The sample of IPOs in this study has two main limitations, region and period. As the purpose of this paper is to investigate the IPO long-run performance of firms going public in the European market, where the European market is defined as the 28 countries in the European Union[5]. Companies operating in countries outside the European market but going public on an exchange in the European market is in included in the data sample.

The period investigated is from 2007 to 2016, meaning firms going public in the period from 2007 to 2013 are included in the data sample. The time limitation is set based on macroeconomic effects on the IPO market in the period where both pre and post the Global financial crises and Euro zone crisis years is included. This is to ensure the data collected is not biased towards the crises. The upper limit is set to 2013 to ensure that companies going public in 2013 can be included in the three-year post IPO performance analysis.

[5]list of the members can be fund in the following link `https://europa.eu/european-union/about-eu/countries_en#28members`

4.2 Collection process

The sample contains a large number of observations where information is collected from financial databases and manually collected from IPO prospects and other publicly available reports, measured in euros. Since the purpose of the paper is to investigate pre-IPO variables prediction power of long-run performance, information is only included if it is available before the IPO date. The data collection is based on four different databases, $Zephyr$, $Orbis$, $Thomson\ One\ Banker$ and $Bloomberg$.

Zephyr is used to generate the fundamental data set used as the baseline in the data collection. The data collected from Zephyr is based on three main criteria; deal type, target region and period. The data sample collected from Zephyr contains 2.474 firms, who went public in the period from 1st of January 2007 to 1st of January 2014 within the 28 countries in the European Union. The 2.474 IPOs collected from Zephyr is set as the base, and the further collection process of data shall assign different variables to each IPO in order to construct the final data sample.

From Zephyr the following variables are assigned to each of the 2.474 firms; firm name, ISO 3166-1 country code, BvD ID number, IPO date and Offer price. The first set of variables is mainly used to identify each deal so it is possible to merge the data collected from different databases. Zephyr is used as part of obtaining pre-IPO financial information and the variables collected are; EBIT and total assets.

In the second step The BvD ID from Zephyr is imported to Orbis in order collect the remaining variables. In the data sample from Zephyr 21 firms in missing the BvD ID and are discarded from the data sample, which then contains 2.453 companies. Orbis was not able to match and identify 95 IPOs based on the BvD ID number, and the sample numbers used Orbis is then 2.358 IPOs. In the data pulled from the Orbis database pre-IPO financial information, financial ratios, enterprise identification information and IPO process information is obtained. Specifically, where Country ISO code, ISIN number, IPO date, Date of incorporation, NACE Rev. 2 sector code, EBIT, and Total asset collected from Orbis. The overlap in data collected from Zephyr and Orbis is to ensure the quality of the data since differences in the data points are investigated further.

After the merge of the two data samples from Zephyr and Orbis the first part of finalising the data is constructed. The sample of 2.358 firms is lacking several parameters, which are necessary for the test of the hypothesis. The missing information is attempted collected by Bloomberg. For identification of each IPO deal in Bloomberg the Ticker symbol is used. The ticker is constructed based on the ISIN number from Orbis. From Bloomberg the following parameters are collected; IPO date, Offer type, Offer price, Venture capital or Private equity backed, EBIT and Total asset. After the data is merged with the original data an additional 789 IPOs are eliminated due to missing almost all variables, and the sample then contains 1.569 firms.

Thomson One Banker is used to collect additional information regarding whether the IPOs are venture capital or private equity backed. The information is obtained by using the two collection criteria time and region and the VC/PE variable as criteria. Thomson One Banker provides a list of 545 IPOs the list is match up against the 1.569 IPOs and only 135 is matched and merged.

After the collection of pre-IPO data and company information, the post-IPO data collecting begins. Data collected in this part is stock information and is collected from Bloomberg. The stock information collected consists of the closing price on the first day of trading, monthly closing prices up to 36 months, delisting day and closing price at delisting day. Out of the sample of 1.569 firms the closing price on the first day of trading is missing for 115 IPOs and is excluded. When stock prices are collected from Bloomberg stock prices are only available if the stock has been traded within the last 30 days. In the remaining 1.454 firms, several stock prices were missing, and 172 deals are eliminated. The elimination is justified with the argument that if stock is not traded in the last 30 days it is illiquid and not categorised as a public traded stock.

For the remaining 1.282 observations the manual collection process starts of the missing variables and finalising of the data. Three variables contained the highest number of missing information, the offer price, EBIT and total asset. In total 405 data points were missing and the information was manually collected by using information from the prospectus. The prospectus was either gathered from the firm web page or the stock exchange where the company was traded or is traded. From this process, 93 variables out of 405 were found, and the final data set used in the regression contains of 970 firms.

Collecting pre IPO data is correlated with a significant amount of uncertainty and incorrect information. The reason for the lacking of the quality of the data is mainly due to the fact that data pre of IPOs is harder to obtain for the financial databases in comparison to regularly traded firms. Despite the fact that firms are publishing pre-IPO information in the prospectus, the quality of the prospectus is varying. To avoid the lack of quality and lack of information regarding the firm going public three different databases are used giving the data the highest possible quality. The finalising of the data is further described in the next section 4.3.

4.3 Data description

In this section, an overview of the data sample is provided including a detailed description of the finalising of the data. The main statistical characteristics of the data sample is pointed out and discussed to provide a profound fundamental knowledge of the trends in the data.

4.3.1 Activities

In the first section, the activities of IPOs in the data sample is analysed. In figure 4.1 the number of IPOs from January 1, 2007, to December 31, 2013, is presented to give an indication of the trend of the IPO activity and determine which years are defined as Hot and which are defined as $Cold$. The figure also includes the total number of IPOs in the market determined by Paleari (2014) and the total number is included to ensure that trends which define the definition of the Hot and $Cold$ periods are not correlated with which firms were eliminated in the data collection process. As we can see from the diagram the trend from the data sample is almost in line with the total number of IPOs over the period. 2013 is the only year where there is no match between the trend from the entire market and the data sample. Based on the total market numbers, 2013 seems to be more a natural year and the data sample indicates it is a Hot market and 2013 is therefore categorised as a Hot period. The years 2008, 2009 and 2012 are defined as $Cold$ markets. 2007, 2010, 2011 and 2013 are therefore defined as "Hot" markets where 2007 is the most active year with almost 273 IPOs. By including the total number of IPOs from the European market it

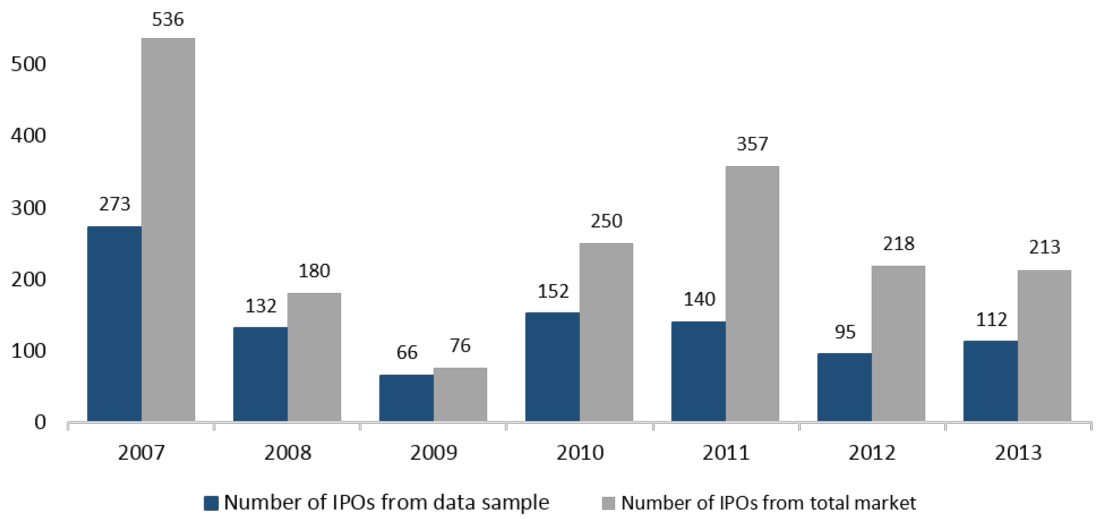


FIGURE 4.1: Number of IPOs by year in the data sample.

is possible to examine any potential bias obtained in the sample data set constructed as described in section 4.2. Bias in the constructed data set could easily occur since the elimination of IPOs from the data set only focuses on missing variables without paying any attention if more IPOs are deleted from a particular period, country, region and other specific characteristics.

Figure 4.2 presents the number of delisting for each year in the period from 2007 to 2016 of firms from the data sample. From the plot, there is a clear growth trend from 2007 to 2012, and from 2012 to 2016 there is a high but stable number of delisting and a total of 327 delisting. This figure includes all firms delisting meaning that even companies which are delisting after the longest period of trading of three years, which this paper analyses, is included. Therefore figure 4.3 is constructed to analyse how many firms are delisting before one, two or three years of trading for their own separate period. The graph shows that for each additional year of trading the number

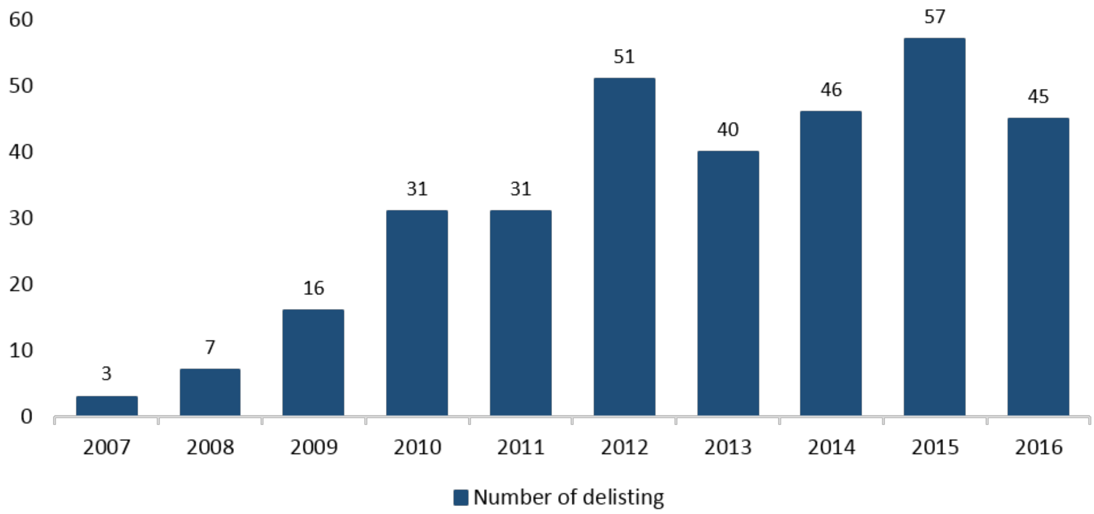


FIGURE 4.2: Number of IPOs delisted by year in the data sample.

of firms being delisted doubles. In the time horizon, up to three years of trading, the total number of delisting is only 106 enterprises from the data sample. As mentioned in both section 2.1.4

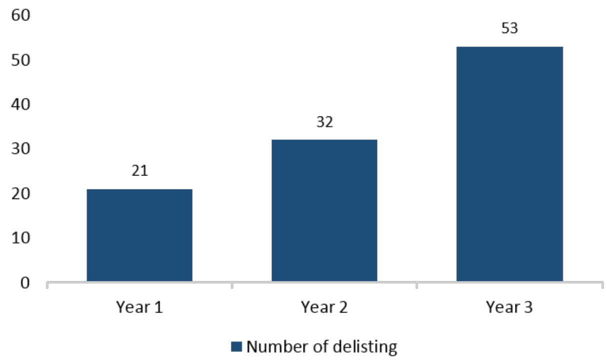


FIGURE 4.3: Number of IPOs delisted before traded for one, two and three years.

and section 2.2.1 argues Ritter (1984) that underpricing is correlated with the volume of IPOs. In figure 4.4 the number of IPOs in each quarter is plotted together with the average performance of trading on the first day after going public. This graph provides several interesting information. First, it confirms the trend of correlation between underpricing and volume of IPOs. Secondly, we can analyse the underpricing trend in the data sample and its relationship to other aspects. The graphs show that when examining underpricing for each year individually the highest underpricing occurs in Q3 or Q4 except for 2011 where it is the opposite. Also, 2011 is the year with the highest underpricing of the first-day return of 79% and the lowest with -33% first-day return. This could indicate that the IPO market in Europe had a bias towards specific events in that period.

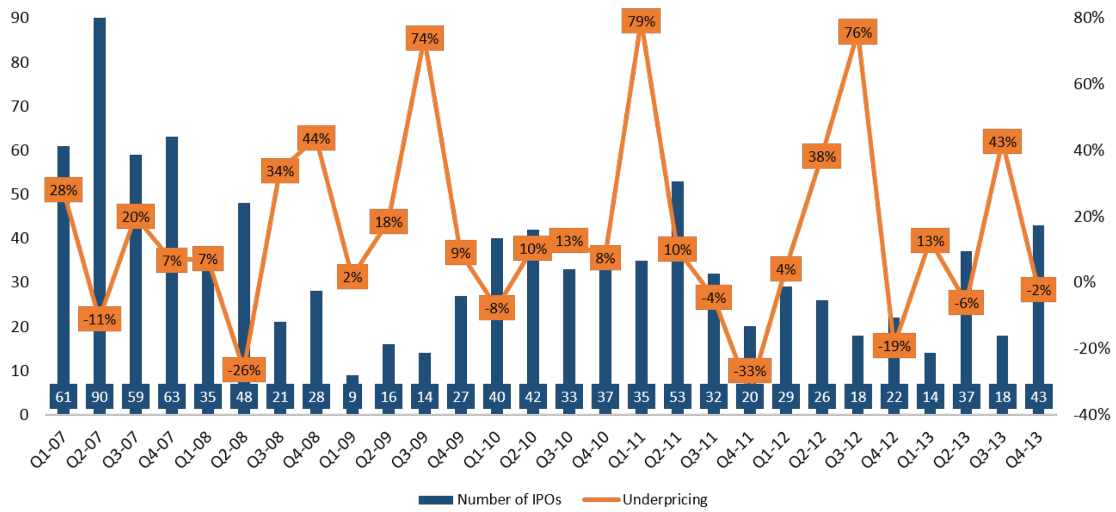


FIGURE 4.4: Number of IPOs by quarter and underpricing.

4.3.2 Industries

As part of the data collecting process described in 4.2 where each IPO in the data sample assigned with a four digit NACE Rev.2 industry code. The NACE Rev.2 industry covers 21 industries in total and is mostly used in America. Therefore, the industry codes are converted to ISIC industry codes. Table 4.1 provides an overview of the sector parameter divided into each ISIC category, the number of observation and a definition of the industry code.

The distribution of IPO activities across the industry is as the following:In 2007 Manufacturing had the highest number of 89 and a share of 32,6% of the total IPOs in 2007. Information and Communication had 42 observations, and a proportion of 15,4% and Financial and Insurance activities had 25 IPOs corresponding to 9,4% of the total market corresponding to the second and third highest activity respectively. The same ranking occurred in 2008 with a market share of 28,0%, 21,2% and 12,1% separately for Manufacturing, Information and Communication and Financial and Insurance activities. In 2009 and 2010 the three industries with the most IPO activity Manufacturing, Information and Communication, and Wholesale and Retail trade; Repair of Motor vehicles and Motorcycles, with a market share in 2009 of 24,2%, 21,2%, and 13,6% respectively and for 2010 the distribution is 28,3%, 14,5%, and 13,8%. The ranking for 2011 is identical to 2007 and 2008. In 2012 Manufacturing, and Information and Communication are the two largest industries while the third most active industry is shared between Wholesale and Retail Trade; Repair of Motor vehicles and Motorcycles, and Financial and Insurance activities with a market share of 11,6% for each sector. In 2013 are the two largest industries are the same as in all the previous years while the industry with the third highest level of IPO activity is Construction with a market share of 9,8%. It is therefore concluded that despite the difference in the distribution of the over all IPO activity the distribution between the different sectors are the same. In table 4.1 the total

TABLE 4.1: Description of industry variables and number of IPOs divided into each categories.

ISIC category code	No. of. obs.	ISIC definition
A	17	Agriculture, Forestry and Fishing.
B	52	Mining and Quarrying.
C	292	Manufacturing.
D	39	Electricity, Gas, Steam, and Air conditioning supply.
F	71	Construction.
G	91	Wholesale and Retail trade; Repair of Motor vehicles and Motorcycles.
H	21	Transportation and Storage.
I	9	Accommodation and Food service activities.
J	162	Information and Communication.
K	92	Financial and Insurance activities.
L	23	Real estate activities.
M	55	Professional, Scientific, and Technical activities.
N	26	Administrative and Support service activities.
Q	9	Human health and Social work activities.
R	11	Arts, Entertainment and Recreation.
Total no. of. obs.	970	

number of IPOs are subdivided between each sector. It is evident to see that the overall distribution of IPO activities between each sector are uneven and some industries have a relatively small number of observations. Therefore, a bias could occur when sectors with a small number of IPOs are analysed. The data sample also has a general bias because not all industries are represented in the data. The ISIC industry codes cover in total 17 sectors, but only 15 are represented and the two missing industries are Education and Water supply; Sewerage, Waste management, and Remediation activities.

4.3.3 Regions

Out of the 28 countries in the European union, 27 countries are represented in the data sample and only Slovakia is not represented. The total distribution of IPOs by countries is presented in the figure 4.6 below in the period from 2007 to 2013. As it is seen in the illustration the distribution of IPOs among the countries in the data sample fluctuates from 1 IPO in Latvia to 226 in Poland. Out of the 27 countries the six nations with the highest number of IPOs represents 83,4% of the total numbers of IPO. When comparing the distribution trend of IPO activity split by countries from the constructed data sample to the entire European IPO market, the trend in the European market data is in line with the constructed data (Paleari, 2014). Poland has the highest share of the total IPO activity with of 23,3% which is a bit higher when compared to the total market share where Poland has 15.4% share, and the United Kingdom has the largest proportion with 15,7%. This results in a bias in the data set towards Poland and the region of Eastern Europe, but no further steps will be taken to avoid the bias. To prevent the lack of activities in separate countries,

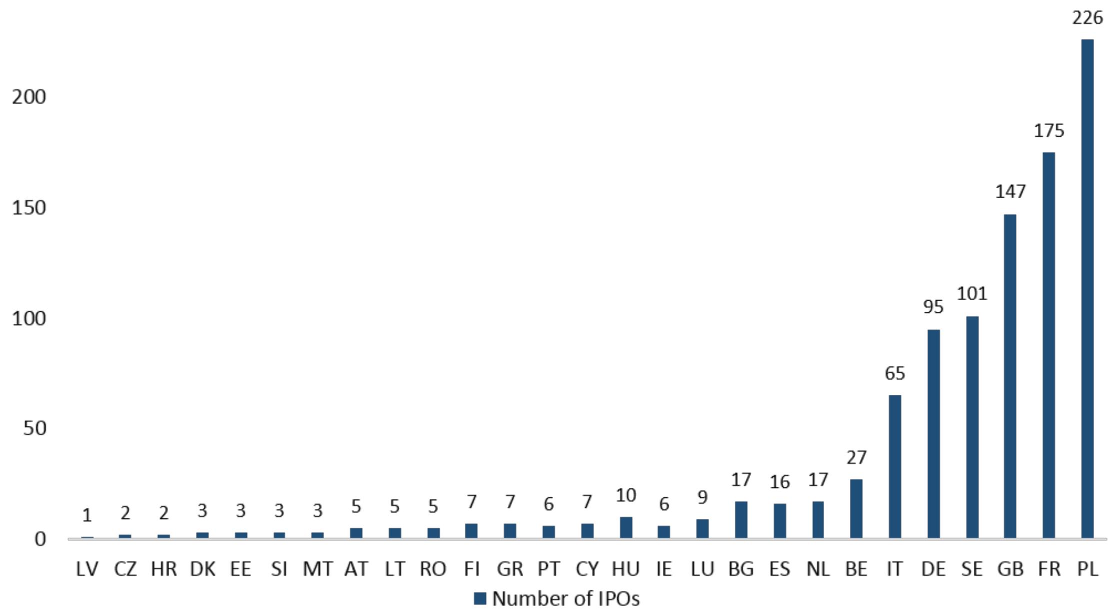


FIGURE 4.5: Number of IPOs by country.

the 27 countries are pooled together in four different regions Northern Europe (NE), Southern Europe (SE), Eastern Europe (EE) and Western Europe (WE). Northern Europe includes the North Germanic, Finnic, and Baltic countries. Southern Europe counts Portugal, Spain, Italy, Greece, Malta and Cyprus. In the Eastern Europe region is Poland, Check Republic, Romanian, Bulgaria, Croatia, Hungary and Slovakia. The last region is Western Europe and covers French, Germany, Great Britain, Belgium, Austria, Ireland, Luxembourg, and the Netherlands. Figure 4.6 shows the distribution of IPO activity in the four areas defined above. There is an uneven distribution with a large gap from Southern Europe with 104 IPOs to Western Europe with 481 IPOs. The Eastern European region has a total of 265 and with Poland's total of 226 the country contributes with 85,2% of the total market share in the Eastern European region. Although a relatively small number in two of the regions it is still assumed that the areas are large enough to avoid any bias of the region variable.

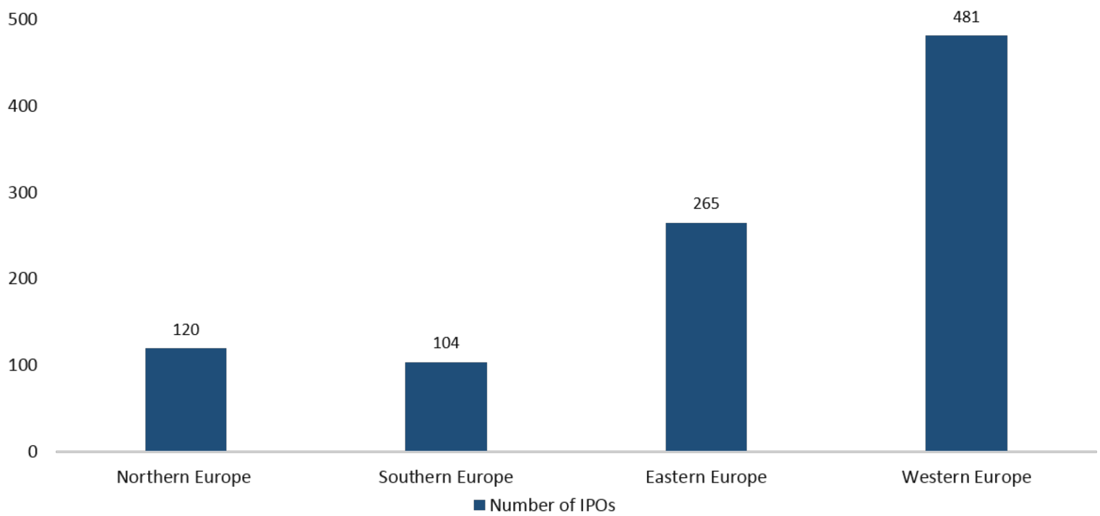


FIGURE 4.6: Number of IPOs by region.

4.3.4 Overview

In table 4.2 the data presents a summary of the characteristics of the individual firms from the data sample. From the overview it is seen that the standard deviations of almost all the parameters are relatively high, which is perceived as a result of a high degree of variation between the companies. Some businesses in the sample are small newly established enterprises with no financial history that went public on smaller and less regulated stock exchange, while others are major conglomerate with several hundred years of successful economic history, extensive financial ratios which went public on broad international stock exchanges. This results in outliers but are defined as natural outliers and contains no bias in the estimation, however, the significant difference in the values will have an influence on the estimated effects and provides less evidence of the detected effects. The outliers have an adverse effect on the linear regression model performance. To avoid the negative effect on the model, the natural logarithms of some of the variables are taken before they are implemented in the regressions. Underpricing presented in table 4.2 is estimated as the return for the stock on the first day of trading minus the market return on the same day. As it is seen in the table the results in this paper confirms existing theory and argues that underpricing is unavoidable when the data sample covers several financial periods. The table also shows that underpricing has a large standard deviation of 124,49% and the minimum value of -644,76% and the maximum value of 891,72% indicates that underpricing has significant outliers. Underpricing as a parameter of long-run performance will be investigated in later sections of this paper. The hot market parameter also confirms that more firms go public in this period since this data sample defines three years as cold market periods and four years as hot market period as discussed in

TABLE 4.2: Statistic characteristics of the data sample.

Variable	Mean	Median	Standard deviation	Minimum	Maximum
Age	15,44	7,70	29,47	0,07	493,99
Size	9,44	9,36	2,51	2,01	17,05
EBIT	6.374,78	44,07	55.874,14	-172.700,00	1.245.000,00
Total assets	318.977,73	11.618,99	1.726.674,45	7,50	25.493.926,21
ROA	0,18	0,00	1,34	-8,44	9,97
Offer price	11,62	3,16	46,92	0,00	945,85
Underpricing	9,93%	2,57%	124,49%	-644,76%	891,72%
VC/PE backed	0,12	0,00	0,32	0,00	1,00
Hot market	0,70	1,00	0,46	0,00	1,00
Delisted	0,34	0,00	0,47	0,00	1,00
Total no. Of. Obs.	970				

section 4.3.1. With an even distribution, the mean should be 0,57, yet the mean of the parameter is 0,70 which means that the hot market is over-represented by 12% compared to the natural distribution between hot and cold market based on the period of the data sample. Delisted is covering if a firm excluded from trading in the whole period from 2007 to 2013 and not if it is delisted before one, two or three years of trading as mentioned in section 4.3.1. It is interesting to see how even in such a short period as seven years 34% of the total population resigned from public trading. It is shown in the table that total assets have the largest outliers and only 12% of the 970 firms in the data sample were backed by either a venture capitalist or a private equity fund.

4.4 Cross-sectional statistics

To obtain a deeper analysis of the parameters, assigned to each company in the data sample are several cross-sectional statistics provided in the next section. The cross-sectional statistics are conducted through a number of regressions across time, industries and regions.

4.4.1 Across time

As it is seen in table 4.3 there is a considerable variation in the different parameter across the period 2007 to 2013. As stated earlier 2007, 2010, 2011 and 2013 are hot markets and 2008, 2009 and 2012 cold markets. Based on the two categories the cross-sectional data from the table will be analysed to see if there are some trends between the two groups as well as within each cluster.

First, by looking at the mean of the age, there is a clear trend that the highest average of age is towards the hot markets. 2007 has the most senior age mean of 18, 89, and 2012 has the lowest age mean of 11,89. When analysing the remaining parameters there is no clear trend between the hot markets and cold markets. Even by changing 2013 to a cold market there is still no pattern between the two groups. This result could indicate that the period when a firm goes public does not have any significant impact on the long-run performance since there are no particular characteristics of the parameters. The lacking of trend could also indicate that the global financial crisis and the later Euro zone crisis had a larger impact on the financial markets and might even have changed some of the fundamental assumptions in the financial world. Of course no conclusions can be made at this point, but the cross-sectional statistics shall be included in the findings of the results from the regressions. As discussed earlier it was a bit unclear if 2013 should be classified as a hot market period or a cold market period. Based on the ambiguous statistics from the cross-sectional analysis 2013 is assumed to be a hot market period.

Despite the blurry picture from the cross-sectional statistics it is still possible to touch upon some trends between the parameters. First, it seems like there is a link between size and age. All the hot years of the highest age also have the highest level of size the only exception is 2012 which has a relatively high level of size, but 2012 also have the highest standard deviation, which could indicate that there is a bias in size for that year. The offer price was fluctuating from an all time high of 18,81 in 2008 to 5,77 in 2012. The reason for an all time high in 2008 could be related to delay of the outbreak of the global financial crisis starting in the US and then later in the European market. If 2008 is assumed to be a mismatch of timing, all the years have defined as hot markets the highest level of offer prices. Lastly, there also seems to be some correlation between the VC/PE backed parameter and underpricing. The two largest outliers in underpricing are 2008 with a value of -1,36% and 2009 with 29,89%. These two years are also the one with the lowest number of VC/PE backed IPOs. Therefore, it could be concluded that an IPO, which is backed have a more moderate level of underpricing. Furthermore, the hot market years have the highest number of VC/PE backed IPOs besides 2012, which has the second highest level in the total data sample.

As seen in the last section there is still some correlation between the period, and some of the parameters picked in this paper as investigating long-run performance of IPOs. The relationship is not clear and there is a high-level of noise between both periods as well as parameters.

TABLE 4.3: Statistic characteristics of the data sample across time.

Variable	Periods 2007	2008	2009	2010	2011	2012	2013
Age	18,89	12,33	13,18	14,31	15,75	11,89	16,23
	32,84	*21,77*	*14,92*	*23,32*	*24,53*	*17,13*	*49,50*
Size	9,67	9,21	9,04	9,48	9,24	9,44	9,62
	2,45	*2,51*	*2,34*	*2,41*	*2,68*	*2,81*	*2,37*
EBIT	5.565,18	5.206,12	1.151,46	1.084,46	1.917,69	14.978,54	18.256,80
	29.194,18	*33.117,51*	*6.881,30*	*7.146,75*	*12.073,94*	*127.818,70*	*97.297,15*
Total assets	270.803,59	395.096,49	423.823,20	296.482,71	430.966,87	269.255,02	217.624,91
	1.222.662,45	*2.224.585,18*	*2.986.892,30*	*1.414.724,54*	*2.378.195,88*	*959.417,63*	*816.133,18*
ROA	0,26	0,16	0,09	-0,05	0,12	0,02	0,52
	1,47	*1,31*	*0,84*	*0,68*	*1,04*	*1,30*	*2,07*
Offer price	15,33	18,81	5,78	6,60	11,98	5,77	8,88
	46,68	*62,27*	*12,91*	*19,37*	*79,82*	*13,88*	*19,22*
Underpricing	8,58%	-1,36%	29,89%	4,85%	19,18%	16,27%	4,76%
	1,25	*1,67*	*1,44*	*0,95*	*1,32*	*0,96*	*0,99*
VC/PE backed	0,13	0,05	0,05	0,12	0,09	0,14	0,21
	0,33	*0,22*	*0,21*	*0,32*	*0,28*	*0,35*	*0,41*
Delisted	0,40	0,42	0,41	0,33	0,34	0,26	0,12
	0,49	*0,49*	*0,50*	*0,47*	*0,47*	*0,44*	*0,32*
Total no. Of. Obs.	273	132	66	152	140	95	112

4.4.2 Across industries

As mentioned in section 4.3.2 there is only 15 different industries represented in the data sample. In the table, 4.4 the variation of parameters is presented across the 15 industries. Manufacturing (C) has the largest activity level of IPOs with 292 number of observations. Manufacturing has the second highest mean of age, which confirms the expectation that manufacturing is a mature industry and will have a relatively high mean of age. The industry has a low profitability measured by ROA. The low ROA was expected to give a high underpricing, but as seen in the table the underpricing is the lowest positive value across all industries. There is no clear indication of this unexpected result when the different parameters are compared with other sectors. One indication could be that Manufacturing has a relatively high number of firms, which are backed by VC or PE since out of the total number of companies across all industries, which are VC or PE backed is the overall share 45,13%. Combined with the maturity of the sector, which indicates that industry is more stable and conservative and therefore do not have any significant outliers in the different parameters.

The cross-sectional statistics of industries are not provided with any clear trends from the individual sectors. Despite that there is no clear trend in the data set across industries some characteristics can still be contracted from the table. The three industries with the lowest number of observation are Accommodation and Food service activities (I), Human health and Social work activities (Q) and Arts, Entertainment and Recreation (R). These three sectors have the lowest underpricing with a significant gap to the rest of the areas. When mapping the underpricing to Age is (I) and (Q) the two sectors with the lowest average and (R) have a relatively high mean, but with a large standard error. The large standard error is due to a significant outlier and is the outlier excluded is (R) the industry with the lowest Age mean. Accommodation and food service activities have the most moderate size across all sectors and Arts, entertainment and recreation are at the bottom end as well indicating size may also is correlated to some of the parameters mentioned above.

TABLE 4.4: Statistic characteristics of the data sample across industries.

	Industries														
Variable	A	B	C	D	F	G	H	I	J	K	L	M	N	Q	R
Age	9,04	8,99	21,66	10,16	14,87	15,11	40,19	5,24	10,61	12,63	11,50	11,49	11,75	8,27	14,41
	6,52	*18,23*	*34,37*	*12,40*	*14,74*	*22,79*	*107,12*	*4,77*	*15,95*	*24,99*	*19,01*	*23,25*	*10,90*	*4,33*	*29,96*
Size	9,29	9,82	9,50	9,60	9,76	9,16	11,16	8,61	9,38	9,26	9,01	9,20	9,07	9,49	9,00
	1,64	*2,95*	*2,39*	*2,62*	*2,46*	*2,87*	*2,06*	*2,75*	*2,37*	*2,89*	*2,00*	*2,06*	*2,39*	*2,79*	*2,73*
EBIT	5.266,49	5.834,50	5.074,33	5.182,24	2.144,39	2.030,15	28.636,32	299,44	9.064,90	16.669,03	2.244,24	1.453,05	617,26	2.202,34	-338,35
	9.516,76	*46.266,72*	*38.311,08*	*11.595,86*	*30.675,06*	*15.294,41*	*82.026,23*	*1.404,57*	*64.415,50*	*130.432,09*	*8.523,14*	*13.803,35*	*2.788,70*	*22.454,14*	*1.456,70*
Total assets	27.831,81	673.816,83	258.376,52	297.150,55	145.124,43	572.612,39	316.368,38	142.155,63	389.937,66	421.537,48	64.169,18	104.511,90	116.883,08	60.230,16	342.803,69
	35.053,86	*2.343.903,13*	*1.635.125,60*	*864.496,91*	*375.521,28*	*2.698.233,53*	*523.069,77*	*381.932,69*	*2.292.730,50*	*1.610.018,34*	*224.000,75*	*349.721,31*	*424.651,01*	*83.143,93*	*1.062.208,86*
ROA	0,66	-0,09	0,14	0,27	0,18	0,16	0,22	-0,02	0,20	0,25	0,79	-0,04	-0,08	1,13	0,14
	1,05	*1,60*	*1,58*	*1,18*	*1,47*	*0,92*	*0,85*	*0,22*	*1,19*	*1,00*	*2,37*	*0,72*	*0,89*	*2,30*	*0,39*
Offer price	5,48	20,44	9,99	47,83	8,86	7,45	13,50	1,97	12,59	8,27	12,84	6,31	4,20	7,86	9,45
	6,50	*77,66*	*24,62*	*171,01*	*13,35*	*17,33*	*14,48*	*2,36*	*42,82*	*34,66*	*29,11*	*8,25*	*4,37*	*9,22*	*12,87*
Underpricing	9,04%	14,74%	3,47%	27,49%	-3,88%	27,94%	12,48%	-33,10%	9,52%	23,14%	3,58%	13,26%	23,52%	-19,89%	-46,78%
	0,81	*2,20*	*1,09*	*1,44*	*1,19*	*1,52*	*0,12*	*0,69*	*1,12*	*1,31*	*1,52*	*0,95*	*0,94*	*0,71*	*1,02*
VC/PE backed	0,00	0,06	0,17	0,03	0,01	0,13	0,14	0,11	0,12	0,04	0,13	0,13	0,04	0,44	0,18
	0,00	*0,24*	*0,38*	*0,16*	*0,12*	*0,34*	*0,36*	*0,33*	*0,33*	*0,21*	*0,34*	*0,34*	*0,20*	*0,53*	*0,40*
Delisted	0,24	0,27	0,33	0,26	0,27	0,46	0,19	0,67	0,38	0,34	0,22	0,36	0,31	0,33	0,27
	0,44	*0,45*	*0,47*	*0,44*	*0,45*	*0,50*	*0,40*	*0,50*	*0,49*	*0,48*	*0,42*	*0,49*	*0,47*	*0,50*	*0,47*
Hot market	0,71	0,67	0,75	0,49	0,72	0,68	0,67	0,67	0,64	0,66	0,74	0,76	0,77	0,89	0,73
	0,47	*0,47*	*0,43*	*0,51*	*0,45*	*0,47*	*0,48*	*0,50*	*0,48*	*0,48*	*0,45*	*0,43*	*0,43*	*0,33*	*0,47*
Total no. Of. Obs.	17	52	292	39	71	91	21	9	162	92	23	55	26	9	11

4.4.3 Across regions

By analysing the characteristics of the parameters across regions, we can see several different results confirming existing theory. As seen in table 4.5 Southern Europe has the oldest average followed by Eastern Europe of the IPOs in the constructed data sample across the four regions. While the Northern European region has the lowest age followed by Western Europe with the second lowest mean of age. The low level of age in the Northern Europe and Western Europe region could be due to several different theories as mentioned earlier in this paper. In section 2.1.2, the legal systems and requirements are discussed. Based on this the difference in age could come from a better legal system and country specific requirements. The largest grey markets are represented in the Western European region and the grey markets stock exchanges are often more attractive for younger firms. Another interesting correlation between age and assets are found in

TABLE 4.5: Statistic characteristics of the data sample across regions.

	Regions			
Variable	Northern Europe	Southern Europe	Eastern Europe	Western Europe
Age	12,05	25,70	16,93	13,25
	23,59	*52,92*	*26,79*	*24,25*
Size	9,13	9,76	9,09	9,65
	2,70	*2,31*	*2,55*	*2,45*
EBIT	188,60	13.173,28	5.329,16	7.024,23
	3.448,84	*71.741,35*	*32.050,04*	*67.892,81*
Total assets	353.204,06	210.215,18	209.570,09	394.231,72
	1.704.948,87	*826.121,12*	*1.014.072,74*	*2.137.961,11*
ROA	-0,20	0,57	0,32	0,10
	0,90	*1,40*	*1,22*	*1,45*
Offer price	6,18	5,87	11,55	14,26
	23,67	*6,02*	*69,17*	*40,56*
Underpricing	33,83%	18,56%	7,34%	3,54%
	1,81	*0,80*	*1,15*	*1,20*
VC/PE backed	0,08	0,11	0,03	0,18
	0,26	*0,31*	*0,16*	*0,38*
Delisted	0,54	0,25	0,25	0,35
	0,50	*0,44*	*0,43*	*0,48*
Hot market	0,64	0,73	0,65	0,73
	0,48	*0,45*	*0,48*	*0,44*
Total no. Of. Obs.	120	104	265	481

the cross-sectional statistic. Despite that Northern Europe and Western Europe have the youngest the two regions have the highest mean of assets across the four regions. First of all, this could indicate that existing theory arguing a correlation between age and asset is not necessarily the case. Another possible reason for this result could be that in a well-developed financial system firms have earlier access to financing and growth. Further analysis shows that both Northern Europe and Western Europe have an over-representation of industries characterised as assets intensive. Furthermore, it can be seen that there is a correlation between ROA and age. The statistics are indicating that younger firms have a lower profitability than older firms. This can be interpreted as a signal of a healthy legal system since investors are willing to take a larger risk and investor in business with a low level of profitability. Which is also in line with the findings mentioned above.

Northern Europe is the region with the highest number of delisting[6] followed by Western Europe. Based on the low profitability as well as age the scores of delisting are expected and are in line with theory. However, it is still too early to make any conclusion about the long-run performance of IPOs in the two regions.

As seen from the three different cross-sectional statistics, time, industries, and regions vary the potential of finding any characteristics or trends. The two first cross-sectional statistics indicate that time and industries had a great extended mix of levels of the different parameters which could mean that the two parameters are lacking any prediction power on long-run performance. On the other hand where the trends across regions are more apparent, and it was possible to relate the core statistics from the constructed data set to existing theory.

[6]The number of delisting is based on the full period from 2007 to 2013 and not only the period of up to three years of trading.

Chapter 5

Empirical results

The next section will outline the results from the regressions formulated in section 3.3.3. A broad discussion of the results will be conducted as well as an analysis of a test of the hypotheses from section 2.4. The empirical results will be used to answer the question for research formulated in section 1.2.1. A discussion of the findings in this paper concerning the existing literature within the field will be discussed as well.

The presentation of the empirical results will include long-run performance, estimations based on both the CAR model as well as the BHAR model. The results will take arise from the simplest model where only a single parameter is tested separately and then build up to the largest model including all parameters and dummy variables. Since the two dummy variables, industry and region consist of a group of variables a particular baseline will be assigned to the regressions. The baseline for the industry dummy variable is Agriculture, Forestry and Fishing (A) and for the region dummy variable the baseline is the Eastern European region (EE). Finally, a robustness test of the results from the regression will be made. The robustness test will be evaluated by calculating Type I and Type II errors for each model.

5.1 Firm specific information

The following section will outline the correlation between firm specific parameters; Age, $\log(Size)$, ROA, and $Underpricing$ and the long-run performance as discussed in section 3.2.1. Regression 3.13 and 3.14 aims to test 2.4, and establish whether Age have any prediction power of long-run IPO performance. Regression 3.13 contains only one predictor Age which will be tested on over a time horizon of one, two and three years after the IPO date. Moreover, four sets of dummy variables will be added to the model which is represented in regression 3.14. Namely; $Industry$, $Region$, VC/PE, and Hot, are included as controls for fixed effects.

The empirical results from regression 3.13 is presented in table 5.1. As it is seen in the table, Age does not have any significant correlation with long-run IPO performance based on the CAR model. When the regression is testing Age based on the BHAR model, there is a significant correlation between Age and IPO performance in a period of one and three years of trading. The estimated coefficient of Age is positive, but have a relatively small value. The simple test of Age indicates that the prediction power of Age is highly depending on which model of estimating performance is used. The positive coefficient estimate of Age can be interpreted, as older firms has a high probability of outperforming the market but only marginal due to the relatively small value. This is in line with existing theory, but since the CAR model shows no significant correlation it is not possible to draw any clear conclusion. Therefore, due to the insignificant $p - values$ for the majority of the regression it is not possible to reject the null hypothesis, there is no positive correlation between IPO performance and Age.

TABLE 5.1: Empirical results from simple test of Age

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	-0,217574	-0,206493	-0,112944	-0,588176	-0,819217	-0,921402
	*0,00298 ** *	*0,00462 ** *	*0,123*	*9,16e-15 **** *	*<2e-16 **** *	*<2e-16 **** *
Age	0,002575	0,001318	0,002783	0,004162	0,003412	0,00524
	0,25417	*0,54845*	*0,225*	*0,0733 .*	*0,135*	*0,0278 **
AIC	1339,7	1340	1346	1281,2	1214,2	1187,3

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

Despite that the simple regression indicates that Age is insignificant based on the CAR model, regression 3.14 will still be used based on the CAR model as well as the BHAR model. The table 5.2 displays the test results of Age prediction power of long-run IPO performance when the set of four dummy variables are included. The results presented in table 5.2 are the final results after the backward elimination has been conducted to the regression[7]. The results from the regression and the backward elimination show that Age has a significant correlation with long-run IPO performance estimated both with CAR and BHAR. Only CAR in two year period of trading shows any correlation between Age and IPO performance. The variable Age appeared to have a positive significant impact on long-run IPO abnormal performance in one, two and three years of trading. The findings of the influence of Age supports empirical findings from (Ritter, 1991), and supports the argument that more mature firms have a higher probability of outperforming the market. The backward elimination finds that the best model is including the dummy variable; $Region$, VC/PE and Hot, where VC/PE is only presented in a time horizon of three years and Hot is only significant based on BHAR in a two-year' time horizon. From the table, it is seen that the region Western Europe has a high positive significant coefficient except

[7]The full empirical result from the regression before backwards elimination is presented in appendix B.1, B.6.

in performance measured with CAR in a three year period. It means that if the baseline region EE is changed to WE the probability of abnormal returns increases. In general, the Southern Europe region has a negatively estimated coefficient. The Northern Europe region is positive except from in BHAR one year where the estimated coefficient is negative. Both, the SE region and the NE region lacks a consistent significance level across the periods and performance method. The Intercept is corresponding to the baseline region, which is the Eastern Europe region EE. The Intercept is negative but significant across all the regression meaning that the log odds for CAR or BHAR to be negative in the EE region are higher than CAR or BHAR to be positive.

TABLE 5.2: Empirical results after backwards selection test of Age and dummy variables

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	-0,48893 *0,000212 ****		-0,298726 *0,0217 **	-0,968745 *9,06e-12 ****	-1,121235 *3,74e-10 ****	-1,412136 *< 2e-16 ****
Age	0,003948 *0,088569 .*		0,003963 *0,0917 .*	0,005353 *0,0259 **	0,005122 *0,03228 **	0,00628 *0,010411 **
Region NE	0,274281 *0,21815*		0,558148 *0,0126 **	-0,026624 *0,9132*	0,555393 *0,02545 **	0,61234 *0,012831 **
Region SE	-0,427058 *0,086008 .*		-0,621272 *0,0120 **	-0,125168 *0,6295*	0,187535 *0,49511*	-0,307844 *0,299122*
Region WE	0,516149 *0,000922 ****		0,225291 *0,1524*	0,727013 *9,03e-06 ****	0,909724 *3,64e-07 ****	0,606834 *0,000830 ****
VC/PE backed			0,434051 *0,0395 **			0,809025 *0,000128 ****
Hot market					-0,427009 *0,00492 ***	
AIC	1322,8		1327,9	1255,3	1185,7	1153,2
AUC	0,5855		0,5911	0,6122	0,6238	0,6374

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

The dummy variable Hot has a negative value which is indicating that firms going public in hot market period have a high probability of underperforming based on a two year trading period measured with BHAR. The fact that the dummy variable is only present once, is in line with the cross-sectional statistic of time. Here it was difficult to see any clear trends difference between the hot market and the cold market period.

The dummy variable $VC/PE\ backed$ has a positive coefficient, which is confirming the current theory that firms supported by a professional investor have a higher probability of outperforming the market. But as it is seen in table 5.2 the variable is only significant when the investment horizon is three years.

The significant $p - values$ for the coefficient Age means that the null hypothesis, there is no positive relation between Age and IPO performance, is rejected when the dummy variable $Region$ is included.

The suggested model for Age correlation with performance have the lowest AIC values in the case of returns estimated with BHAR. This indicates that BHAR is a better approach for evaluating performance. It is also clear to see that the AUC is significantly higher for BHAR than CAR. But in both cases, the AUC has a relatively small value indicating that the model does not have the highest quality. A more detailed analysis of the quality of the model will be conducted in section 5.3.

The second firm specific parameter which is tested for correlation to IPO performance is $\log(Size)$. In table 5.3 the empirical results from the simple regression 3.15 is presented. As it is seen from the table, the simple model indicates that $\log(Size)$ has no significant impact on the long-run IPO performance. The insignificant $p - values$ across all the different regression for the coefficient $\log(Size)$ means that the null hypothesis, there is no positive relation between $\log(Size)$ and IPO performance, is not rejected.

TABLE 5.3: Empirical results from simple test of $\log(Size)$

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	0,10588	0,10238	0,32949	-0,530749	-0,752212	-0,913787
	0,674	*0,684*	*0,19*	*0,0404 **	*0,00514 ***	*0,000817 ****
Size	-0,0301	-0,0306	-0,0423	0,0008268	-0,001372	0,008069
	0,244	*0,236*	*0,1*	*0,9751*	*0,96024*	*0,77229*
AIC	1339,7	1338,9	1344,8	1284,6	1216,5	1192,4

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

As in the case with analysing the Age variable a set of dummy variables is added to the simple model. The regression after the added dummy variables is mathematically defined in equation 3.16. The results from the regression after the backward elimination is applied are presented in table 5.4[8]. As presented $\log(Size)$ only has a significant correlation in a three year period of trading based on the CAR estimation of performance. The estimated coefficient of $\log(Size)$ has a negative value and therefore, the probability of large firm underperforming is higher. This is not in line with existing theory (Loughran and Ritter, 1995) and the empirical results from firm age.

[8]The full empirical result from the regression before backwards elimination is shown in appendix B.2, B.7.

It was expected that firm age and size of the business was correlated. But as mentioned in the cross-sectional statistic in section 4.4, there was no clear relation between the age of a company and the size. This could indicate that within the European Union the financial system is well developed and therefore, younger firms can quickly obtain a large asset base. But this could also be correlated towards that more asset light companies, such as technology companies go public, and with these two opposite directed trends the prediction power based on the size of the firm will decline. The region dummy variable indicates that the regions; EE, NE, and WE have positives log odds while the SE region has negative log odds. The regional dummy variables EE and WE have a high $p - value$ which makes them insignificant. The $VC/PE\ backed$ dummy variable is again indicates that in a time horizon of three years firms backed by professional investors have a high probability of outperforming the market.

TABLE 5.4: Empirical results after backwards selection test of $\log(Size)$ and dummy variables

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept			0,18116 *0,4988*			
Size			-0,0456 *0,0825 .*			
Region NE			0,53811 *0,0160 **			
Region SE			-0,5583 *0,0234 **			
Region WE			0,22951 *0,1454*			
VC/PE backed			0,47632 *0,0237 **			
Hot market						
AIC			1327,8			
AUC			0,6017			

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

With only a significant $p - value$ of $\log(Size)$ when estimated based on three year period of trading based on the CAR estimate is the null hypothesis, there is no positive relation between $\log(Size)$ and IPO performance is not rejected. Therefore, it is argued that $\log(Size)$ does not have any predictive power of long-run IPO performance.

The third hypothesis 2.4 states that there is a positive relation between IPO performance and firm profitability. This assumption is tested by regression 3.17, where ROA is the indication of firm profitability. Table 5.5 displays the empirical results from the simple test of only ROA's prediction power of long-run IPO abnormal returns. The results from the regression indicate that ROA is significant positive correlated with IPO performance based on the BHAR estimation of return with a trading horizon of two and three years. Since the two significant coefficients of ROA are both positive it indicates that firms are more profitable and performing better than less profitable firms.

The empirical results for CAR one, two, and three years and BHAR one year show no significant $p-values$ the null hypothesis is not rejected. Therefore, it is argued that ROA does not have any predictive power of long-run IPO performance. While the empirical results for BHAR two and three years have significant $p-values$ and therefore, the null hypothesis is rejected.

TABLE 5.5: Empirical results from simple test of ROA

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	-0,18149 *0,00527 ***	-0,19356 *0,00295 ***	-0,0763 *0,239*	-0,53156 *2,43e-15 ****	-0,78738 *<2e-16 ****	-0,86081 *<2e-16 ****
ROA	0,02092 *0,66381*	0,04186 *0,38654*	0,0351 *0,467*	0,04653 *0,346*	0,10775 *0,0354 **	0,11095 *0,0317 **
AIC	1340,9	1339,6	1347	1283,7	1212	1187,8

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

To make a more narrow and precise test of ROA the same set of dummy variables as in the previous two test are added to the regression as specified in equation 3.18. The results for the larger model gives the same results as the simple model after the backward elimination[9]. It is interpreted that it is only in a time horizon of two and three years based on the BHAR calculation that ROA is significant. But, after the backward elimination the dummy variable region is added to the model in both periods, and the Hot market dummy variable is added in a two years horizon, while in the three horizons the $VC/PE\ backed$ variable is added.

From the results we do not see a clear picture of ROA prediction power of long-run IPO performance. This was also the indication from the cross-sectional statistical analysis, where the conducted study did not find any clear trend between ROA and different parameters. This could be due to the fact, as stated in section 2.1.2, that more unprofitable firms are entering the market. While the performance outcome for less profitable firms entering the market is blurry it is therefore, not possible to find any clear trend. But as most of the existing theory concludes more

[9] The full empirical result from the regression before backwards elimination is shown in appendix B.3, B.8.

profitable firms perform better. The results are indicating that firms going public in the NE and WE region have high log odds of positive outperformance. If the company is backed by a venture capitalist or private equity investor the probability of abnormal returns is higher. If the firm, however, goes public in a period of a hot market the chance of underperforming is higher in a trading period of two and three years.

TABLE 5.6: Empirical results after backwards selection test of ROA and dummy variables

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept					-1,08786 *6,39e-10 ****	-1,35365 *< 2e-16 ****
ROA					0,14334 *0,00690 ***	0,14484 *0,007466 ***
Region NE					0,60824 *0,01506 **	0,65594 *0,007954 ***
Region SE					0,20098 *0,46316*	-0,28421 *0,334109*
Region WE					0,92346 *2,69e-07 ****	0,60968 *0,000791 ****
VC/PE backed						0,82876 *8,25e-05 ****
Hot market					-0,42209 *0,00543 ***	
AIC					1183	1152,8
AUC					0,6296	0,6405

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

But as concluded by the simple model BHAR two and three years have significant $p - values$ and therefore, the null hypothesis is rejected. The rest of the regressions fails to reject the null hypothesis.

The last parameter tested is underpricing, which is defined as the abnormal return reached on the first-day trading. Underpricing is a short-run theorem as mentioned in section 2.2.1 but several studies have indicated that underpricing have a significant impact on long-run IPO performance. As stated in hypothesis 2.4, it is expected that underpricing would have a negative relation to IPO performance. To test the hypothesis the regressions 3.19 and 3.20 are formulated. The results from the simple regression, 3.19 testing only the parameter $Underpricing$ is presented in the table 5.7.

The empirical results from the simple regressions are explicitly confirms existing theory (Ritter, 1991) since all the $p - values$ shows a significant level expect from CAR one year. The results also indicates the longer the trading period the higher the negative estimated coefficient, meaning that the log odds for underperformance are increasing. The evident results from the simple regression mean that the null hypothesis is rejected. To get a more detailed picture of $Underpricings$ prediction the second regression will still be completed.

TABLE 5.7: Empirical results from simple test of $Underpricing$

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	-0,17019	-0,1745	-0,05528	-0,51528	-0,75639	-0,83157
	*0,00854 ***	*0,00711 ***	*0,39299*	*1,03e-14 ****	*< 2e-16 ****	*< 2e-16 ****
Underpricing	-0,08178	-0,13185	-0,16183	-0,09604	-0,22279	-0,28249
	0,12251	*0,01564 **	*0,00347 ***	*0,0821 .*	*0,000367 ****	*2,07e-05 ****
AIC	1338,6	1334,2	1338,4	1281,5	1202,4	1171,5

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

The empirical results from regression 3.20 after backward elimination is in line with the results from the simple regression 3.19. As seen in table 5.8[10] CAR one year is insignificant but also BHAR one year is insignificant in the larger model. In the simple model BHAR one year had the highest $p-value$ therefore, the rejection of BHAR one year, in the larger model, is not a controversial result. The insignificant correlation between $Underpricing$ and IPO performance in a time horizon of one year could be because $Underpricing$ is a short-run theorem. This means $Underpricing$ makes the stock performance volatile in the relatively short period of one year. Thus, the high level of volatility covers for the prediction power seen, $Underpricing$ has a longer trading period.

The dummy variables included in the model after the backward elimination are in line with results from the previous three regressions. This indicates that hot market periods have a negative log odds outcome, financially backed IPOs have a positive log odds result. Both dummy variables have a significant $p - value$ in all the outcome where it is included. The estimated coefficient for the region SE has a negative value but only a significant $p - value$ in CAR three years. Also in CAR three years the region WE does not not have a significant $p - value$. This could mean that the two regions are offsetting each other since the regression testing $\log(Size)$ have the same result.

In the empirical results showing significant $p - values$ for $Underpricing$, there is no negative relation between $Underpricing$ and IPO performance, meaning the null hypothesis is rejected. This is based on a trading period of two and three years. From the AIC and AUC scores it is concluded that BHAR used for estimating IPO abnormal returns is the best model since the AIC

[10]The full empirical result from the regression before backwards elimination is shown in appendix B.4, B.9.

scores are significantly lower than for CAR based estimation. The AUC score for BHAR three years is the highest value across all the regressions and therefore it is concluded that $Underpricing$ have the highest prediction power of IPO performance based on three year period.

TABLE 5.8: Empirical results after backwards selection test of $Underpricing$ and dummy variables

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept		-0,29944 *0,05446 .*	-0,2223 *0,07382 .*		-1,04253 *2,93e-09 ****	-1,30717 *< 2e-16 ****
Underpricing		-0,13229 *0,01653 **	-0,1665 *0,00283 ***		-0,22397 *0,000491 ****	-0,28541 *2,41e-05 ****
Region NE		0,46863 *0,03678 **	0,58708 *0,00934 ***		0,57837 *0,020910 **	0,63106 *0,01125 **
Region SE		-0,27168 *0,27*	-0,56753 *0,02108 **		0,26996 *0,324118*	-0,19905 *0,49809*
Region WE		0,42736 *0,00752 ***	0,20337 *0,19663*		0,89008 *6,72e-07 ****	0,57466 *0,00159 ***
VC/PE backed		0,462 *0,02730 **	0,45455 *0,03099 **			0,85396 *5,33e-05 ****
Hot market		-0,24937 *0,08192 .*			-0,4102 *0,007111 ***	
AIC		1319,2	1321,4		1176,8	1139,5
AUC		0,6042	0,612		0,6273	0,6545

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

5.2 Combined test results

The four regressions performed above test the parameters Age, $\log(Size)$, ROA, and $Underpricing$ separately but as mentioned previously. Age and $\log(Size)$ could have a significant correlation. To test if there are any correlation between the parameters and to test if it is possible to construct a model with a higher prediction power, all the parameters are included at once and not only separately.

The first regression 3.21 tests only the four parameters without any dummy variables. The output from the regression is presented in table 5.9. When comparing the results when all the parameters are included in the individual models there are only some few changes in the results. Age is now only significant in BHAR three years while in the simple model BHAR one year also indicates a significant $p-value$. $\log(Size)$ has a significant negative estimated coefficient in CAR three years while in the simple there were no significant estimations. Both ROA and $Underpricing$ follows the same results as in the simple model. When comparing the AIC score there is no clear indication that the larger model is significantly better in prediction IPO performance compared to the simple models. Only BHAR three years have a lower AIC across all the regressions.

TABLE 5.9: Empirical results from simple test of all parameters

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	0,135724	0,161474	0,404282	-0,51932	-0,717316	-0,879966
	0,595	*0,5276*	*0,11441*	*0,0482 **	*0,009288 ***	*0,00176 ***
Age	0,00258	0,001109	0,002661	0,003822	0,002646	0,004344
	0,259	*0,6174*	*0,25133*	*0,1016*	*0,248589*	*0,06855 .*
Size	-0,03684	-0,038043	-0,053409	-0,0067	-0,010877	-0,004723
	0,158	*0,1462*	*0,04159 **	*0,803*	*0,699492*	*0,86926*
ROA	0,01267	0,036098	0,02512	0,038049	0,102229	0,104558
	0,793	*0,4575*	*0,606*	*0,4435*	*0,048130 **	*0,04650 **
Underpricing	-0,085931	-0,138298	-0,169832	-0,09227	-0,223849	-0,282012
	0,108	*0,0119 **	*0,00239 ***	*0,0981 .*	*0,000396 ****	*2,64e-05 ****
AIC	1341,4	1337,2	1338,7	1283,8	1202,4	1169,2

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

Again the dummy variables are included and a backward elimination is performed. When conducting the backward elimination it can now also be seen which parameter beyond the dummy variables should be added to reach the highest prediction power[11] . Table 5.10 provides the empirical results from the regression. Overall, by including every parameter in the model, the AIC as well as the AUC levels increase and therefore, the largest model is still a better model to predict IPO performance. In general, the parameters included in the new model are the same parameters, which had a significant coefficient in the simple regressions as well. The largest difference from the largest model to the simple models is that $\log(Size)$ is included when returns are based on CAR. It is also seen that both Age and $\log(Size)$ is present in CAR one year and CAR three years. But Age has positive coefficient while $\log(Size)$ has a negative. This means that there is a correlation between the two parameters but in different directions. This is not in line with existing theory (Ritter, 1991).

In the largest model the industry is excluded from the model. This is in line with the cross-sectional statistic, indicating that there was no clear trend across industries. This result is therefore, pointing out that the assumption that some industries have a better performance than others as argued by (Ritter, 1991) is questionable. An explanation for this finding could be those industries,

[11]The full empirical result from the regression before backwards elimination is shown in appendix B.5, B.10.

which are characterised by, e.g. older and larger firms. This is probably still the case if the industries are analysed post-IPO. Nevertheless, companies going public today in different industries do not have any clear trends since the markets are more dynamic and a more diversified group of companies than those 30 years ago are conducting IPOs.

TABLE 5.10: Empirical results for all parameters and dummy variables after backward selection

	CAR			BHAR		
	Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Intercept	-0,087857 *0,75*	0,09662 *0,73*	0,240058 *0,37793*	-0,968745 *9,06e-12 ****	-1,145561 *2,39e-10 ****	-1,454076 *< 2e-16 ****
Age	0,004069 *0,081946 .*		0,003961 *0,09477 .*	0,005353 *0,0259 **	0,004245 *0,074602 .*	0,005313 *0,029299 **
Size	-0,043913 *0,098054 .*	-0,04427 *0,09592 .*	-0,058298 *0,02881 **			
ROA					0,13658 *0,010143 **	0,138116 *0,011202 **
Underpricing		-0,14153 *0,01086 **	-0,173887 *0,00203 ***		-0,219673 *0,000699 ****	-0,281946 *3,61e-05 ****
Region NE	0,298078 *0,183062*	0,47253 *0,03556 **	0,614211 *0,00684 ***	-0,026624 *0,9132*	0,675304 *0,007716 ***	0,736403 *0,003568 ***
Region SE	-0,390869 *0,117589*	-0,24486 *0,3217*	-0,569063 *0,02215 **	-0,125168 *0,6295*	0,197246 *0,476671*	-0,296503 *0,321702*
Region WE	0,540153 *0,000583 ****	0,44857 *0,00525 ***	0,250039 *0,1161*	0,727013 *9,03e-06 ****	0,944871 *1,84e-07 ****	0,63587 *0,000551 ****
VC/PE backed		0,48315 *0,02139 **	0,463592 *0,02895 **			0,818923 *0,000126 ****
Hot market		-0,24056 *0,09402 .*			-0,448915 *0,003533 ***	
AIC	1322,3	1318,4	1318,1	1255,3	1170,1	1131
AUC	0,5988	0,6084	0,6225	0,6122	0,6443	0,6687

*Note: Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1*

From the empirical results, there are some clear trends, especially within the different regions. The WE region is the variable, which across both CAR and BHAR is most significant and only in BHAR one year is the $p-value$ insignificant. The estimated coefficient for WE is all positive as well as for the NE region except from BHAR one year where the coefficient is negative but insignificant. Both for the EE and for SE region the trend is more blurry and more values are insignificant.

The trend between the individual models and the larger model indicates that the tested parameters have some correlation with IPO performance. The larger model has a significantly better prediction power than the individual models, when comparing both the AIC values and the area under the ROC curve.

5.3 Robustness test

In the next section a more detailed analysis of the quality of the models will be presented. The test performed to evaluate the predictability of the models is examined by the prediction results for $Type\ I$ and $Type\ II\ errors$. The cutoff levels used in the robustness test level is set at 0,5.

$Type\ I\ error$ classifies an IPOs performance as over performance while the IPO is underperforming. A $Type\ II\ error$ classifies an IPOs performance as underperforming while the real classification is overperforming.

TABLE 5.11: Empirical results from robustness test

		CAR			BHAR		
		Year 1	Year 2	Year 3	Year 1	Year 2	Year 3
Age	Type I	186	292	303	343	301	244
	Type II	234	105	124	10	6	50
Size	Type I	192	292	251	361	308	247
	Type II	231	105	156	0	0	48
ROA	Type I	192	285	342	361	299	246
	Type II	231	110	86	0	7	44
Underpricing	Type I	198	289	299	354	295	237
	Type II	222	106	102	8	9	53
Total	Type I	244	269	226	343	280	245
	Type II	177	114	171	10	20	43

The true condition of Y is in the interval $0 \leq y_c \leq 1$. The cutoff level is determined by y_c which can be set at any level between 0 and 1. $\hat{Y}$ denotes the sample prediction value, which is divided into two categories. The first category defines the number of $Type\ I\ errors$, defined as all values of y_c which are less than $\hat{Y}$. The second category defines the number of $Type\ II\ errors$, which is all the values detected as equal or above the cutoff level y_c.

$$Type\ I\ error : \hat{Y} < y_c | Y = 1 \tag{5.1}$$

$$Type\ II\ error : \hat{Y} \geq y_c | Y = 0 \tag{5.2}$$

The true condition of positive and negative are expressed as:

$$Type\ I\ error : \hat{Y} < y_c | Y = 0 \tag{5.3}$$

$$Type\ II\ error : \hat{Y} \geq y_c | Y = 1 \tag{5.4}$$

$\hat{Y}$ is forced to take the value in the interval $0 \leq \hat{Y} \leq 1$ meaning that all negative outcomes are assigned with the value 0 and all values above 1 are equated to 1.

In table 5.11 the errors of the estimated values are presented. As it is seen from the table all the models have a high level of $TypeI$ and $TypeII$ errors. This results is in line with the AUC values where the highest value was 0,6687 from the largest model based on a three years trading period estimated with BHAR. Therefore, the robustness test indicates that the model has a significant prediction power but the quality of the models lay in the lower end of what is acceptable. To increase the prediction power of the model two different approaches can be used. First, a larger data sample would increase the power of the model. But based on the data collection process in this paper a significantly higher number of observations would be almost unachievable. Secondly, all the outliers from the data sample can be excluded. This would increase the prediction power, but would make the model less correlated with the real world. Therefore, the increasing of prediction power of statistically models is a trade-off between prediction power, time used in the data collection process, and correlation with the real world.

Chapter 6

Conclusion

The existing theory has conducted a considerable amount of time to IPO related topics. Especially, IPO long-run performance has received a significant amount of attention and performance indicating parameters.

This paper follows existing theory and an enormous amount of research centred on IPO performance. A larger section of the research concentrates on a few factors, based on existing theory, which predicts performance irregularities. Moreover, a set of control variables; industry, region, and time varying effects on IPO performance have been researched. The paper tests well-known parameters impact on IPO performance, based on a custom-made data set, including all IPOs conducted in the European Union. This limitation makes this research unique since the geographical area of all the 28 countries in the European Union have received a minimum amount of attention in existing papers.

The statistical part of the paper is based on a logistic regression model that controls for time, regions, and industry specific effects. The regression is used to determinate the abnormal return of IPOs.

The main conclusion from this paper is generally confirming existing theory, but with some discrepancies. The first parameter tested for prediction power was Age of a firm. It is concluded that firm age is a highly significant factor for abnormal returns. The older the firms are when conducting an IPO process, the higher is the chance for performing better than the market. It is concluded that the size of a company does not have any statistical significance for IPO performance. Despite the lack of statistical power of firm size, when tested separately, size has a statistical significance, when included together with the rest of the tested parameters.

Profitability defined as ROA is concluded to be an irrelevant and insignificant predictor of IPO performance. Existing theory argues that profitability is a determining factor for more successful IPOs (AlOmoush and Al-Shubiri, 2013), (Skogsvik and Skogsvik, 2010). However, if the rest of the tested parameters are included together with ROA, a statistical significance of ROA occurred in a period of two and three-years of trading based on BHAR.

Furthermore, underpricing has shown statistical significance. Therefore, it is concluded that the higher the underpricing the higher is the long-run IPO performance in a period of two and three years.

The controlling variables; industry, region, market period, and financially backed is lacking a clear statistical prediction power. Especially, the variable industry shows no significant statistic power both when tested in combination with all the parameters together and individually. This finding is highly surprising since existing theory finds that industries have a significant difference in performance (Ritter, 1991). While the region parameter, in general, shows a significant prediction power of IPO performance despite that some regions are lacking some statistical significance in some of the models.

The last two parameters $VC/PE\ backed$, and $Hot\ market$ have a statistical significance only in a specific time horizon. The findings show that firms going public in a hot market period are underperforming, which is in line with existing theory. The same is the case with the VC/PE backed parameter, which is an indication of whether enterprises supported by a professional investor are performing better than the market based on a period of three years of trading.

Based on the research conducted in this paper is it possible to answer the question of research: based on a set of available public parameters is it feasible to estimate the performance of an IPO.

Appendix A

RStudio code

A.1 CAR Age

```
DATA <- CAR_ready
attach(DATA)
##########################################################################
# Logistics regression Age year 1
glm_mod_Car12_Age <- glm(I(CAR_12>0) ~ 1 + Age,family = binomial(link="logit")
    )
summary(glm_mod_Car12_Age)
confint(glm_mod_Car12_Age)
exp(coef(glm_mod_Car12_Age))
exp(confint(glm_mod_Car12_Age))
predict(glm_mod_Car12_Age, type="response")
residuals(glm_mod_Car12_Age, type="deviance")
library(stargazer)
stargazer(glm_mod_Car12_Age, title="Results", align=TRUE)
# Logistics regression Age year 1 with dummy #################
glm_mod_Car12_Age_dum <- glm(I(CAR_12>0) ~ 1 + Age + as.factor(Industry) + as.
    factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
    binomial(link="logit"))
summary(glm_mod_Car12_Age_dum)
confint(glm_mod_Car12_Age_dum)
exp(coef(glm_mod_Car12_Age_dum))
exp(confint(glm_mod_Car12_Age_dum))
stargazer(glm_mod_Car12_Age_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_Car12_Age_dum)
glm_mod_Car12_Age_dum_step<-glm(formula = I(CAR_12 > 0) ~ Age + as.factor(
    Region), family = binomial(link = "logit"))
summary(glm_mod_Car12_Age_dum_step)
```

```
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car12_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car12_Age_dum_step,type="response")>
   threshold,1,0)
actual_values<-glm_mod_Car12_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car12_Age_dum_step, title="Results", align=TRUE)

##########################################################################
# Logistics regression Age year 2
glm_mod_Car24_Age <- glm(I(CAR_24>0) ~ 1 + Age,family = binomial(link="logit")
   )
summary(glm_mod_Car24_Age)
confint(glm_mod_Car24_Age)
exp(coef(glm_mod_Car24_Age))
exp(confint(glm_mod_Car24_Age))
predict(glm_mod_Car24_Age, type="response")
residuals(glm_mod_Car24_Age, type="deviance")
stargazer(glm_mod_Car24_Age, title="Results", align=TRUE)
# Logistics regression Age year 2 with dummy #################
glm_mod_Car24_Age_dum <- glm(I(CAR_24>0) ~ 1 + Age + as.factor(Industry) + as.
   factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
   binomial(link="logit"))
summary(glm_mod_Car24_Age_dum)
confint(glm_mod_Car24_Age_dum)
exp(coef(glm_mod_Car24_Age_dum))
exp(confint(glm_mod_Car24_Age_dum))
stargazer(glm_mod_Car24_Age_dum, title="Results", align=TRUE)
####STEP###########################################################
step(glm_mod_Car24_Age_dum)
```

```
glm_mod_Car24_Age_dum_step<-glm(formula = I(CAR_24 > 0) ~ as.factor(Region) +
    as.factor(VC_PE_backed) + as.factor(Hot_market), family = binomial(link =
    "logit"))
summary(glm_mod_Car24_Age_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car24_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car24_Age_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_Car24_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car24_Age_dum_step, title="Results", align=TRUE)

##########################################################################
# Logistics regression Age year 3
glm_mod_Car36_Age <- glm(I(CAR_36>0) ~ 1 + Age,family = binomial(link="logit")
    )
summary(glm_mod_Car36_Age)
confint(glm_mod_Car36_Age)
exp(coef(glm_mod_Car36_Age))
exp(confint(glm_mod_Car36_Age))
predict(glm_mod_Car36_Age, type="response")
residuals(glm_mod_Car36_Age, type="deviance")
stargazer(glm_mod_Car36_Age, title="Results", align=TRUE)
# Logistics regression Age year 3 with dummy #################
glm_mod_Car36_Age_dum <- glm(I(CAR_36>0) ~ 1 + Age + as.factor(Industry) + as.
    factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
    binomial(link="logit"))
summary(glm_mod_Car36_Age_dum)
confint(glm_mod_Car36_Age_dum)
exp(coef(glm_mod_Car36_Age_dum))
exp(confint(glm_mod_Car36_Age_dum))
```

```
stargazer(glm_mod_Car36_Age_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_Car36_Age_dum)
glm_mod_Car36_Age_dum_step<-glm(formula = I(CAR_36 > 0) ~ Age + as.factor(
    Region) + as.factor(VC_PE_backed),family = binomial(link = "logit"))
summary(glm_mod_Car36_Age_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car36_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car36_Age_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_Car36_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car36_Age_dum_step, title="Results", align=TRUE)
```

A.2 CAR Size

```
DATA <- CAR_ready
attach(DATA)
##########################################################################
# Logistics regression Ln_Size year 1
glm_mod_Car12_Ln_Size <- glm(I(CAR_12>0) ~ 1 + Ln_Size,family = binomial(link=
    "logit"))
summary(glm_mod_Car12_Ln_Size)
confint(glm_mod_Car12_Ln_Size)
exp(coef(glm_mod_Car12_Ln_Size))
exp(confint(glm_mod_Car12_Ln_Size))
predict(glm_mod_Car12_Ln_Size, type="response")
residuals(glm_mod_Car12_Ln_Size, type="deviance")
stargazer(glm_mod_Car12_Ln_Size, title="Results", align=TRUE)
```

```
# Logistics regression Ln_Size year 1 with dummy #################
glm_mod_Car12_Ln_Size_dum <- glm(I(CAR_12>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_Car12_Ln_Size_dum)
confint(glm_mod_Car12_Ln_Size_dum)
exp(coef(glm_mod_Car12_Ln_Size_dum))
exp(confint(glm_mod_Car12_Ln_Size_dum))
stargazer(glm_mod_Car12_Ln_Size_dum, title="Results", align=TRUE)
#####STEP############################################################
step(glm_mod_Car12_Ln_Size_dum)
glm_mod_Car12_Ln_Size_dum_step<-glm(formula = I(CAR_12 > 0) ~ as.factor(Region
    ), family = binomial(link = "logit"))
summary(glm_mod_Car12_Ln_Size_dum_step)
#ROC curve     #######################################################
predpr <- predict(glm_mod_Car12_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car12_Ln_Size_dum_step,type="response
    ")>threshold,1,0)
actual_values<-glm_mod_Car12_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car12_Ln_Size_dum_step, title="Results", align=TRUE)

#########################################################################
# Logistics regression Ln_Size year 2
glm_mod_Car24_Ln_Size <- glm(I(CAR_24>0) ~ 1 + Ln_Size,family = binomial(link=
    "logit"))
summary(glm_mod_Car24_Ln_Size)
confint(glm_mod_Car24_Ln_Size)
exp(coef(glm_mod_Car24_Ln_Size))
exp(confint(glm_mod_Car24_Ln_Size))
predict(glm_mod_Car24_Ln_Size, type="response")
```

```
residuals(glm_mod_Car24_Ln_Size, type="deviance")
stargazer(glm_mod_Car24_Ln_Size, title="Results", align=TRUE)
# Logistics regression Ln_Size year 2 with dummy #################
glm_mod_Car24_Ln_Size_dum <- glm(I(CAR_24>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_Car24_Ln_Size_dum)
confint(glm_mod_Car24_Ln_Size_dum)
exp(coef(glm_mod_Car24_Ln_Size_dum))
exp(confint(glm_mod_Car24_Ln_Size_dum))
stargazer(glm_mod_Car24_Ln_Size_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_Car24_Ln_Size_dum)
glm_mod_Car24_Ln_Size_dum_step<- glm(formula = I(CAR_24 > 0) ~ as.factor(
    Region) + as.factor(VC_PE_backed) +
                                         as.factor(Hot_market), family =
                                            binomial(link = "logit"))
summary(glm_mod_Car24_Ln_Size_dum_step)
#ROC curve     #########################################################
predpr <- predict(glm_mod_Car24_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car24_Ln_Size_dum_step,type="response
    ")>threshold,1,0)
actual_values<-glm_mod_Car24_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car24_Ln_Size_dum_step, title="Results", align=TRUE)

##########################################################################
# Logistics regression Ln_Size year 3
glm_mod_Car36_Ln_Size <- glm(I(CAR_36>0) ~ 1 + Ln_Size,family = binomial(link=
    "logit"))
summary(glm_mod_Car36_Ln_Size)
```

```
confint(glm_mod_Car36_Ln_Size)
exp(coef(glm_mod_Car36_Ln_Size))
exp(confint(glm_mod_Car36_Ln_Size))
predict(glm_mod_Car36_Ln_Size, type="response")
residuals(glm_mod_Car36_Ln_Size, type="deviance")
stargazer(glm_mod_Car36_Ln_Size, title="Results", align=TRUE)
# Logistics regression Ln_Size year 3 with dummy #################
glm_mod_Car36_Ln_Size_dum <- glm(I(CAR_36>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_Car36_Ln_Size_dum)
confint(glm_mod_Car36_Ln_Size_dum)
exp(coef(glm_mod_Car36_Ln_Size_dum))
exp(confint(glm_mod_Car36_Ln_Size_dum))
stargazer(glm_mod_Car36_Ln_Size_dum, title="Results", align=TRUE)
#####STEP#############################################################
step(glm_mod_Car36_Ln_Size_dum)
glm_mod_Car36_Ln_Size_dum_step<-glm(formula = I(CAR_36 > 0) ~ Ln_Size + as.
    factor(Region) + as.factor(VC_PE_backed),
                                    family = binomial(link = "logit"))
summary(glm_mod_Car36_Ln_Size_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car36_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car36_Ln_Size_dum_step,type="response
    ")>threshold,1,0)
actual_values<-glm_mod_Car36_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car36_Ln_Size_dum_step, title="Results", align=TRUE)
```

A.3 CAR ROA

```
DATA <- CAR_ready
attach(DATA)
###################################################################
# Logistics regression ROA year 1
glm_mod_Car12_ROA <- glm(I(CAR_12>0) ~ 1 + ROA,family = binomial(link="logit")
   )
summary(glm_mod_Car12_ROA)
confint(glm_mod_Car12_ROA)
exp(coef(glm_mod_Car12_ROA))
exp(confint(glm_mod_Car12_ROA))
predict(glm_mod_Car12_ROA, type="response")
residuals(glm_mod_Car12_ROA, type="deviance")
stargazer(glm_mod_Car12_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 1 with dummy ################
glm_mod_Car12_ROA_dum <- glm(I(CAR_12>0) ~ 1 + ROA + as.factor(Industry) + as.
   factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
   binomial(link="logit"))
summary(glm_mod_Car12_ROA_dum)
confint(glm_mod_Car12_ROA_dum)
exp(coef(glm_mod_Car12_ROA_dum))
exp(confint(glm_mod_Car12_ROA_dum))
stargazer(glm_mod_Car12_ROA_dum, title="Results", align=TRUE)
#####STEP###########################################################
step(glm_mod_Car12_ROA_dum)
glm_mod_Car12_ROA_dum_step<-glm(formula = I(CAR_12 > 0) ~ as.factor(Region),
   family = binomial(link = "logit"))
summary(glm_mod_Car12_ROA_dum_step)
#ROC curve     #########################################################
predpr <- predict(glm_mod_Car12_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car12_ROA_dum_step,type="response")>
   threshold,1,0)
actual_values<-glm_mod_Car12_ROA_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
```

```
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car12_ROA_dum_step, title="Results", align=TRUE)

####################################################################
# Logistics regression ROA year 2
glm_mod_Car24_ROA <- glm(I(CAR_24>0) ~ 1 + ROA,family = binomial(link="logit")
    )
summary(glm_mod_Car24_ROA)
confint(glm_mod_Car24_ROA)
exp(coef(glm_mod_Car24_ROA))
exp(confint(glm_mod_Car24_ROA))
predict(glm_mod_Car24_ROA, type="response")
residuals(glm_mod_Car24_ROA, type="deviance")
stargazer(glm_mod_Car24_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 2 with dummy #################
glm_mod_Car24_ROA_dum <- glm(I(CAR_24>0) ~ 1 + ROA + as.factor(Industry) + as.
    factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
    binomial(link="logit"))
summary(glm_mod_Car24_ROA_dum)
confint(glm_mod_Car24_ROA_dum)
exp(coef(glm_mod_Car24_ROA_dum))
exp(confint(glm_mod_Car24_ROA_dum))
stargazer(glm_mod_Car24_ROA_dum, title="Results", align=TRUE)
#####STEP##################################################################
step(glm_mod_Car24_ROA_dum)
glm_mod_Car24_ROA_dum_step<-glm(formula = I(CAR_24 > 0) ~ ROA + as.factor(
    Region) + as.factor(VC_PE_backed) +
                                     as.factor(Hot_market), family = binomial(
                                        link = "logit"))
summary(glm_mod_Car24_ROA_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car24_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
```

```
predicted_values<-ifelse(predict(glm_mod_Car24_ROA_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_Car24_ROA_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car24_ROA_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression ROA year 3
glm_mod_Car36_ROA <- glm(I(CAR_36>0) ~ 1 + ROA,family = binomial(link="logit")
    )
summary(glm_mod_Car36_ROA)
confint(glm_mod_Car36_ROA)
exp(coef(glm_mod_Car36_ROA))
exp(confint(glm_mod_Car36_ROA))
predict(glm_mod_Car36_ROA, type="response")
residuals(glm_mod_Car36_ROA, type="deviance")
stargazer(glm_mod_Car36_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 3 with dummy #################
glm_mod_Car36_ROA_dum <- glm(I(CAR_36>0) ~ 1 + ROA + as.factor(Industry) + as.
    factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
    binomial(link="logit"))
summary(glm_mod_Car36_ROA_dum)
confint(glm_mod_Car36_ROA_dum)
exp(coef(glm_mod_Car36_ROA_dum))
exp(confint(glm_mod_Car36_ROA_dum))
stargazer(glm_mod_Car36_ROA_dum, title="Results", align=TRUE)
####STEP#################################################################
step(glm_mod_Car36_ROA_dum)
glm_mod_Car36_ROA_dum_step<-glm(formula = I(CAR_36 > 0) ~ as.factor(Region) +
    as.factor(VC_PE_backed),
                                    family = binomial(link = "logit"))
summary(glm_mod_Car36_ROA_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car36_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
```

```
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car36_ROA_dum_step,type="response")>
   threshold,1,0)
actual_values<-glm_mod_Car36_ROA_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car36_ROA_dum_step, title="Results", align=TRUE)
```

A.4 CAR Underpricing

```
DATA <- CAR_ready
attach(DATA)
###################################################################
# Logistics regression Underpricing year 1
glm_mod_Car12_Underpricing <- glm(I(CAR_12>0) ~ 1 + Underpricing,family =
   binomial(link="logit"))
summary(glm_mod_Car12_Underpricing)
confint(glm_mod_Car12_Underpricing)
exp(coef(glm_mod_Car12_Underpricing))
exp(confint(glm_mod_Car12_Underpricing))
predict(glm_mod_Car12_Underpricing, type="response")
residuals(glm_mod_Car12_Underpricing, type="deviance")
stargazer(glm_mod_Car12_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 1 with dummy #################
glm_mod_Car12_Underpricing_dum <- glm(I(CAR_12>0) ~ 1 + Underpricing + as.
   factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
   Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car12_Underpricing_dum)
confint(glm_mod_Car12_Underpricing_dum)
exp(coef(glm_mod_Car12_Underpricing_dum))
exp(confint(glm_mod_Car12_Underpricing_dum))
stargazer(glm_mod_Car12_Underpricing_dum, title="Results", align=TRUE)
#####STEP#############################################################
step(glm_mod_Car12_Underpricing_dum)
```

```
glm_mod_Car12_Underpricing_dum_step<-glm(formula = I(CAR_12 > 0) ~
    Underpricing + as.factor(Region),
                                              family = binomial(link = "logit"))
summary(glm_mod_Car12_Underpricing_dum_step)
#ROC curve      ##########################################################
predpr <- predict(glm_mod_Car12_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car12_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_Car12_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car12_Underpricing_dum_step, title="Results", align=TRUE)

####################################################################
# Logistics regression Underpricing year 2
glm_mod_Car24_Underpricing <- glm(I(CAR_24>0) ~ 1 + Underpricing,family =
    binomial(link="logit"))
summary(glm_mod_Car24_Underpricing)
confint(glm_mod_Car24_Underpricing)
exp(coef(glm_mod_Car24_Underpricing))
exp(confint(glm_mod_Car24_Underpricing))
predict(glm_mod_Car24_Underpricing, type="response")
residuals(glm_mod_Car24_Underpricing, type="deviance")
stargazer(glm_mod_Car24_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 2 with dummy #################
glm_mod_Car24_Underpricing_dum <- glm(I(CAR_24>0) ~ 1 + Underpricing + as.
    factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
    Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car24_Underpricing_dum)
confint(glm_mod_Car24_Underpricing_dum)
exp(coef(glm_mod_Car24_Underpricing_dum))
exp(confint(glm_mod_Car24_Underpricing_dum))
```

```
stargazer(glm_mod_Car24_Underpricing_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_Car24_Underpricing_dum)
glm_mod_Car24_Underpricing_dum_step<-glm(formula = I(CAR_24 > 0) ~
    Underpricing + as.factor(Region) +
                                              as.factor(VC_PE_backed) + as.
                                                 factor(Hot_market), family =
                                                  binomial(link = "logit"))
summary(glm_mod_Car24_Underpricing_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car24_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car24_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_Car24_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car24_Underpricing_dum_step, title="Results", align=TRUE)

####################################################################
# Logistics regression Underpricing year 3
glm_mod_Car36_Underpricing <- glm(I(CAR_36>0) ~ 1 + Underpricing,family =
    binomial(link="logit"))
summary(glm_mod_Car36_Underpricing)
confint(glm_mod_Car36_Underpricing)
exp(coef(glm_mod_Car36_Underpricing))
exp(confint(glm_mod_Car36_Underpricing))
predict(glm_mod_Car36_Underpricing, type="response")
residuals(glm_mod_Car36_Underpricing, type="deviance")
stargazer(glm_mod_Car36_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 3 with dummy #################
```

```
glm_mod_Car36_Underpricing_dum <- glm(I(CAR_36>0) ~ 1 + Underpricing + as.
    factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
    Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car36_Underpricing_dum)
confint(glm_mod_Car36_Underpricing_dum)
exp(coef(glm_mod_Car36_Underpricing_dum))
exp(confint(glm_mod_Car36_Underpricing_dum))
stargazer(glm_mod_Car36_Underpricing_dum, title="Results", align=TRUE)
#####STEP#############################################################
step(glm_mod_Car36_Underpricing_dum)

glm_mod_Car36_Underpricing_dum_step<-glm(formula = I(CAR_36 > 0) ~
    Underpricing + as.factor(Region) +
                                           as.factor(VC_PE_backed), family
                                              = binomial(link = "logit"))
summary(glm_mod_Car36_Underpricing_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car36_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car36_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_Car36_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car36_Underpricing_dum_step, title="Results", align=TRUE)
```

A.5 CAR Total

```
DATA <- CAR_ready
attach(DATA)
####################################################################
```

```
# Logistics regression total year 1
glm_mod_Car12_total <- glm(I(CAR_12>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing,family = binomial(link="logit"))
summary(glm_mod_Car12_total)
confint(glm_mod_Car12_total)
exp(coef(glm_mod_Car12_total))
exp(confint(glm_mod_Car12_total))
predict(glm_mod_Car12_total, type="response")
residuals(glm_mod_Car12_total, type="deviance")
stargazer(glm_mod_Car12_total, title="Results", align=TRUE)
# Logistics regression total year 1 with dummy #################
glm_mod_Car12_total_dum <- glm(I(CAR_12>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
    backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car12_total_dum)
confint(glm_mod_Car12_total_dum)
exp(coef(glm_mod_Car12_total_dum))
exp(confint(glm_mod_Car12_total_dum))
stargazer(glm_mod_Car12_total_dum, title="Results", align=TRUE)
#####STEP#############################################################
step(glm_mod_Car12_total_dum)
glm_mod_Car12_total_dum_step<-glm(formula = I(CAR_12 > 0) ~ Age + Ln_Size +
    Underpricing +
                                  as.factor(Region), family = binomial(link
                                      = "logit"))
summary(glm_mod_Car12_total_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car12_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car12_total_dum_step,type="response")
    >threshold,1,0)
actual_values<-glm_mod_Car12_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
```

```
specificity(conf_matrix)
stargazer(glm_mod_Car12_total_dum_step, title="Results", align=TRUE)

#####################################################################
# Logistics regression total year 2
glm_mod_Car24_total <- glm(I(CAR_24>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing,family = binomial(link="logit"))
summary(glm_mod_Car24_total)
confint(glm_mod_Car24_total)
exp(coef(glm_mod_Car24_total))
exp(confint(glm_mod_Car24_total))
predict(glm_mod_Car24_total, type="response")
residuals(glm_mod_Car24_total, type="deviance")
stargazer(glm_mod_Car24_total, title="Results", align=TRUE)
# Logistics regression total year 2 with dummy #################
glm_mod_Car24_total_dum <- glm(I(CAR_24>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
    backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car24_total_dum)
confint(glm_mod_Car24_total_dum)
exp(coef(glm_mod_Car24_total_dum))
exp(confint(glm_mod_Car24_total_dum))
stargazer(glm_mod_Car24_total_dum, title="Results", align=TRUE)
#####STEP################################################################
step(glm_mod_Car24_total_dum)
glm_mod_Car24_total_dum_step<-glm(formula = I(CAR_24 > 0) ~ Ln_Size +
    Underpricing + as.factor(Region) +
                                  as.factor(VC_PE_backed) + as.factor(Hot_
                                     market), family = binomial(link = "
                                     logit"))
summary(glm_mod_Car24_total_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_Car24_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car24_total_dum_step,type="response")
    >threshold,1,0)
```

```
actual_values<-glm_mod_Car24_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car24_total_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression total year 3
glm_mod_Car36_total <- glm(I(CAR_36>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing,family = binomial(link="logit"))
summary(glm_mod_Car36_total)
confint(glm_mod_Car36_total)
exp(coef(glm_mod_Car36_total))
exp(confint(glm_mod_Car36_total))
predict(glm_mod_Car36_total, type="response")
residuals(glm_mod_Car36_total, type="deviance")
stargazer(glm_mod_Car36_total, title="Results", align=TRUE)
# Logistics regression total year 3 with dummy #################
glm_mod_Car36_total_dum <- glm(I(CAR_36>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
    backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_Car36_total_dum)
confint(glm_mod_Car36_total_dum)
exp(coef(glm_mod_Car36_total_dum))
exp(confint(glm_mod_Car36_total_dum))
stargazer(glm_mod_Car36_total_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_Car36_total_dum)
glm_mod_Car36_total_dum_step<-glm(formula = I(CAR_36 > 0) ~ Age + Ln_Size +
    Underpricing +
                                    as.factor(Region) + as.factor(VC_PE_backed
                                    ), family = binomial(link = "logit"))
summary(glm_mod_Car36_total_dum_step)
#ROC curve       #########################################################
predpr <- predict(glm_mod_Car36_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(CAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
```

```
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_Car36_total_dum_step,type="response")
    >threshold,1,0)
actual_values<-glm_mod_Car36_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_Car36_total_dum_step, title="Results", align=TRUE)
```

A.6 BHAR Age

```
DATA2 <- BHAR_ready
attach(DATA2)
#################################################################
# Logistics regression Age year 1
glm_mod_BHAR12_Age <- glm(I(BHAR_12>0) ~ 1 + Age,family = binomial(link="logit
    "))
summary(glm_mod_BHAR12_Age)
confint(glm_mod_BHAR12_Age)
exp(coef(glm_mod_BHAR12_Age))
exp(confint(glm_mod_BHAR12_Age))
predict(glm_mod_BHAR12_Age, type="response")
residuals(glm_mod_BHAR12_Age, type="deviance")
stargazer(glm_mod_BHAR12_Age, title="Results", align=TRUE)
# Logistics regression Age year 1 with dummy #################
glm_mod_BHAR12_Age_dum <- glm(I(BHAR_12>0) ~ 1 + Age + as.factor(Industry) +
    as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
     binomial(link="logit"))
summary(glm_mod_BHAR12_Age_dum)
confint(glm_mod_BHAR12_Age_dum)
exp(coef(glm_mod_BHAR12_Age_dum))
exp(confint(glm_mod_BHAR12_Age_dum))
stargazer(glm_mod_BHAR12_Age_dum, title="Results", align=TRUE)
#####STEP########################################################
step(glm_mod_BHAR12_Age_dum)
glm_mod_BHAR12_Age_dum_step<-glm(formula = I(BHAR_12 > 0) ~ Age + as.factor(
    Region), family = binomial(link = "logit"))
```

```
summary(glm_mod_BHAR12_Age_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR12_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR12_Age_dum_step,type="response")>
   threshold,1,0)
actual_values<-glm_mod_BHAR12_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR12_Age_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression Age year 2
glm_mod_BHAR24_Age <- glm(I(BHAR_24>0) ~ 1 + Age,family = binomial(link="logit
   "))
summary(glm_mod_BHAR24_Age)
confint(glm_mod_BHAR24_Age)
exp(coef(glm_mod_BHAR24_Age))
exp(confint(glm_mod_BHAR24_Age))
predict(glm_mod_BHAR24_Age, type="response")
residuals(glm_mod_BHAR24_Age, type="deviance")
stargazer(glm_mod_BHAR24_Age, title="Results", align=TRUE)
# Logistics regression Age year 2 with dummy #################
glm_mod_BHAR24_Age_dum <- glm(I(BHAR_24>0) ~ 1 + Age + as.factor(Industry) +
   as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
    binomial(link="logit"))
summary(glm_mod_BHAR24_Age_dum)
confint(glm_mod_BHAR24_Age_dum)
exp(coef(glm_mod_BHAR24_Age_dum))
exp(confint(glm_mod_BHAR24_Age_dum))
stargazer(glm_mod_BHAR24_Age_dum, title="Results", align=TRUE)
#####STEP#############################################################
step(glm_mod_BHAR24_Age_dum)
```

```
glm_mod_BHAR24_Age_dum_step<-glm(formula = I(BHAR_24 > 0) ~ Age + as.factor(
    Region) + as.factor(Hot_market),
                                              family = binomial(link = "logit"))
summary(glm_mod_BHAR24_Age_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR24_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR24_Age_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_BHAR24_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR24_Age_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression Age year 3
glm_mod_BHAR36_Age <- glm(I(BHAR_36>0) ~ 1 + Age,family = binomial(link="logit
    "))
summary(glm_mod_BHAR36_Age)
confint(glm_mod_BHAR36_Age)
exp(coef(glm_mod_BHAR36_Age))
exp(confint(glm_mod_BHAR36_Age))
predict(glm_mod_BHAR36_Age, type="response")
residuals(glm_mod_BHAR36_Age, type="deviance")
stargazer(glm_mod_BHAR36_Age, title="Results", align=TRUE)
# Logistics regression Age year 3 with dummy #################
glm_mod_BHAR36_Age_dum <- glm(I(BHAR_36>0) ~ 1 + Age + as.factor(Industry) +
    as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
     binomial(link="logit"))
summary(glm_mod_BHAR36_Age_dum)
confint(glm_mod_BHAR36_Age_dum)
exp(coef(glm_mod_BHAR36_Age_dum))
exp(confint(glm_mod_BHAR36_Age_dum))
```

```
stargazer(glm_mod_BHAR36_Age_dum, title="Results", align=TRUE)
#####STEP#######################################################################
step(glm_mod_BHAR36_Age_dum)
glm_mod_BHAR36_Age_dum_step<-glm(formula = I(BHAR_36 > 0) ~ Age + as.factor(
    Region) + as.factor(VC_PE_backed),family = binomial(link = "logit"))
summary(glm_mod_BHAR36_Age_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR36_Age_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR36_Age_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_BHAR36_Age_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR36_Age_dum_step, title="Results", align=TRUE)
```

A.7 BHAR Size

```
DATA2 <- BHAR_ready
attach(DATA2)
##################################################################
# Logistics regression Ln_Size year 1
glm_mod_BHAR12_Ln_Size <- glm(I(BHAR_12>0) ~ 1 + Ln_Size,family = binomial(
    link="logit"))
summary(glm_mod_BHAR12_Ln_Size)
confint(glm_mod_BHAR12_Ln_Size)
exp(coef(glm_mod_BHAR12_Ln_Size))
exp(confint(glm_mod_BHAR12_Ln_Size))
predict(glm_mod_BHAR12_Ln_Size, type="response")
residuals(glm_mod_BHAR12_Ln_Size, type="deviance")
stargazer(glm_mod_BHAR12_Ln_Size, title="Results", align=TRUE)
```

```
# Logistics regression Ln_Size year 1 with dummy ################
glm_mod_BHAR12_Ln_Size_dum <- glm(I(BHAR_12>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_BHAR12_Ln_Size_dum)
confint(glm_mod_BHAR12_Ln_Size_dum)
exp(coef(glm_mod_BHAR12_Ln_Size_dum))
exp(confint(glm_mod_BHAR12_Ln_Size_dum))
#####STEP########################################################
step(glm_mod_BHAR12_Ln_Size_dum)
glm_mod_BHAR12_Ln_Size_dum_step<-glm(formula = I(BHAR_12 > 0) ~ as.factor(
    Region), family = binomial(link = "logit"))
summary(glm_mod_BHAR12_Ln_Size_dum_step)
#ROC curve    ######################################################
predpr <- predict(glm_mod_BHAR12_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR12_Ln_Size_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR12_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR12_Ln_Size_dum_re_step, title="Results", align=TRUE)

#################################################################
# Logistics regression Ln_Size year 2
glm_mod_BHAR24_Ln_Size <- glm(I(BHAR_24>0) ~ 1 + Ln_Size,family = binomial(
    link="logit"))
summary(glm_mod_BHAR24_Ln_Size)
confint(glm_mod_BHAR24_Ln_Size)
exp(coef(glm_mod_BHAR24_Ln_Size))
exp(confint(glm_mod_BHAR24_Ln_Size))
predict(glm_mod_BHAR24_Ln_Size, type="response")
residuals(glm_mod_BHAR24_Ln_Size, type="deviance")
```

```
stargazer(glm_mod_BHAR24_Ln_Size, title="Results", align=TRUE)
# Logistics regression Ln_Size year 2 with dummy #################
glm_mod_BHAR24_Ln_Size_dum <- glm(I(BHAR_24>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_BHAR24_Ln_Size_dum)
confint(glm_mod_BHAR24_Ln_Size_dum)
exp(coef(glm_mod_BHAR24_Ln_Size_dum))
exp(confint(glm_mod_BHAR24_Ln_Size_dum))
stargazer(glm_mod_BHAR24_Ln_Size_dum, title="Results", align=TRUE)
#####STEP##############################################################
step(glm_mod_BHAR24_Ln_Size_dum)
glm_mod_BHAR24_Ln_Size_dum_step<- glm(formula = I(BHAR_24 > 0) ~ as.factor(
    Region) + as.factor(Hot_market),
                                        family = binomial(link = "logit"))
summary(glm_mod_BHAR24_Ln_Size_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR24_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR24_Ln_Size_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR24_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR24_Ln_Size_dum_step, title="Results", align=TRUE)

##################################################################
# Logistics regression Ln_Size year 3
glm_mod_BHAR36_Ln_Size <- glm(I(BHAR_36>0) ~ 1 + Ln_Size,family = binomial(
    link="logit"))
summary(glm_mod_BHAR36_Ln_Size)
confint(glm_mod_BHAR36_Ln_Size)
exp(coef(glm_mod_BHAR36_Ln_Size))
```

```
exp(confint(glm_mod_BHAR36_Ln_Size))
predict(glm_mod_BHAR36_Ln_Size, type="response")
residuals(glm_mod_BHAR36_Ln_Size, type="deviance")
stargazer(glm_mod_BHAR36_Ln_Size, title="Results", align=TRUE)
# Logistics regression Ln_Size year 3 with dummy #################
glm_mod_BHAR36_Ln_Size_dum <- glm(I(BHAR_36>0) ~ 1 + Ln_Size + as.factor(
    Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_
    market),family = binomial(link="logit"))
summary(glm_mod_BHAR36_Ln_Size_dum)
confint(glm_mod_BHAR36_Ln_Size_dum)
exp(coef(glm_mod_BHAR36_Ln_Size_dum))
exp(confint(glm_mod_BHAR36_Ln_Size_dum))
stargazer(glm_mod_BHAR36_Ln_Size_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_BHAR36_Ln_Size_dum)
glm_mod_BHAR36_Ln_Size_dum_step<-glm(formula = I(BHAR_36 > 0) ~ as.factor(
    Region) + as.factor(VC_PE_backed),
                                        family = binomial(link = "logit"))
summary(glm_mod_BHAR36_Ln_Size_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR36_Ln_Size_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR36_Ln_Size_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR36_Ln_Size_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR36_Ln_Size_dum_step, title="Results", align=TRUE)
```

A.8 BHAR ROA

```
DATA2 <- BHAR_ready
attach(DATA2)
##################################################################
# Logistics regression ROA year 1
glm_mod_BHAR12_ROA <- glm(I(BHAR_12>0) ~ 1 + ROA,family = binomial(link="logit
    "))
summary(glm_mod_BHAR12_ROA)
confint(glm_mod_BHAR12_ROA)
exp(coef(glm_mod_BHAR12_ROA))
exp(confint(glm_mod_BHAR12_ROA))
predict(glm_mod_BHAR12_ROA, type="response")
residuals(glm_mod_BHAR12_ROA, type="deviance")
stargazer(glm_mod_BHAR12_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 1 with dummy #################
glm_mod_BHAR12_ROA_dum <- glm(I(BHAR_12>0) ~ 1 + ROA + as.factor(Industry) +
    as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
     binomial(link="logit"))
summary(glm_mod_BHAR12_ROA_dum)
confint(glm_mod_BHAR12_ROA_dum)
exp(coef(glm_mod_BHAR12_ROA_dum))
exp(confint(glm_mod_BHAR12_ROA_dum))
stargazer(glm_mod_BHAR12_ROA_dum, title="Results", align=TRUE)
####STEP#############################################################
step(glm_mod_BHAR12_ROA_dum)
glm_mod_BHAR12_ROA_dum_step<-glm(formula = I(BHAR_12 > 0) ~ as.factor(Region),
     family = binomial(link = "logit"))
summary(glm_mod_BHAR12_ROA_dum_step)
#ROC curve     #########################################################
predpr <- predict(glm_mod_BHAR12_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR12_ROA_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_BHAR12_ROA_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
```

```
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR12_ROA_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression ROA year 2
glm_mod_BHAR24_ROA <- glm(I(BHAR_24>0) ~ 1 + ROA,family = binomial(link="logit
    "))
summary(glm_mod_BHAR24_ROA)
confint(glm_mod_BHAR24_ROA)
exp(coef(glm_mod_BHAR24_ROA))
exp(confint(glm_mod_BHAR24_ROA))
predict(glm_mod_BHAR24_ROA, type="response")
residuals(glm_mod_BHAR24_ROA, type="deviance")
stargazer(glm_mod_BHAR24_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 2 with dummy #################
glm_mod_BHAR24_ROA_dum <- glm(I(BHAR_24>0) ~ 1 + ROA + as.factor(Industry) +
    as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
     binomial(link="logit"))
summary(glm_mod_BHAR24_ROA_dum)
confint(glm_mod_BHAR24_ROA_dum)
exp(coef(glm_mod_BHAR24_ROA_dum))
exp(confint(glm_mod_BHAR24_ROA_dum))
stargazer(glm_mod_BHAR24_ROA_dum, title="Results", align=TRUE)
#####STEP###########################################################
step(glm_mod_BHAR24_ROA_dum)
glm_mod_BHAR24_ROA_dum_step<-glm(formula = I(BHAR_24 > 0) ~ ROA + as.factor(
    Region) + as.factor(Hot_market),
                                    family = binomial(link = "logit"))
summary(glm_mod_BHAR24_ROA_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR24_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR24_ROA_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_BHAR24_ROA_dum_step$y
```

```
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR24_ROA_dum_step, title="Results", align=TRUE)

#################################################################
# Logistics regression ROA year 3
glm_mod_BHAR36_ROA <- glm(I(BHAR_36>0) ~ 1 + ROA,family = binomial(link="logit
    "))
summary(glm_mod_BHAR36_ROA)
confint(glm_mod_BHAR36_ROA)
exp(coef(glm_mod_BHAR36_ROA))
exp(confint(glm_mod_BHAR36_ROA))
predict(glm_mod_BHAR36_ROA, type="response")
residuals(glm_mod_BHAR36_ROA, type="deviance")
stargazer(glm_mod_BHAR36_ROA, title="Results", align=TRUE)
# Logistics regression ROA year 3 with dummy #################
glm_mod_BHAR36_ROA_dum <- glm(I(BHAR_36>0) ~ 1 + ROA + as.factor(Industry) +
    as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(Hot_market),family =
     binomial(link="logit"))
summary(glm_mod_BHAR36_ROA_dum)
confint(glm_mod_BHAR36_ROA_dum)
exp(coef(glm_mod_BHAR36_ROA_dum))
exp(confint(glm_mod_BHAR36_ROA_dum))
stargazer(glm_mod_BHAR36_ROA_dum, title="Results", align=TRUE)
#####STEP##############################################################
step(glm_mod_BHAR36_ROA_dum)

glm_mod_BHAR36_ROA_dum_step<-glm(formula = I(BHAR_36 > 0) ~ ROA + as.factor(
    Region) + as.factor(VC_PE_backed),
                                  family = binomial(link = "logit"))
summary(glm_mod_BHAR36_ROA_dum_step)
#ROC curve     #########################################################
predpr <- predict(glm_mod_BHAR36_ROA_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
```

```
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR36_ROA_dum_step,type="response")>
    threshold,1,0)
actual_values<-glm_mod_BHAR36_ROA_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR36_ROA_dum_step, title="Results", align=TRUE)
```

A.9 BHAR Underpricing

```
DATA2 <- BHAR_ready
attach(DATA2)
###################################################################
# Logistics regression Underpricing year 1
glm_mod_BHAR12_Underpricing <- glm(I(BHAR_12>0) ~ 1 + Underpricing,family =
    binomial(link="logit"))
summary(glm_mod_BHAR12_Underpricing)
confint(glm_mod_BHAR12_Underpricing)
exp(coef(glm_mod_BHAR12_Underpricing))
exp(confint(glm_mod_BHAR12_Underpricing))
predict(glm_mod_BHAR12_Underpricing, type="response")
residuals(glm_mod_BHAR12_Underpricing, type="deviance")
stargazer(glm_mod_BHAR12_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 1 with dummy #################
glm_mod_BHAR12_Underpricing_dum <- glm(I(BHAR_12>0) ~ 1 + Underpricing + as.
    factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
    Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR12_Underpricing_dum)
confint(glm_mod_BHAR12_Underpricing_dum)
exp(coef(glm_mod_BHAR12_Underpricing_dum))
exp(confint(glm_mod_BHAR12_Underpricing_dum))
stargazer(glm_mod_BHAR12_Underpricing_dum, title="Results", align=TRUE)
#####STEP############################################################
step(glm_mod_BHAR12_Underpricing_dum)
glm_mod_BHAR12_Underpricing_dum_step<-glm(formula = I(BHAR_12 > 0) ~
    Underpricing + as.factor(Region),
                                             family = binomial(link = "logit"))
```

```
summary(glm_mod_BHAR12_Underpricing_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR12_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR12_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR12_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR12_Underpricing_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression Underpricing year 2
glm_mod_BHAR24_Underpricing <- glm(I(BHAR_24>0) ~ 1 + Underpricing,family =
    binomial(link="logit"))
summary(glm_mod_BHAR24_Underpricing)
confint(glm_mod_BHAR24_Underpricing)
exp(coef(glm_mod_BHAR24_Underpricing))
exp(confint(glm_mod_BHAR24_Underpricing))
predict(glm_mod_BHAR24_Underpricing, type="response")
residuals(glm_mod_BHAR24_Underpricing, type="deviance")
stargazer(glm_mod_BHAR24_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 2 with dummy #################
glm_mod_BHAR24_Underpricing_dum <- glm(I(BHAR_24>0) ~ 1 + Underpricing + as.
    factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
    Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR24_Underpricing_dum)
confint(glm_mod_BHAR24_Underpricing_dum)
exp(coef(glm_mod_BHAR24_Underpricing_dum))
exp(confint(glm_mod_BHAR24_Underpricing_dum))
stargazer(glm_mod_BHAR24_Underpricing_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_BHAR24_Underpricing_dum)
```

```
glm_mod_BHAR24_Underpricing_dum_step<-glm(formula = I(BHAR_24 > 0) ~
    Underpricing + as.factor(Region) +
                                            as.factor(Hot_market), family =
                                                binomial(link = "logit"))
summary(glm_mod_BHAR24_Underpricing_dum_step)
#ROC curve     #########################################################
predpr <- predict(glm_mod_BHAR24_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR24_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR24_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR24_Underpricing_dum_step, title="Results", align=TRUE)

##################################################################
# Logistics regression Underpricing year 3
glm_mod_BHAR36_Underpricing <- glm(I(BHAR_36>0) ~ 1 + Underpricing,family =
    binomial(link="logit"))
summary(glm_mod_BHAR36_Underpricing)
confint(glm_mod_BHAR36_Underpricing)
exp(coef(glm_mod_BHAR36_Underpricing))
exp(confint(glm_mod_BHAR36_Underpricing))
predict(glm_mod_BHAR36_Underpricing, type="response")
residuals(glm_mod_BHAR36_Underpricing, type="deviance")
stargazer(glm_mod_BHAR36_Underpricing, title="Results", align=TRUE)
# Logistics regression Underpricing year 3 with dummy #################
glm_mod_BHAR36_Underpricing_dum <- glm(I(BHAR_36>0) ~ 1 + Underpricing + as.
    factor(Industry) + as.factor(Region)+ as.factor(VC_PE_backed)+ as.factor(
    Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR36_Underpricing_dum)
confint(glm_mod_BHAR36_Underpricing_dum)
exp(coef(glm_mod_BHAR36_Underpricing_dum))
```

```
exp(confint(glm_mod_BHAR36_Underpricing_dum))
stargazer(glm_mod_BHAR36_Underpricing_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_BHAR36_Underpricing_dum)
glm_mod_BHAR36_Underpricing_dum_step<-glm(formula = I(BHAR_36 > 0) ~
    Underpricing + as.factor(Region) +
                                                as.factor(VC_PE_backed), family
                                                   = binomial(link = "logit"))
summary(glm_mod_BHAR36_Underpricing_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR36_Underpricing_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR36_Underpricing_dum_step,type="
    response")>threshold,1,0)
actual_values<-glm_mod_BHAR36_Underpricing_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR36_Underpricing_dum_step, title="Results", align=TRUE)
```

A.10 BHAR Total

```
DATA2 <- BHAR_ready
attach(DATA2)
###################################################################

# Logistics regression total year 1
glm_mod_BHAR12_total <- glm(I(BHAR_12>0) ~ 1 + Age + Ln_Size + ROA +
    Underpricing,family = binomial(link="logit"))
summary(glm_mod_BHAR12_total)
confint(glm_mod_BHAR12_total)
exp(coef(glm_mod_BHAR12_total))
```

```
exp(confint(glm_mod_BHAR12_total))
predict(glm_mod_BHAR12_total, type="response")
residuals(glm_mod_BHAR12_total, type="deviance")
stargazer(glm_mod_BHAR12_total, title="Results", align=TRUE)
# Logistics regression total year 1 with dummy #################
glm_mod_BHAR12_total_dum <- glm(I(BHAR_12>0) ~ 1 + Age + Ln_Size + ROA +
   Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
   backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR12_total_dum)
confint(glm_mod_BHAR12_total_dum)
exp(coef(glm_mod_BHAR12_total_dum))
exp(confint(glm_mod_BHAR12_total_dum))
stargazer(glm_mod_BHAR12_total_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_BHAR12_total_dum)
glm_mod_BHAR12_total_dum_step<-glm(formula = I(BHAR_12 > 0) ~ Age + as.factor(
   Region), family = binomial(link = "logit"))
summary(glm_mod_BHAR12_total_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR12_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_12 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR12_total_dum_step,type="response"
   )>threshold,1,0)
actual_values<-glm_mod_BHAR12_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR12_total_dum_step, title="Results", align=TRUE)

###################################################################
# Logistics regression total year 2
glm_mod_BHAR24_total <- glm(I(BHAR_24>0) ~ 1 + Age + Ln_Size + ROA +
   Underpricing,family = binomial(link="logit"))
summary(glm_mod_BHAR24_total)
```

```
confint(glm_mod_BHAR24_total)
exp(coef(glm_mod_BHAR24_total))
exp(confint(glm_mod_BHAR24_total))
predict(glm_mod_BHAR24_total, type="response")
residuals(glm_mod_BHAR24_total, type="deviance")
stargazer(glm_mod_BHAR24_total, title="Results", align=TRUE)
# Logistics regression total year 2 with dummy #################
glm_mod_BHAR24_total_dum <- glm(I(BHAR_24>0) ~ 1 + Age + Ln_Size + ROA +
   Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
   backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR24_total_dum)
confint(glm_mod_BHAR24_total_dum)
exp(coef(glm_mod_BHAR24_total_dum))
exp(confint(glm_mod_BHAR24_total_dum))
stargazer(glm_mod_BHAR24_total_dum, title="Results", align=TRUE)
#####STEP##########################################################
step(glm_mod_BHAR24_total_dum)
glm_mod_BHAR24_total_dum_step<-glm(formula = I(BHAR_24 > 0) ~ Age + ROA +
   Underpricing + as.factor(Region) +
                                 as.factor(Hot_market), family = binomial(
                                    link = "logit"))
summary(glm_mod_BHAR24_total_dum_step)
#ROC curve     ########################################################
predpr <- predict(glm_mod_BHAR24_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_24 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR24_total_dum_step,type="response"
   )>threshold,1,0)
actual_values<-glm_mod_BHAR24_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
specificity(conf_matrix)
stargazer(glm_mod_BHAR24_total_dum_step, title="Results", align=TRUE)

##################################################################
```

```
# Logistics regression total year 3
glm_mod_BHAR36_total <- glm(I(BHAR_36>0) ~ 1 + Age + Ln_Size + ROA +
   Underpricing,family = binomial(link="logit"))
summary(glm_mod_BHAR36_total)
confint(glm_mod_BHAR36_total)
exp(coef(glm_mod_BHAR36_total))
exp(confint(glm_mod_BHAR36_total))
predict(glm_mod_BHAR36_total, type="response")
residuals(glm_mod_BHAR36_total, type="deviance")
stargazer(glm_mod_BHAR36_total, title="Results", align=TRUE)
# Logistics regression total year 3 with dummy ################
glm_mod_BHAR36_total_dum <- glm(I(BHAR_36>0) ~ 1 + Age + Ln_Size + ROA +
   Underpricing + as.factor(Industry) + as.factor(Region)+ as.factor(VC_PE_
   backed)+ as.factor(Hot_market),family = binomial(link="logit"))
summary(glm_mod_BHAR36_total_dum)
confint(glm_mod_BHAR36_total_dum)
exp(coef(glm_mod_BHAR36_total_dum))
exp(confint(glm_mod_BHAR36_total_dum))
stargazer(glm_mod_BHAR36_total_dum, title="Results", align=TRUE)
#####STEP#################################################################
step(glm_mod_BHAR36_total_dum)
glm_mod_BHAR36_total_dum_step<-glm(formula = I(BHAR_36 > 0) ~ Age + ROA +
   Underpricing + as.factor(Region) +
                                     as.factor(VC_PE_backed), family =
                                        binomial(link = "logit"))
summary(glm_mod_BHAR36_total_dum_step)
#ROC curve     ##########################################################
predpr <- predict(glm_mod_BHAR36_total_dum_step,type=c("response"))
library(pROC)
roccurve <- roc(I(BHAR_36 > 0) ~ predpr)
plot(roccurve)
auc(roccurve)
plot(ci.thresholds(roccurve))
#Type 1 and 2 error estimating ################################
threshold=0.5
predicted_values<-ifelse(predict(glm_mod_BHAR36_total_dum_step,type="response"
   )>threshold,1,0)
actual_values<-glm_mod_BHAR36_total_dum_step$y
conf_matrix<-table(predicted_values,actual_values)
conf_matrix
library(caret)
sensitivity(conf_matrix)
```

```
specificity(conf_matrix)
stargazer(glm_mod_BHAR36_total_dum_step, title="Results", align=TRUE)
```

Appendix B

Empirical results

B.1 CAR Age

TABLE B.1: Results

	Dependent variable:
	I(CAR_12 >0)
Age	0.004*
	(0.002)
as.factor(Industry)B	0.071
	(0.574)
as.factor(Industry)C	0.268
	(0.517)
as.factor(Industry)D	0.253
	(0.602)
as.factor(Industry)F	-0.110
	(0.562)
as.factor(Industry)G	0.353
	(0.547)
as.factor(Industry)H	-0.209
	(0.684)
as.factor(Industry)I	-0.959
	(0.949)
as.factor(Industry)J	0.127
	(0.527)
as.factor(Industry)K	0.579
	(0.546)
as.factor(Industry)L	-0.316
	(0.670)
as.factor(Industry)M	0.014
	(0.574)
as.factor(Industry)N	0.517
	(0.641)
as.factor(Industry)Q	-0.742
	(0.958)
as.factor(Industry)R	0.353
	(0.796)
as.factor(Region)NE	0.266
	(0.229)
as.factor(Region)SE	-0.443*
	(0.254)
as.factor(Region)WE	0.555***
	(0.164)
as.factor(VC_PE_backed)1	-0.065
	(0.214)
as.factor(Hot_market)1	-0.088
	(0.146)
Constant	-0.628
	(0.528)
Note:	$^{*}p < 0.1$;**p<0.05; ***p<0.01

TABLE B.2: Results

	Dependent variable:
	I(CAR_24 >0)
Age	0.002
	(0.002)
as.factor(Industry)B	−1.081*
	(0.580)
as.factor(Industry)C	−0.445
	(0.517)
as.factor(Industry)D	−0.891
	(0.606)
as.factor(Industry)F	−1.043*
	(0.565)
as.factor(Industry)G	−0.633
	(0.547)
as.factor(Industry)H	−0.750
	(0.676)
as.factor(Industry)I	−1.770*
	(0.950)
as.factor(Industry)J	−0.600
	(0.527)
as.factor(Industry)K	−0.461
	(0.545)
as.factor(Industry)L	−0.172
	(0.658)
as.factor(Industry)M	−0.518
	(0.572)
as.factor(Industry)N	−1.026
	(0.652)
as.factor(Industry)Q	−0.159
	(0.853)
as.factor(Industry)R	−0.518
	(0.795)
as.factor(Region)NE	0.403*
	(0.229)
as.factor(Region)SE	−0.382
	(0.252)
as.factor(Region)WE	0.428***
	(0.164)
as.factor(VC_PE_backed)1	0.376*
	(0.215)
as.factor(Hot_market)1	−0.281*
	(0.146)
Constant	0.310
	(0.527)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.3: Results

	Dependent variable:
	I(CAR_36 >0)
Age	0.004*
	(0.002)
as.factor(Industry)B	−0.366
	(0.568)
as.factor(Industry)C	−0.217
	(0.511)
as.factor(Industry)D	−0.611
	(0.600)
as.factor(Industry)F	−0.581
	(0.554)
as.factor(Industry)G	−0.300
	(0.541)
as.factor(Industry)H	−1.001
	(0.689)
as.factor(Industry)I	−0.085
	(0.837)
as.factor(Industry)J	−0.118
	(0.520)
as.factor(Industry)K	−0.301
	(0.539)
as.factor(Industry)L	−0.147
	(0.650)
as.factor(Industry)M	−0.470
	(0.567)
as.factor(Industry)N	−0.702
	(0.640)
as.factor(Industry)Q	−1.009
	(0.877)
as.factor(Industry)R	0.848
	(0.849)
as.factor(Region)NE	0.462**
	(0.229)
as.factor(Region)SE	−0.700***
	(0.253)
as.factor(Region)WE	0.179
	(0.162)
as.factor(VC_PE_backed)1	0.408*
	(0.217)
as.factor(Hot_market)1	0.040
	(0.145)
Constant	0.015
	(0.521)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.2 CAR Size

TABLE B.4: Results

	Dependent variable:
	I(CAR_12 >0)
Ln_Size	−0.034
	(0.027)
as.factor(Industry)B	0.092
	(0.574)
as.factor(Industry)C	0.318
	(0.516)
as.factor(Industry)D	0.264
	(0.601)
as.factor(Industry)F	−0.073
	(0.561)
as.factor(Industry)G	0.372
	(0.546)
as.factor(Industry)H	−0.026
	(0.679)
as.factor(Industry)I	−0.998
	(0.950)
as.factor(Industry)J	0.134
	(0.526)
as.factor(Industry)K	0.587
	(0.544)
as.factor(Industry)L	−0.318
	(0.669)
as.factor(Industry)M	0.016
	(0.573)
as.factor(Industry)N	0.520
	(0.640)
as.factor(Industry)Q	−0.759
	(0.958)
as.factor(Industry)R	0.352
	(0.797)
as.factor(Region)NE	0.244
	(0.228)
as.factor(Region)SE	−0.390
	(0.253)
as.factor(Region)WE	0.551***
	(0.164)
as.factor(VC_PE_backed)1	−0.032
	(0.214)
as.factor(Hot_market)1	−0.066
	(0.145)
Constant	−0.295
	(0.576)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.5: Results

	Dependent variable:
	I(CAR_24 >0)
Ln_Size	−0.034
	(0.027)
as.factor(Industry)B	−1.062*
	(0.579)
as.factor(Industry)C	−0.412
	(0.515)
as.factor(Industry)D	−0.876
	(0.606)
as.factor(Industry)F	−1.011*
	(0.564)
as.factor(Industry)G	−0.621
	(0.546)
as.factor(Industry)H	−0.620
	(0.673)
as.factor(Industry)I	−1.798*
	(0.950)
as.factor(Industry)J	−0.592
	(0.525)
as.factor(Industry)K	−0.452
	(0.543)
as.factor(Industry)L	−0.177
	(0.657)
as.factor(Industry)M	−0.517
	(0.571)
as.factor(Industry)N	−1.024
	(0.651)
as.factor(Industry)Q	−0.163
	(0.853)
as.factor(Industry)R	−0.521
	(0.797)
as.factor(Region)NE	0.392*
	(0.229)
as.factor(Region)SE	−0.344
	(0.251)
as.factor(Region)WE	0.435***
	(0.165)
as.factor(VC_PE_backed)1	0.401*
	(0.215)
as.factor(Hot_market)1	−0.266*
	(0.146)
Constant	0.626
	(0.576)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.6: Results

	Dependent variable:
	I(CAR_36 >0)
Ln_Size	−0.041
	(0.027)
as.factor(Industry)B	−0.336
	(0.567)
as.factor(Industry)C	−0.160
	(0.508)
as.factor(Industry)D	−0.595
	(0.599)
as.factor(Industry)F	−0.536
	(0.553)
as.factor(Industry)G	−0.278
	(0.539)
as.factor(Industry)H	−0.795
	(0.683)
as.factor(Industry)I	−0.120
	(0.838)
as.factor(Industry)J	−0.107
	(0.518)
as.factor(Industry)K	−0.287
	(0.537)
as.factor(Industry)L	−0.149
	(0.649)
as.factor(Industry)M	−0.466
	(0.565)
as.factor(Industry)N	−0.697
	(0.639)
as.factor(Industry)Q	−1.021
	(0.876)
as.factor(Industry)R	0.858
	(0.850)
as.factor(Region)NE	0.438^{*}
	(0.229)
as.factor(Region)SE	-0.641^{**}
	(0.252)
as.factor(Region)WE	0.178
	(0.162)
as.factor(VC_PE_backed)1	0.443^{**}
	(0.217)
as.factor(Hot_market)1	0.063
	(0.145)
Constant	0.411
	(0.568)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.3 CAR ROA

TABLE B.7: Results

	Dependent variable:
	I(CAR_12 >0)
ROA	0.059
	(0.050)
as.factor(Industry)B	0.118
	(0.576)
as.factor(Industry)C	0.347
	(0.518)
as.factor(Industry)D	0.282
	(0.602)
as.factor(Industry)F	−0.058
	(0.563)
as.factor(Industry)G	0.407
	(0.548)
as.factor(Industry)H	−0.055
	(0.679)
as.factor(Industry)I	−0.933
	(0.950)
as.factor(Industry)J	0.161
	(0.528)
as.factor(Industry)K	0.616
	(0.546)
as.factor(Industry)L	−0.314
	(0.671)
as.factor(Industry)M	0.066
	(0.575)
as.factor(Industry)N	0.572
	(0.643)
as.factor(Industry)Q	−0.784
	(0.960)
as.factor(Industry)R	0.398
	(0.798)
as.factor(Region)NE	0.272
	(0.230)
as.factor(Region)SE	−0.423*
	(0.253)
as.factor(Region)WE	0.549***
	(0.164)
as.factor(VC_PE_backed)1	−0.055
	(0.214)
as.factor(Hot_market)1	−0.080
	(0.145)
Constant	−0.637
	(0.529)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.8: Results

	Dependent variable:
	I(CAR_24 >0)
ROA	0.062
	(0.051)
as.factor(Industry)B	−1.037*
	(0.581)
as.factor(Industry)C	−0.387
	(0.518)
as.factor(Industry)D	−0.860
	(0.607)
as.factor(Industry)F	−1.002*
	(0.566)
as.factor(Industry)G	−0.589
	(0.549)
as.factor(Industry)H	−0.655
	(0.673)
as.factor(Industry)I	−1.740*
	(0.951)
as.factor(Industry)J	−0.569
	(0.528)
as.factor(Industry)K	−0.428
	(0.546)
as.factor(Industry)L	−0.172
	(0.660)
as.factor(Industry)M	−0.469
	(0.574)
as.factor(Industry)N	−0.975
	(0.653)
as.factor(Industry)Q	−0.189
	(0.856)
as.factor(Industry)R	−0.479
	(0.798)
as.factor(Region)NE	0.422*
	(0.230)
as.factor(Region)SE	−0.377
	(0.251)
as.factor(Region)WE	0.433***
	(0.165)
as.factor(VC_PE_backed)1	0.379*
	(0.214)
as.factor(Hot_market)1	−0.280*
	(0.146)
Constant	0.284
	(0.529)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.9: Results

	Dependent variable:
	I(CAR_36 >0)
ROA	0.066
	(0.051)
as.factor(Industry)B	−0.315
	(0.570)
as.factor(Industry)C	−0.133
	(0.511)
as.factor(Industry)D	−0.579
	(0.601)
as.factor(Industry)F	−0.525
	(0.555)
as.factor(Industry)G	−0.243
	(0.542)
as.factor(Industry)H	−0.838
	(0.683)
as.factor(Industry)I	−0.055
	(0.838)
as.factor(Industry)J	−0.081
	(0.521)
as.factor(Industry)K	−0.260
	(0.540)
as.factor(Industry)L	−0.142
	(0.652)
as.factor(Industry)M	−0.413
	(0.569)
as.factor(Industry)N	−0.641
	(0.642)
as.factor(Industry)Q	−1.051
	(0.879)
as.factor(Industry)R	0.907
	(0.852)
as.factor(Region)NE	0.470^{**}
	(0.230)
as.factor(Region)SE	-0.681^{***}
	(0.252)
as.factor(Region)WE	0.174
	(0.162)
as.factor(VC_PE_backed)1	0.415^{*}
	(0.216)
as.factor(Hot_market)1	0.047
	(0.145)
Constant	0.003
	(0.522)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.4 CAR Underpricing

TABLE B.10: Results

	Dependent variable:
	I(CAR_12 >0)
Underpricing	−0.087
	(0.054)
as.factor(Industry)B	0.083
	(0.575)
as.factor(Industry)C	0.309
	(0.516)
as.factor(Industry)D	0.269
	(0.601)
as.factor(Industry)F	−0.101
	(0.562)
as.factor(Industry)G	0.392
	(0.547)
as.factor(Industry)H	−0.081
	(0.678)
as.factor(Industry)I	−1.007
	(0.949)
as.factor(Industry)J	0.134
	(0.527)
as.factor(Industry)K	0.602
	(0.545)
as.factor(Industry)L	−0.314
	(0.670)
as.factor(Industry)M	0.030
	(0.573)
as.factor(Industry)N	0.541
	(0.641)
as.factor(Industry)Q	−0.782
	(0.957)
as.factor(Industry)R	0.310
	(0.797)
as.factor(Region)NE	0.268
	(0.229)
as.factor(Region)SE	−0.401
	(0.253)
as.factor(Region)WE	0.532^{***}
	(0.164)
as.factor(VC_PE_backed)1	−0.047
	(0.214)
as.factor(Hot_market)1	−0.073
	(0.145)
Constant	−0.588
	(0.527)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.11: Results

	Dependent variable:
	I(CAR_24 >0)
Underpricing	-0.134^{**}
	(0.056)
as.factor(Industry)B	-1.079^{*}
	(0.582)
as.factor(Industry)C	-0.429
	(0.517)
as.factor(Industry)D	-0.867
	(0.607)
as.factor(Industry)F	-1.053^{*}
	(0.566)
as.factor(Industry)G	-0.598
	(0.549)
as.factor(Industry)H	-0.678
	(0.673)
as.factor(Industry)I	-1.835^{*}
	(0.951)
as.factor(Industry)J	-0.596
	(0.528)
as.factor(Industry)K	-0.437
	(0.546)
as.factor(Industry)L	-0.175
	(0.661)
as.factor(Industry)M	-0.504
	(0.573)
as.factor(Industry)N	-1.002
	(0.653)
as.factor(Industry)Q	-0.210
	(0.853)
as.factor(Industry)R	-0.598
	(0.797)
as.factor(Region)NE	0.432^{*}
	(0.230)
as.factor(Region)SE	-0.346
	(0.251)
as.factor(Region)WE	0.416^{**}
	(0.164)
as.factor(VC_PE_backed)1	0.389^{*}
	(0.215)
as.factor(Hot_market)1	-0.274^{*}
	(0.146)
Constant	0.336
	(0.528)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.12: Results

	Dependent variable:
	I(CAR_36 >0)
Underpricing	-0.161^{***}
	(0.056)
as.factor(Industry)B	-0.347
	(0.571)
as.factor(Industry)C	-0.178
	(0.511)
as.factor(Industry)D	-0.589
	(0.602)
as.factor(Industry)F	-0.582
	(0.556)
as.factor(Industry)G	-0.246
	(0.542)
as.factor(Industry)H	-0.862
	(0.683)
as.factor(Industry)I	-0.163
	(0.840)
as.factor(Industry)J	-0.108
	(0.521)
as.factor(Industry)K	-0.266
	(0.540)
as.factor(Industry)L	-0.148
	(0.654)
as.factor(Industry)M	-0.449
	(0.568)
as.factor(Industry)N	-0.666
	(0.642)
as.factor(Industry)Q	-1.072
	(0.878)
as.factor(Industry)R	0.770
	(0.852)
as.factor(Region)NE	0.489^{**}
	(0.231)
as.factor(Region)SE	-0.646^{**}
	(0.252)
as.factor(Region)WE	0.154
	(0.162)
as.factor(VC_PE_backed)1	0.426^{**}
	(0.217)
as.factor(Hot_market)1	0.055
	(0.146)
Constant	0.059
	(0.521)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.5 CAR Total

TABLE B.13: Results

	Dependent variable:
	I(CAR_12 >0)
Age	0.004
	(0.002)
Ln_Size	−0.041
	(0.027)
ROA	0.050
	(0.050)
Underpricing	−0.091*
	(0.054)
as.factor(Industry)B	0.133
	(0.578)
as.factor(Industry)C	0.300
	(0.520)
as.factor(Industry)D	0.315
	(0.604)
as.factor(Industry)F	−0.068
	(0.565)
as.factor(Industry)G	0.397
	(0.550)
as.factor(Industry)H	−0.094
	(0.688)
as.factor(Industry)I	−1.001
	(0.953)
as.factor(Industry)J	0.157
	(0.529)
as.factor(Industry)K	0.623
	(0.548)
as.factor(Industry)L	−0.353
	(0.674)
as.factor(Industry)M	0.057
	(0.577)
as.factor(Industry)N	0.567
	(0.645)
as.factor(Industry)Q	−0.790
	(0.961)
as.factor(Industry)R	0.317
	(0.799)
as.factor(Region)NE	0.323
	(0.232)
as.factor(Region)SE	−0.420
	(0.256)
as.factor(Region)WE	0.586***
	(0.166)
as.factor(VC_PE_backed)1	−0.047
	(0.215)
as.factor(Hot_market)1	−0.083
	(0.146)
Constant	−0.308
	(0.582)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.14: Results

	Dependent variable:
	I(CAR_24 >0)
Age	0.002
	(0.002)
Ln_Size	−0.042
	(0.027)
ROA	0.055
	(0.051)
Underpricing	−0.141**
	(0.056)
as.factor(Industry)B	−1.029*
	(0.585)
as.factor(Industry)C	−0.418
	(0.520)
as.factor(Industry)D	−0.826
	(0.609)
as.factor(Industry)F	−1.015*
	(0.569)
as.factor(Industry)G	−0.587
	(0.551)
as.factor(Industry)H	−0.631
	(0.680)
as.factor(Industry)I	−1.836*
	(0.953)
as.factor(Industry)J	−0.573
	(0.530)
as.factor(Industry)K	−0.417
	(0.548)
as.factor(Industry)L	−0.200
	(0.665)
as.factor(Industry)M	−0.474
	(0.575)
as.factor(Industry)N	−0.977
	(0.655)
as.factor(Industry)Q	−0.220
	(0.858)
as.factor(Industry)R	−0.591
	(0.800)
as.factor(Region)NE	0.479**
	(0.233)
as.factor(Region)SE	−0.348
	(0.253)
as.factor(Region)WE	0.460***
	(0.166)
as.factor(VC_PE_backed)1	0.398*
	(0.216)
as.factor(Hot_market)1	−0.277*
	(0.147)
Constant	0.647
	(0.582)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.15: Results

	Dependent variable:
	I(CAR_36 >0)
Age	0.004
	(0.002)
Ln_Size	-0.051^{*}
	(0.027)
ROA	0.056
	(0.051)
Underpricing	-0.167^{***}
	(0.057)
as.factor(Industry)B	-0.293
	(0.575)
as.factor(Industry)C	-0.189
	(0.515)
as.factor(Industry)D	-0.543
	(0.605)
as.factor(Industry)F	-0.546
	(0.559)
as.factor(Industry)G	-0.246
	(0.546)
as.factor(Industry)H	-0.862
	(0.694)
as.factor(Industry)I	-0.157
	(0.845)
as.factor(Industry)J	-0.086
	(0.524)
as.factor(Industry)K	-0.252
	(0.543)
as.factor(Industry)L	-0.186
	(0.659)
as.factor(Industry)M	-0.425
	(0.572)
as.factor(Industry)N	-0.646
	(0.645)
as.factor(Industry)Q	-1.083
	(0.881)
as.factor(Industry)R	0.762
	(0.853)
as.factor(Region)NE	0.549^{**}
	(0.234)
as.factor(Region)SE	-0.663^{***}
	(0.255)
as.factor(Region)WE	0.211
	(0.164)
as.factor(VC_PE_backed)1	0.436^{**}
	(0.218)
as.factor(Hot_market)1	0.048
	(0.147)
Constant	0.436
	(0.577)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.6 BHAR Age

TABLE B.16: Results

	Dependent variable:
	I(BHAR_12 >0)
Age	0.005**
	(0.002)
as.factor(Industry)B	−0.058
	(0.594)
as.factor(Industry)C	0.082
	(0.534)
as.factor(Industry)D	0.302
	(0.618)
as.factor(Industry)F	−0.290
	(0.586)
as.factor(Industry)G	0.287
	(0.564)
as.factor(Industry)H	−0.022
	(0.700)
as.factor(Industry)I	−0.652
	(0.962)
as.factor(Industry)J	0.124
	(0.543)
as.factor(Industry)K	0.441
	(0.561)
as.factor(Industry)L	−0.288
	(0.693)
as.factor(Industry)M	−0.170
	(0.595)
as.factor(Industry)N	0.820
	(0.655)
as.factor(Industry)Q	−1.387
	(1.197)
as.factor(Industry)R	−0.181
	(0.862)
as.factor(Region)NE	−0.028
	(0.250)
as.factor(Region)SE	−0.136
	(0.265)
as.factor(Region)WE	0.736***
	(0.172)
as.factor(VC_PE_backed)1	0.135
	(0.217)
as.factor(Hot_market)1	−0.042
	(0.151)
Constant	−1.055*
	(0.546)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.17: Results

	Dependent variable:
	I(BHAR_24 >0)
Age	0.005*
	(0.002)
as.factor(Industry)B	−0.263
	(0.673)
as.factor(Industry)C	0.610
	(0.594)
as.factor(Industry)D	0.069
	(0.699)
as.factor(Industry)F	−0.225
	(0.664)
as.factor(Industry)G	0.458
	(0.626)
as.factor(Industry)H	0.409
	(0.755)
as.factor(Industry)I	−0.921
	(1.214)
as.factor(Industry)J	0.595
	(0.603)
as.factor(Industry)K	0.570
	(0.623)
as.factor(Industry)L	0.939
	(0.722)
as.factor(Industry)M	0.678
	(0.647)
as.factor(Industry)N	0.481
	(0.723)
as.factor(Industry)Q	0.783
	(0.933)
as.factor(Industry)R	0.363
	(0.899)
as.factor(Region)NE	0.510**
	(0.255)
as.factor(Region)SE	0.133
	(0.280)
as.factor(Region)WE	0.881***
	(0.186)
as.factor(VC_PE_backed)1	0.083
	(0.221)
as.factor(Hot_market)1	−0.464***
	(0.155)
Constant	−1.529**
	(0.607)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.18: Results

	Dependent variable:
	I(BHAR_36 >0)
Age	0.007***
	(0.002)
as.factor(Industry)B	−0.097
	(0.663)
as.factor(Industry)C	0.281
	(0.595)
as.factor(Industry)D	−0.402
	(0.734)
as.factor(Industry)F	−0.043
	(0.651)
as.factor(Industry)G	0.144
	(0.628)
as.factor(Industry)H	−0.285
	(0.796)
as.factor(Industry)I	−0.236
	(0.997)
as.factor(Industry)J	0.401
	(0.604)
as.factor(Industry)K	0.556
	(0.622)
as.factor(Industry)L	1.174
	(0.719)
as.factor(Industry)M	0.369
	(0.651)
as.factor(Industry)N	−0.886
	(0.846)
as.factor(Industry)Q	0.296
	(0.946)
as.factor(Industry)R	0.157
	(0.903)
as.factor(Region)NE	0.596**
	(0.253)
as.factor(Region)SE	−0.343
	(0.301)
as.factor(Region)WE	0.601***
	(0.187)
as.factor(VC_PE_backed)1	0.777***
	(0.218)
as.factor(Hot_market)1	−0.157
	(0.160)
Constant	−1.532**
	(0.609)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.7 BHAR Size

TABLE B.19: Results

	Dependent variable:
	I(BHAR_12 >0)
Ln_Size	−0.010
	(0.028)
as.factor(Industry)B	−0.047
	(0.593)
as.factor(Industry)C	0.147
	(0.532)
as.factor(Industry)D	0.303
	(0.617)
as.factor(Industry)F	−0.258
	(0.585)
as.factor(Industry)G	0.318
	(0.562)
as.factor(Industry)H	0.159
	(0.692)
as.factor(Industry)I	−0.679
	(0.963)
as.factor(Industry)J	0.133
	(0.542)
as.factor(Industry)K	0.454
	(0.559)
as.factor(Industry)L	−0.274
	(0.691)
as.factor(Industry)M	−0.159
	(0.594)
as.factor(Industry)N	0.831
	(0.654)
as.factor(Industry)Q	−1.408
	(1.196)
as.factor(Industry)R	−0.171
	(0.862)
as.factor(Region)NE	−0.060
	(0.249)
as.factor(Region)SE	−0.087
	(0.263)
as.factor(Region)WE	0.712***
	(0.172)
as.factor(VC_PE_backed)1	0.160
	(0.216)
as.factor(Hot_market)1	−0.021
	(0.151)
Constant	−0.913
	(0.596)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.20: Results

	Dependent variable:
	I(BHAR_24 >0)
Ln_Size	−0.007
	(0.029)
as.factor(Industry)B	−0.261
	(0.673)
as.factor(Industry)C	0.661
	(0.593)
as.factor(Industry)D	0.067
	(0.698)
as.factor(Industry)F	−0.199
	(0.663)
as.factor(Industry)G	0.485
	(0.625)
as.factor(Industry)H	0.566
	(0.748)
as.factor(Industry)I	−0.943
	(1.214)
as.factor(Industry)J	0.601
	(0.602)
as.factor(Industry)K	0.582
	(0.622)
as.factor(Industry)L	0.944
	(0.721)
as.factor(Industry)M	0.683
	(0.645)
as.factor(Industry)N	0.491
	(0.722)
as.factor(Industry)Q	0.761
	(0.932)
as.factor(Industry)R	0.368
	(0.899)
as.factor(Region)NE	0.482*
	(0.254)
as.factor(Region)SE	0.176
	(0.279)
as.factor(Region)WE	0.862***
	(0.186)
as.factor(VC_PE_backed)1	0.104
	(0.220)
as.factor(Hot_market)1	−0.445***
	(0.155)
Constant	−1.424**
	(0.657)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.21: Results

	Dependent variable:
	I(BHAR_36 >0)
Ln_Size	0.0004
	(0.029)
as.factor(Industry)B	−0.095
	(0.662)
as.factor(Industry)C	0.355
	(0.593)
as.factor(Industry)D	−0.409
	(0.732)
as.factor(Industry)F	−0.015
	(0.650)
as.factor(Industry)G	0.183
	(0.627)
as.factor(Industry)H	−0.059
	(0.781)
as.factor(Industry)I	−0.260
	(0.997)
as.factor(Industry)J	0.408
	(0.602)
as.factor(Industry)K	0.572
	(0.620)
as.factor(Industry)L	1.182*
	(0.717)
as.factor(Industry)M	0.378
	(0.649)
as.factor(Industry)N	−0.868
	(0.845)
as.factor(Industry)Q	0.264
	(0.944)
as.factor(Industry)R	0.161
	(0.904)
as.factor(Region)NE	0.551**
	(0.251)
as.factor(Region)SE	−0.284
	(0.299)
as.factor(Region)WE	0.564***
	(0.187)
as.factor(VC_PE_backed)1	0.797***
	(0.217)
as.factor(Hot_market)1	−0.133
	(0.159)
Constant	−1.467**
	(0.659)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.8 BHAR ROA

TABLE B.22: Results

	Dependent variable:
	I(BHAR_12 >0)
ROA	0.073
	(0.051)
as.factor(Industry)B	0.001
	(0.595)
as.factor(Industry)C	0.186
	(0.534)
as.factor(Industry)D	0.337
	(0.618)
as.factor(Industry)F	−0.226
	(0.587)
as.factor(Industry)G	0.357
	(0.564)
as.factor(Industry)H	0.179
	(0.691)
as.factor(Industry)I	−0.624
	(0.963)
as.factor(Industry)J	0.167
	(0.544)
as.factor(Industry)K	0.489
	(0.561)
as.factor(Industry)L	−0.287
	(0.695)
as.factor(Industry)M	−0.103
	(0.596)
as.factor(Industry)N	0.889
	(0.657)
as.factor(Industry)Q	−1.443
	(1.198)
as.factor(Industry)R	−0.124
	(0.863)
as.factor(Region)NE	−0.022
	(0.251)
as.factor(Region)SE	−0.110
	(0.263)
as.factor(Region)WE	0.725^{***}
	(0.172)
as.factor(VC_PE_backed)1	0.148
	(0.216)
as.factor(Hot_market)1	−0.032
	(0.151)
Constant	-1.061^{*}
	(0.547)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.23: Results

	Dependent variable:
	I(BHAR_24 >0)
ROA	0.142***
	(0.054)
as.factor(Industry)B	−0.148
	(0.676)
as.factor(Industry)C	0.750
	(0.597)
as.factor(Industry)D	0.152
	(0.701)
as.factor(Industry)F	−0.132
	(0.667)
as.factor(Industry)G	0.569
	(0.629)
as.factor(Industry)H	0.642
	(0.750)
as.factor(Industry)I	−0.838
	(1.215)
as.factor(Industry)J	0.679
	(0.606)
as.factor(Industry)K	0.658
	(0.626)
as.factor(Industry)L	0.944
	(0.728)
as.factor(Industry)M	0.805
	(0.651)
as.factor(Industry)N	0.609
	(0.727)
as.factor(Industry)Q	0.704
	(0.944)
as.factor(Industry)R	0.461
	(0.903)
as.factor(Region)NE	0.563**
	(0.256)
as.factor(Region)SE	0.142
	(0.279)
as.factor(Region)WE	0.896***
	(0.186)
as.factor(VC_PE_backed)1	0.089
	(0.221)
as.factor(Hot_market)1	−0.465***
	(0.156)
Constant	−1.611***
	(0.611)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.24: Results

	Dependent variable:
	I(BHAR_36 >0)
ROA	0.138**
	(0.055)
as.factor(Industry)B	0.015
	(0.666)
as.factor(Industry)C	0.443
	(0.597)
as.factor(Industry)D	−0.320
	(0.736)
as.factor(Industry)F	0.065
	(0.655)
as.factor(Industry)G	0.267
	(0.631)
as.factor(Industry)H	0.030
	(0.783)
as.factor(Industry)I	−0.161
	(1.000)
as.factor(Industry)J	0.487
	(0.607)
as.factor(Industry)K	0.650
	(0.625)
as.factor(Industry)L	1.203*
	(0.725)
as.factor(Industry)M	0.497
	(0.654)
as.factor(Industry)N	−0.753
	(0.849)
as.factor(Industry)Q	0.217
	(0.951)
as.factor(Industry)R	0.253
	(0.908)
as.factor(Region)NE	0.628**
	(0.253)
as.factor(Region)SE	−0.313
	(0.299)
as.factor(Region)WE	0.600***
	(0.187)
as.factor(VC_PE_backed)1	0.791***
	(0.218)
as.factor(Hot_market)1	−0.147
	(0.160)
Constant	−1.590***
	(0.613)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.9 BHAR Underpricing

TABLE B.25: Results

	Dependent variable:
	I(BHAR_12 >0)
Underpricing	−0.099*
	(0.057)
as.factor(Industry)B	−0.050
	(0.594)
as.factor(Industry)C	0.138
	(0.532)
as.factor(Industry)D	0.319
	(0.617)
as.factor(Industry)F	−0.275
	(0.585)
as.factor(Industry)G	0.333
	(0.563)
as.factor(Industry)H	0.147
	(0.690)
as.factor(Industry)I	−0.709
	(0.962)
as.factor(Industry)J	0.133
	(0.542)
as.factor(Industry)K	0.469
	(0.560)
as.factor(Industry)L	−0.283
	(0.692)
as.factor(Industry)M	−0.149
	(0.594)
as.factor(Industry)N	0.848
	(0.655)
as.factor(Industry)Q	−1.432
	(1.195)
as.factor(Industry)R	−0.226
	(0.861)
as.factor(Region)NE	−0.033
	(0.250)
as.factor(Region)SE	−0.081
	(0.263)
as.factor(Region)WE	0.704***
	(0.171)
as.factor(VC_PE_backed)1	0.158
	(0.216)
as.factor(Hot_market)1	−0.021
	(0.151)
Constant	−1.002*
	(0.545)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.26: Results

	Dependent variable:
	I(BHAR_24 >0)
Underpricing	-0.239^{***}
	(0.066)
as.factor(Industry)B	-0.305
	(0.681)
as.factor(Industry)C	0.653
	(0.595)
as.factor(Industry)D	0.126
	(0.700)
as.factor(Industry)F	-0.238
	(0.666)
as.factor(Industry)G	0.517
	(0.628)
as.factor(Industry)H	0.583
	(0.748)
as.factor(Industry)I	-1.023
	(1.214)
as.factor(Industry)J	0.613
	(0.605)
as.factor(Industry)K	0.615
	(0.624)
as.factor(Industry)L	0.935
	(0.726)
as.factor(Industry)M	0.716
	(0.648)
as.factor(Industry)N	0.528
	(0.725)
as.factor(Industry)Q	0.712
	(0.932)
as.factor(Industry)R	0.242
	(0.900)
as.factor(Region)NE	0.549^{**}
	(0.257)
as.factor(Region)SE	0.209
	(0.279)
as.factor(Region)WE	0.861^{***}
	(0.186)
as.factor(VC_PE_backed)1	0.110
	(0.221)
as.factor(Hot_market)1	-0.450^{***}
	(0.156)
Constant	-1.495^{**}
	(0.609)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.27: Results

	Dependent variable:
	I(BHAR_36 >0)
Underpricing	-0.292^{***}
	(0.069)
as.factor(Industry)B	-0.128
	(0.672)
as.factor(Industry)C	0.347
	(0.596)
as.factor(Industry)D	-0.365
	(0.737)
as.factor(Industry)F	-0.058
	(0.655)
as.factor(Industry)G	0.220
	(0.631)
as.factor(Industry)H	-0.019
	(0.782)
as.factor(Industry)I	-0.369
	(1.000)
as.factor(Industry)J	0.422
	(0.606)
as.factor(Industry)K	0.617
	(0.624)
as.factor(Industry)L	1.176
	(0.726)
as.factor(Industry)M	0.416
	(0.653)
as.factor(Industry)N	-0.824
	(0.848)
as.factor(Industry)Q	0.205
	(0.942)
as.factor(Industry)R	0.003
	(0.907)
as.factor(Region)NE	0.619^{**}
	(0.255)
as.factor(Region)SE	-0.233
	(0.299)
as.factor(Region)WE	0.566^{***}
	(0.188)
as.factor(VC_PE_backed)1	0.817^{***}
	(0.219)
as.factor(Hot_market)1	-0.130
	(0.161)
Constant	-1.481^{**}
	(0.611)
Note:	*p<0.1; **p<0.05; ***p<0.01

B.10 BHAR Total

TABLE B.28: Results

	Dependent variable:
	I(BHAR_12 >0)
Age	0.005**
	(0.002)
Ln_Size	−0.018
	(0.028)
ROA	0.062
	(0.051)
Underpricing	−0.096*
	(0.058)
as.factor(Industry)B	−0.004
	(0.597)
as.factor(Industry)C	0.117
	(0.536)
as.factor(Industry)D	0.362
	(0.620)
as.factor(Industry)F	−0.254
	(0.588)
as.factor(Industry)G	0.337
	(0.566)
as.factor(Industry)H	0.063
	(0.703)
as.factor(Industry)I	−0.667
	(0.964)
as.factor(Industry)J	0.158
	(0.545)
as.factor(Industry)K	0.489
	(0.563)
as.factor(Industry)L	−0.323
	(0.697)
as.factor(Industry)M	−0.114
	(0.597)
as.factor(Industry)N	0.885
	(0.659)
as.factor(Industry)Q	−1.439
	(1.198)
as.factor(Industry)R	−0.208
	(0.865)
as.factor(Region)NE	0.031
	(0.253)
as.factor(Region)SE	−0.126
	(0.266)
as.factor(Region)WE	0.757***
	(0.174)
as.factor(VC_PE_backed)1	0.142
	(0.217)
as.factor(Hot_market)1	−0.043
	(0.152)
Constant	−0.941
	(0.601)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.29: Results

	Dependent variable:
	I(BHAR_24 >0)
Age	0.004
	(0.002)
Ln_Size	−0.018
	(0.030)
ROA	0.134**
	(0.054)
Underpricing	−0.240***
	(0.067)
as.factor(Industry)B	−0.192
	(0.686)
as.factor(Industry)C	0.702
	(0.602)
as.factor(Industry)D	0.213
	(0.705)
as.factor(Industry)F	−0.175
	(0.672)
as.factor(Industry)G	0.582
	(0.634)
as.factor(Industry)H	0.583
	(0.762)
as.factor(Industry)I	−0.929
	(1.218)
as.factor(Industry)J	0.688
	(0.610)
as.factor(Industry)K	0.682
	(0.630)
as.factor(Industry)L	0.914
	(0.735)
as.factor(Industry)M	0.827
	(0.654)
as.factor(Industry)N	0.626
	(0.732)
as.factor(Industry)Q	0.684
	(0.944)
as.factor(Industry)R	0.320
	(0.905)
as.factor(Region)NE	0.649**
	(0.260)
as.factor(Region)SE	0.156
	(0.283)
as.factor(Region)WE	0.925***
	(0.189)
as.factor(VC_PE_backed)1	0.091
	(0.223)
as.factor(Hot_market)1	−0.477***
	(0.157)
Constant	−1.498**
	(0.669)
Note:	*p<0.1; **p<0.05; ***p<0.01

TABLE B.30: Results

	Dependent variable:
	I(BHAR_36 >0)
Age	0.006**
	(0.002)
Ln_Size	−0.013
	(0.030)
ROA	0.130**
	(0.056)
Underpricing	−0.291***
	(0.070)
as.factor(Industry)B	−0.026
	(0.678)
as.factor(Industry)C	0.368
	(0.604)
as.factor(Industry)D	−0.276
	(0.742)
as.factor(Industry)F	0.004
	(0.661)
as.factor(Industry)G	0.273
	(0.638)
as.factor(Industry)H	−0.111
	(0.804)
as.factor(Industry)I	−0.266
	(1.006)
as.factor(Industry)J	0.495
	(0.612)
as.factor(Industry)K	0.678
	(0.631)
as.factor(Industry)L	1.175
	(0.736)
as.factor(Industry)M	0.523
	(0.660)
as.factor(Industry)N	−0.736
	(0.854)
as.factor(Industry)Q	0.198
	(0.950)
as.factor(Industry)R	0.088
	(0.911)
as.factor(Region)NE	0.731***
	(0.259)
as.factor(Region)SE	−0.307
	(0.304)
as.factor(Region)WE	0.638***
	(0.191)
as.factor(VC_PE_backed)1	0.800***
	(0.221)
as.factor(Hot_market)1	−0.160
	(0.163)
Constant	−1.548**
	(0.673)
Note:	*p<0.1; **p<0.05; ***p<0.01

Bibliography

Admati, Anant et al. (1998). “Venture capital and the structure of capital markets: banks versus stock markets”. In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/S0304-405X(97)00045-7.

Ak, B. K. et al. (2013). “The use of financial ratio models to help investors predict and interpret significant corporate events”. In: *Australian Journal of Management*. ISSN: 0312-8962. DOI: 10.1177/0312896213510714.

Akaike, Hirotugu (1973). “Maximum likelihood identification of Gaussian autoregressive moving average models”. In: *Biometrika*. ISSN: 0006-3444. DOI: 10.1093/biomet/60.2.255.

Akhigbe, Aigbe, Stephen F Borde, and Ann Marie Whyte (2003). “Does an Industry Effect Exist for Initial Public Offerings?” In: *Financial Review* 38.4, pp. 531–551. ISSN: 1540-6288. DOI: 10.1111/1540-6288.00059. URL: http://dx.doi.org/10.1111/1540-6288.00059.

Al-Khazali, Osamah M. and Taisier A. Zoubi (2005). *Empirical testing of different alternative proxy measures for firm size*. DOI: 10.19030/jabr.v21i3.1471.

Allen, Franklin and Gerald R. Faulhaber (1989). “Signalling by underpricing in the IPO market”. In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/0304-405X(89)90060-3.

AlOmoush, B and F Al-Shubiri (2013). “The impact of multiple approaches financial performance indicators on stocks prices: An empirical study in Jordan”. In: *Journal of Global Business & Economics* 6, pp. 1–11.

Altman, Edward I. (1968). “Financial ratios, Discriminant analysis and The Prediction of corporate bankruptcy”. In: *The Journal of Finance* 23.4, pp. 589–609. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1968.tb00843.x.

— (2002). “Managing Credit Risk : A Challenge for the New Millennium”. In: *Economic Notes* 31.2, pp. 201–214. ISSN: 1468-0300. DOI: 10.1111/1468-0300.00084. URL: http://dx.doi.org/10.1111/1468-0300.00084http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=8&hid=125&sid=082a039b-d550-4a85-8cc9-6ed9ef55e625%40sessionmgr111http://doi.wiley.com/10.1111/1468-0300.00084.

Ang, Andrew and Joseph Chen (2007). “CAPM over the long run: 1926-2001”. In: *Journal of Empirical Finance* 14.1, pp. 1–40. ISSN: 09275398. DOI: 10.1016/j.jempfin.2005.12.001.

Ang, James and Carol Boyer (2009). "Performance differences between IPOs in new industries and IPOs in established industries". In: *Managerial Finance* 35.7, pp. 606–623. ISSN: 0307-4358. DOI: 10.1108/03074350910960346.

B. Chin, Fook Yap, Fie Yong D. Gun, and Ching Poon Wai (2010). "How Well Do Financial Ratios and Multiple Discriminant Analysis Predict Company Failures in Malaysia". In: *International Research Journal of Finance and Economics*. ISSN: 14502887.

Banerjee, Suman, Lili Dai, and Keshab Shrestha (2011). "Cross-country IPOs: What explains differences in underpricing?" In: *Journal of Corporate Finance*. ISSN: 09291199. DOI: 10.1016/j.jcorpfin.2011.06.004.

Baron, David P. (1982). "A Model of the Demand for Investment Banking Advising and Distribution Services for New Issues". In: *The Journal of Finance* 37.4, pp. 955–976. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1982.tb03591.x.

Basti, Eyup, Cemil Kuzey, and Dursun Delen (2015). "Analyzing initial public offerings' short-term performance using decision trees and SVMs". In: *Decision Support Systems*. ISSN: 01679236. DOI: 10.1016/j.dss.2015.02.011.

Beatty, Randolph P. and Jay R. Ritter (1986). "Investment banking, reputation, and the underpricing of initial public offerings". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/0304-405X(86)90055-3.

Beaver, William, Maureen McNichols, and Jung Wu Rhie (2005). "Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy". In: *Review of Accounting Studies* 10.1, pp. 93–122. ISSN: 13806653. DOI: 10.1007/s11142-004-6341-9.

Beaver, William H (1966). "Financial ratios as predictors of failure". In: *Journal of Accounting Research* 4, pp. 71–111. ISSN: 00218456. DOI: 10.2307/2490173. URL: http://www.jstor.org/stable/2490171%5Cnhttp://about.jstor.org/terms.

Bell, R. Greg, Curt B. Moore, and Igor Filatotchev (2012). "Strategic and institutional effects on foreign IPO performance: Examining the impact of country of origin, corporate governance, and host country effects". In: *Journal of Business Venturing* 27.2, pp. 197–216. ISSN: 08839026. DOI: 10.1016/j.jbusvent.2010.11.001.

Benveniste, Lawrence M., Walid Y. Busaba, and William J. Wilhelm (2002). "Information Externalities and the Role of Underwriters in Primary Equity Markets". In: *Journal of Financial Intermediation*. ISSN: 10429573. DOI: 10.1006/jfin.2000.0310.

Bergström et al. (2006). "Underpricing and Long-Run Performance Patterns of European Private-Equity-Backed and Non-Private-Equity-Backed IPOs". In: *The Journal of Private Equity*. ISSN: 1096-5572. DOI: 10.3905/jpe.2006.650457.

Berk, Ales S and Polona Peterle (2015). "Initial and Long-Run IPO Returns in Central and Eastern Europe." In: *Emerging Markets Finance & Trade* 51, S42–S60. ISSN: 1540496X. DOI: 10.1080/

1540496X.2015.1080555. URL: http://10.0.4.56/1540496X.2015.1080555%5Cnhttp://search.ebscohost.com.pxz.iubh.de:8080/login.aspx?direct=true&db=bth&AN=110263031&site=eds-live&scope=site.

Bessler, Wolfgang and Andreas Kurth (2007). "Agency Problems and the Performance of Venture-backed IPOs in Germany: Exit Strategies, Lock-up Periods, and Bank Ownership". In: *The European Journal of Finance*. ISSN: 1351-847X. DOI: 10.1080/13518470600763661.

Black, Fischer and Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities." In: *Journal of Political Economy* 81.3, p. 637. ISSN: 00223808. DOI: 10.1086/260062. URL: http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=5053921&lang=pt-br&site=eds-live&authtype=ip,cookie,uid.

Brau, James C. (2012). "Why Do Firms Go Public?" In: *The Oxford Handbook of Entrepreneurial Finance*. ISBN: 9780199940820. DOI: 10.1093/oxfordhb/9780195391244.013.0016.

Brau, James C. and Stanley E. Fawcett (2006). "Initial public offerings: An analysis of theory and practice". In: *Journal of Finance*. ISSN: 00221082. DOI: 10.1111/j.1540-6261.2006.00840.x.

Brau, James C., Bill Francis, and Ninon Kohers (2003). "The Choice of IPO versus Takeover: Empirical Evidence". In: *The Journal of Business*. ISSN: 0021-9398. DOI: 10.1086/377032.

Brav, Alon, Christopher Geczy, and Paul A. Gompers (2000). "Is the abnormal return following equity issuances anomalous?" In: *Journal of Financial Economics* 56.2, pp. 209–249. ISSN: 0304405X. DOI: 10.1016/S0304-405X(00)00040-4.

Brav, Alon and Paul A. Gompers (1997). "Myth or Reality? The Long-Run Underperformance of Initial Public Offerings: Evidence from Venture and Nonventure Capital-Backed Companies". In: *The Journal of Finance* 52.5, p. 1791. ISSN: 00221082. DOI: 10.2307/2329465. URL: http://www.jstor.org/stable/2329465?origin=crossref.

Brown, Stephen J. and Jerold B. Warner (1980). "Measuring security price performance". In: *Journal of Financial Economics* 8.3, pp. 205–258. ISSN: 0304405X. DOI: 10.1016/S0304-405X(96)00899-9.

Bryman, Alan and Emma Bell (2015). *Business research methods*. 4. ed. Oxford: Oxford University Press, p. 778. ISBN: 9780199668649.

Chemmanur, Thomas J. and Imants Paeglis (2005). "Management quality, certification, and initial public offerings". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2004.10.001.

Claessens, Stijn and Luc Laeven (2003). "Financial Development, Property Rights, and Growth". In: *The Journal of Finance* 58.6, pp. 2401–2436. ISSN: 00221082. DOI: 10.1046/j.1540-6261.2003.00610.x. URL: http://doi.wiley.com/10.1046/j.1540-6261.2003.00610.x.

Clark, David T (2002). "A Study of the Relationship Between Firm Age-at-IPO and Aftermarket Stock Performance". In: *Financial Markets, Institutions & Instruments* 11.4, pp. 385–400. ISSN: 1468-0416. DOI: 10.1111/1468-0416.11406. URL: http://dx.doi.org/10.1111/1468-0416.11406.

Cogliati, Giordano M., Stefano Paleari, and Silvio Vismara (2011). "IPO pricing: Growth rates implied in offer prices". In: *Annals of Finance*. ISSN: 16142446. DOI: 10.1007/s10436-010-0170-6.

Dainow, Joseph (1966). "The Civil Law and the Common Law: Some Points of Comparison". In: *Source: The American Journal of Comparative Law* 15.3, p. 419435. ISSN: 0002919X. DOI: 10.2307/838275.

De Vaus, David (2001). *Research Design in Social Research*. ISBN: 9780761953470. DOI: 10.1177/1473325011401471.

Degeorge, François, François Derrien, and Kent L. Womack (2007). "Analyst hype in IPOs: Explaining the popularity of bookbuilding". In: *Review of Financial Studies*. ISSN: 08939454. DOI: 10.1093/revfin/hhm010.

Demers, Elizabeth and Philip Joos (2007). "IPO failure risk". In: *Journal of Accounting Research* 45.2, pp. 333–371. ISSN: 00218456. DOI: 10.1111/j.1475-679X.2007.00236.x.

Demers, Elizabeth and Katharina Lewellen (2003). "The marketing role of IPOs: evidence from internet stocks". In: *Journal of Financial Economics* 68.3, pp. 413–437. ISSN: 0304405X. DOI: 10.1016/S0304-405X(03)00072-2. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304405X03000722.

Dunbar, Craig G. and Stephen R. Foerster (2008). "Second time lucky? Withdrawn IPOs that return to the market". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2006.08.007.

Engelen, Peter Jan and Marc van Essen (2010). "Underpricing of IPOs: Firm-, issue- and country-specific characteristics". In: *Journal of Banking and Finance*. ISSN: 03784266. DOI: 10.1016/j.jbankfin.2010.01.002.

Espenlaub, Susanne, Alan Gregory, and Ian Tonks (2000). "Re-assessing the long-term underperformance of UK Initial Public Offerings". In: *European Financial Management*. ISSN: 1354-7798. DOI: 10.1111/1468-036X.00127.

Fama, Eugene F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2, p. 383. ISSN: 00221082. DOI: 10.2307/2325486. URL: http://www.jstor.org/stable/2325486?origin=crossref.

— (1998). "Market efficiency, long-term returns, and behavioral finance". In: *Journal of Financial Economics* 49.3, pp. 283–306. ISSN: 0304405X. DOI: 10.1016/S0304-405X(98)00026-9. URL: http://www.sciencedirect.com/science/article/pii/S0304405X98000269.

Fama, Eugene F. and Kenneth R. French (1992). "The Cross-Section of Expected Stock Returns". In: *The Journal of Finance* 47.2, p. 427. ISSN: 00221082. DOI: 10.2307/2329112. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1992.tb04398.x/full%5Cn%3CGotoISI%3E://WOS:A1992HX63800001http://doi.wiley.com/10.1111/j.1540-6261.1992.tb04398.xhttp://www.jstor.org/stable/2329112?origin=crossref.

— (2004a). "New lists: Fundamentals and survival rates". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2003.04.001.

— (2004b). "The Capital Asset Pricing Model: Theory and Evidence". In: *Journal of Economic Perspectives* 18.3, pp. 25–46. ISSN: 0895-3309. DOI: 10.1257/0895330042162430. URL: http.

— (2015). "A five-factor asset pricing model". In: *Journal of Financial Economics* 116.1, pp. 1–22. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2014.10.010.

Fauzi, Fitriya, Nirosha Hewa-Wellalage, and Stuart Locke (2012). "The Global Financial Crisis' Impact on Short-term Performance of IPO: The Case Study of New Zealand Firms' IPOs". In: *Asian Journal of Finance & Accounting* 4.2. ISSN: 1946-052X. DOI: 10.5296/ajfa.v4i2.1829. URL: http://www.macrothink.org/journal/index.php/ajfa/article/view/1829.

Fisher, Irving (1907). *The Rate of Interest*. McMaster University Archive for the History of Economic Thought. URL: http://econpapers.repec.org/RePEc:hay:hetboo:fisher1907.

— (1930). *The Theory of Interest*. New York: The Macmillan Co. URL: http://www.econlib.org/library/YPDBooks/Fisher/fshToI.html.

Freeman, R. Edward, Andrew C. Wicks, and Bidhan Parmar (2004). "Stakeholder Theory and "The Corporate Objective Revisited"". In: *Organization Science*. ISSN: 1047-7039. DOI: 10.1287/orsc.1040.0066.

Friedman, Milton (1975). *There's no such thing as a free lunch*. Lasalle III, p. 330.

Gaunt, Clive (2004). "Size and book to market effects and the Fama French three factor asset pricing model: Evidence from the Australian stockmarket". In: *Accounting and Finance* 44.1, pp. 27–44. ISSN: 08105391. DOI: 10.1111/j.1467-629x.2004.00100.x.

Goergen, Marc, Arif Khurshed, and Ram Mudambi (2007). "The long-run performance of UK IPOs: can it be predicted?" In: *Managerial Finance* 33.6. Ed. by Arif Khurshed, pp. 401–419. ISSN: 0307-4358. DOI: 10.1108/03074350710748759. URL: http://www.emeraldinsight.com/doi/10.1108/03074350710748759.

Grinblatt, Mark and Chuan Yang Hwang (1989). "Signalling and the Pricing of New Issues". In: *The Journal of Finance* 44.2, pp. 393–420. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1989.tb05063.x.

Guo, Hui (2011). "IPO First-Day Return and Ex Ante Equity Premium". In: *Journal of Financial and Quantitative Analysis*. ISSN: 0022-1090. DOI: 10.1017/S002210901100010X.

H. Stock, James and Mark W. Watson (2010). *Introduction to Econometrics (3rd Edition) (Addison-Wesley Series in Economics)*. Vol. 1, p. 840. ISBN: 0138009007.

Helwege, Jean and Nellie Liang (2004). “Initial public offerings in hot and cold markets”. In: *Journal of Financial and Quantitative Analysis*. ISSN: 0022-1090. DOI: 10.1017/S0022109000004026.

Huyghebaert, Nancy, Linda Van De Gucht, and Cynthia Van Hulle (2007). “The choice between bank debt and trace credit in business start-ups”. In: *Small Business Economics*. ISSN: 0921898X. DOI: 10.1007/s11187-006-9005-2.

Ibbotson, R. (1975). “Price performance of common stock new issues”. In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/0304-405X(75)90015-X.

Ibbotson, Roger G. and Jeffrey F. Jaffe (1975). “"Hot Issue" Markets”. In: *The Journal of Finance* 30.4, pp. 1027–1042. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1975.tb01019.x.

Jain, Bharat A. and Omesh Kini (1994). “The Post-Issue Operating Performance of IPO Firms”. In: *The Journal of Finance* 49.5, pp. 1699–1726. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1994.tb04778.x.

Jardin, Philippe du (2009). “Bankruptcy prediction models: How to choose the most relevant variables?” In: *Bankers, Markets & Investors* 98, pp. 39–46. ISSN: 0148-396X. DOI: 10.1227/01.NEU.0000349921.14519.2A.

Jelic, Ranko, Brahim Saadouni, and Mike Wright (2005). *Performance of private to public MBOs: The role of venture capital*. DOI: 10.1111/j.0306-686X.2005.00608.x.

Jenkinson, Tim and Alexander Ljungqvist (2001). *Going public: the theory and evidence on how companies raise equity finance*. Oxford: Oxford University Press, 241 s. ISBN: 0198295995.

Jovanovic, Boyan and Peter L. Rousseau (2001). “Why wait? A century of life before IPO”. In: *American Economic Review* 91.2, pp. 336–341. ISSN: 00028282. DOI: 10.1257/aer.91.2.336.

Kaplan, Robert S (1984). “The Evolution of Management Accounting”. In: *The Accounting Review* 59.3, pp. 390–418. ISSN: 00014826. DOI: 10.2307/246701.

Katz, Sharon P. (2009). “Earnings quality and ownership structure: The role of private equity sponsors”. In: *Accounting Review*. ISSN: 00014826. DOI: 10.2308/accr.2009.84.3.623.

Krishnan, C. N. V. et al. (2011). “Venture Capital Reputation, Post-IPO Performance, and Corporate Governance”. In: *Journal of Financial and Quantitative Analysis*. ISSN: 0022-1090. DOI: 10.1017/S0022109011000251.

La Porta, Rafael et al. (1997). *Trust in Large Organizations*. DOI: 10.1126/science.151.3712.867-a.

La Porta, Rafael et al. (2000). “Investor protection and corporate governance”. In: *Journal of Financial Economics*. ISSN: 00144797. DOI: http://dx.doi.org/10.1016/S0304-405X(00)00065-9.

Lastra, Rosa María (2006). *Legal foundations of international monetary stability*. Oxford: Oxford University Press, p. 547. ISBN: 0199269343.

Levis, Mario (2011). “The Performance of Private Equity-Backed IPOs”. In: *Financial Management* 40.1, pp. 253–277. ISSN: 00463892. DOI: 10.1111/j.1755-053X.2010.01141.x.

Lintner, John (1965). “The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets”. In: *The Review of Economics and Statistics* 47.1, p. 13. ISSN: 00346535. DOI: 10.2307/1924119. URL: http://www.jstor.org/stable/1924119http://www.jstor.org/stable/1924119?origin=crossref.

Ljungqvist, Alexander P, Tim Jenkinson, and William J Wilhelm Jr. (2003). “Global Integration in Primary Equity Markets: The Role of U.S. Banks and U.S. Investors”. In: *Review of Financial Studies*. ISSN: 0893-9454. DOI: 10.1093/rfs/16.1.63.

Loughran, Tim and Jay R. Ritter (1995). “The New Issues Puzzle”. In: *The Journal of Finance* 50.1, pp. 23–51. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1995.tb05166.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1995.tb05166.x.

— (2000). “Uniformly least powerful tests of market efficiency”. In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/S0304-405X(99)00054-9.

Lyandres, Evgeny, Le Sun, and Lu Zhang (2005). “Investment-Based Underperformance Following Seasoned Equity Offerings”. In: *National Bureau of Economic Research Working Paper Series*. DOI: 10.3386/w11459.

Lyon, John D., Brad M. Barber, and Chih-Ling Tsai (1999). “Improved Methods for Tests of Long-Run Abnormal Stock Returns”. In: *The Jounal of Finance* 54.1, pp. 165–201. ISSN: 0022-1082. DOI: 10.1111/0022-1082.00101.

Maksimovic, Vojislav and Pegaret Pichler (2001). “Technological innovation and initial public offerings”. In: *Review of Financial Studies*. ISSN: 08939454. DOI: 10.1093/rfs/14.2.459.

Merton, Robert C (1974). “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”. In: *The Journal of Finance* 29.2, p. 449. ISSN: 00221082. DOI: 10.2307/2978814. URL: papers3://publication/doi/10.2307/2978814?ref=search-gateway:670998d5695f540594a51eb7d23a1a16%5Cnhttp://www.jstor.org/stable/2978814?origin=crossref.

Michaely, R. and W. H. Shaw (1994). “The Pricing of Initial Public Offerings - Tests of Adverse-Selection and Signaling Theories”. In: *Review of Financial Studies*. ISSN: 08939454. DOI: 10.1093/rfs/7.2.279.

Miller, Edward M. (1977). “Risk, Uncertainty and Divergence of opinion”. In: *The Journal of Finance*. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1977.tb03317.x.

Mossin, Jan (1966). “Equilibrium in a Capital Asset Market”. In: *Source: Econometrica Econometrica* 34.4, pp. 768–783. ISSN: 00129682. DOI: 10.2307/1910098. URL: http://www.jstor.org/stable/1910098%5Cnhttp://www.jstor.org/%5Cnhttp://www.jstor.org/action/showPublisher?publisherCode=econosoc.%5Cnhttp://www.jstor.

org%5Cnhttp://www.jstor.org/stable/1910098%5Cnhttp://www.jstor.org/%5Cnhttp://www.jstor.org/action/showPu.

Muscarella, Chris J. and Michael R. Vetsuypens (1989). "The underpricing of "Second" initial public offerings". In: *Journal of Financial Research* 12.3, pp. 183–192. ISSN: 14756803. DOI: 10.1111/j.1475-6803.1989.tb00512.x.

Novy-Marx, Robert (2013). "The other side of value: The gross profitability premium". In: *Journal of Financial Economics* 108.1, pp. 1–28. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2013.01.003.

Ohlson, James a. (1980). "Financial Ratios and the Probabilistic Prediction of Bankruptcy". In: *Journal of Accounting Research* 18.1, p. 109. ISSN: 00218456. DOI: 10.2307/2490395. URL: http://www.jstor.org/stable/10.2307/2490395?origin=crossref%5Cnpapers3://publication/doi/10.2307/2490395%5Cnhttp://www.jstor.org/stable/10.2307/2490395?origin=crossref%5Cnhttp://www.jstor.org/stable/10.2307/2490395?origin=crossref.

Pagano, Marco, Fabio Panetta, and Luigi Zingales (1998). *Why do companies go public? An empirical analysis*. ISBN: 00221082. DOI: 10.1111/0022-1082.25448.

Paleari, Stefano (2014). *Academic EurIPO Fact Book 2014*. Milan: University of Bergamo, p. 121. ISBN: 9781500608309.

Peristiani, Stavros and Gijoon Hong (2004). "Pre-IPO Financial Performance and Aftermarket Survival". In: *Current Issues in Economics and Finance* 10.2, pp. 1–8.

Rajan, Raguram G. (1992). "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt". In: *The Journal of Finance*. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1992.tb04662.x.

Ritter, Jay R (1984). "The Hot Issue Market of 1980". In: *The Journal of Business* 57.2, pp. 215–240. ISSN: 0021-9398. DOI: 10.1086/296260. URL: http://www.jstor.org/stable/2352736%5Cnhttp://www.jstor.org/stable/2352736.

Ritter, Jay R. (1991). "The Long-Run Performance of initial Public Offerings". In: *The Journal of Finance* 46.1, pp. 3–27. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1991.tb03743.x.

Ritter, Jay R. (2003). "Differences between European and American IPO Markets". In: *European Financial Management* 9.4, pp. 421–434. ISSN: 1354-7798. DOI: 10.1111/1468-036X.00230.

Ritter, Jay R (2011). "Equilibrium in the IPO Market". In: *Annual Review of Finance Economics* 3, pp. 1–37. ISSN: 1556-5068. DOI: 10.2139/ssrn.2172332.

Ritter, Jay R., Andrea Signori, and Silvio Vismara (2012). "Economies of Scope and IPO Activity in Europe". In: *SSRN Electronic Journal*, pp. 1–29. ISSN: 1556-5068. DOI: 10.2139/ssrn.2169966. URL: http://www.ssrn.com/abstract=2169966.

Ritter, Jay R. and Ivo Welch (2002). "A Review of IPO Activity, Pricing, and Allocations". In: *The Journal of Finance* 57.4, pp. 1795–1828. ISSN: 00221082. DOI: 10.1111/1540-6261.00478. URL: http://doi.wiley.com/10.1111/1540-6261.00478.

Ruud, Judith S. (1993). "Underwriter price support and the IPO underpricing puzzle". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/0304-405X(93)90015-4.

Savsar, Arif and S S Karaca (2012). "The effect of financial ratios on the firm value: Evidence from Turkey". In: *Journal of Applied Economic Sciences* 1.19, pp. 56–63.

Schmidt, Daniel, Sascha Steffen, and Franziska Szabó (2010). "Exit strategies of buyout investments: An empirical analysis". In: *Journal of Alternative Investments* 12.4, pp. 58–84. ISSN: 15203255. DOI: 10.3905/JAI.2010.12.4.058.

Schultz, Paul (2003). *Pseudo Market Timing and the Long-Run Underperformance of IPOs*. DOI: 10.1111/1540-6261.00535.

Shalit, S and U Sankar (1977). "The Measurement of Firm Size". In: *The Review of Economics and Statistics*.

Sharp, William F. (1964). "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". In: *The Journal of Finance* 19.3, pp. 425–442. ISSN: 00221082. DOI: 10.2307/2329297.

Sherman, Ann E. (2005). "Global trends in IPO methods: Book building versus auctions with endogenous entry". In: *Journal of Financial Economics*. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2004.09.005.

Shiller, Robert J (1990). "Speculative Prices and Popular Models". In: *Journal of Economic Perspectives*. ISSN: 0895-3309. DOI: 10.1257/jep.4.2.55.

Shumway, Tyler (2001). "Forecasting Bankruptcy More Accurately: A Simple Hazard Model". In: *Journal of Business* 74.1, pp. 101–124. ISSN: 0021-9398. DOI: 10.1086/209665.

Skogsvik, Stina and Kenth Skogsvik (2010). "Accounting-based probabilistic prediction of ROE, the residual income valuation model and the assessment of mispricing in the swedish stock market". In: *Abacus* 46.4, pp. 387–418. ISSN: 00013072. DOI: 10.1111/j.1467-6281.2010.00325.x.

Smith, Raymond F. and Arthur H. Winakor (1935). *Changes in the financial structure of unsuccessful industrial corporations*. Urbana: University of Illinois, p. 44.

Stoughton, Neal M., Kit Pong Wong, and Josef Zechner (2001). "IPOs and Product Quality". In: *The Journal of Business*. ISSN: 0021-9398. DOI: 10.1086/321931.

Stringham, Edward Peter and Ivan Chen (2012). "The Alternative of Private Regulation: The London Stock Exchange's Alternative Investment Market as a Model". In: *Economic Affairs*. ISSN: 02650665. DOI: 10.1111/j.1468-0270.2012.02172.x.

Su, Chen and Kenbata Bangassa (2011). "The impact of underwriter reputation on initial returns and long-run performance of Chinese IPOs". In: *Journal of International Financial Markets, Institutions and Money*. ISSN: 10424431. DOI: 10.1016/j.intfin.2011.06.002.

Teoh, Siew Hong, Ivo Welch, and T.J. Wong (1998). "Earnings Management and the Long-Run Market Performance of Initial Public Offerings". In: *The Journal of Finance* 53.6, pp. 1935–1974. ISSN: 00221082. DOI: 10.1111/0022-1082.00079. URL: http://ezproxy.library.capella.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=1311802&site=ehost-live&scope=sitehttp://doi.wiley.com/10.1111/0022-1082.00079.

Titman, Sheridan, K. C. John Wei, and Feixue Xie (2004). "Capital Investments and Stock Returns". In: *Journal of Financial and Quantitative Analysis* 39.4, pp. 677–700. ISSN: 00221090. DOI: 10.1017/S0022109000003173. URL: http://www.jstor.org/stable/30031881.

Treynor, Jack L. (1961). "Market Value, Time, and Risk". URL: http://www.ssrn.com/abstract=2600356.

Welch, Ivo (1989). "Seasoned Offerings, Imitation Costs, and the Underpricing of Initial Public Offerings". In: *The Journal of Finance*. ISSN: 15406261. DOI: 10.1111/j.1540-6261.1989.tb05064.x.

Williams, John Burr. (1938). *The theory of investment value,* English. Cambridge, Mass.: Harvard University Press.

Yung, Chris, Gönül Çolak, and Wei Wang (2008). "Cycles in the IPO market". In: *Journal of Financial Economics* 89.1, pp. 192–208. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2007.06.007.

Zingales, Luigi (1995). "Insider Ownership and the Decision to Go Public". In: *Journal of Financial Economics*. ISSN: 00346527. DOI: 10.1016/0304-405X(94)90035-3.

Zmijewski, Mark E (1984). "Methodological Issues Related to the Estimation of Financial Distress Prediction Models". In: *The Journal of Accounting Research* 221884675.22, pp. 59–82. URL: http://www.jstor.org/stable/2490859%5Cnhttp://www.jstor.org/stable/2490859.