Empirical Essays on Initial Public Offerings

Dissertation to obtain the doctoral degree of Economic Sciences (Dr. oec.)

Faculty of Business, Economics and Social Sciences

University of Hohenheim

Institute of Financial Management

Department of Corporate Finance

submitted by

Annika Reiff

from Reutlingen

2022

Date of Disputation: November 24, 2022

Faculty Dean: Prof. Dr. Jörg Schiller

Supervisor: Prof. Dr. Tereza Tykvová, University of St. Gallen

Second Reviewer: Prof. Dr. Monika Gehde-Trapp, University of Tübingen

Chair of the doctoral committee: Prof. Dr. Andreas Kuckertz, University of Hohenheim

Acknowledgments

I want to thank several people who supported me during the process of writing this thesis. First and foremost, I want to thank my supervisor Professor Dr. Tereza Tykvová for giving me the opportunity to join her team and write this thesis at all. In particular, I am grateful for her guidance and all the fruitful discussions with regard to our joint research project and for her significant contribution to the development of the idea of the third paper of this thesis. Further, I thank my second supervisor Professor Dr. Monika Gehde-Trapp and Professor Dr. Hans-Peter Burghof for joining the advisory team and giving me helpful comments in individual discussions and during presentations at different Ph.D. Finance Seminars in Hohenheim. I also want to thank Professor Dr. Andreas Kuckertz for joining the examination committee. I am very grateful to Dr. Karsten Schweikert for his useful comments and advise regarding the methodological challenges I had to face. I also gratefully acknowledge access to several databases such as Thomson Reuters Securities Data Company (SDC) Platinum, Thomson Reuters Eikon, S&P Capital IQ, Compustat and CRSP provided by the Hohenheim Datalab (DALAHO).

Equally important was the help of my colleagues, fellow researchers and friends Manuel, Sebastian, Robert, Matze and Tanja, who always listened to my problems, offered me advice, and cheered me up in times of doubt and indecision.

I am deeply indebted to my parents and brother for not only supporting me during this work but also always have enabled and encouraged me to pursue my aims.

Finally, I owe the biggest thanks to Lukas for accompanying me through all ups and downs of the working process, offering me endless support, being understanding and patient, and always believing in me. Without him, I wouldn't have had the courage to start this thesis and I would have never finished it.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Going public is one of the most important events in a firm's life cycle. Each year, a great number of firms decide to take this step, thereby raising a large amount of financial resources. In the last two decades, the IPO volume in the US fluctuated between about \$ 19 billion and \$ 75 billion per year, with a peak of even \$ 108 billion in 2000 (see Figure 1.1). Given the substantial amount of money raised in Initial Public Offerings (IPOs), they are of great importance for firms, investors, and the economy as a whole.

Therefore, it is not surprising that IPOs have also attracted the attention of the academic literature and have become the object of a large research field. A fundamental

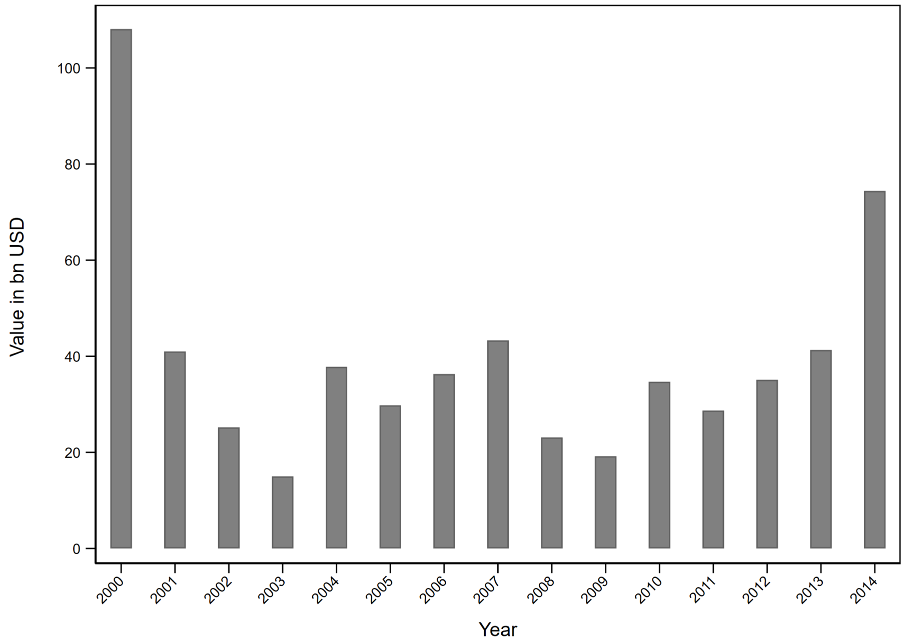


Figure 1.1: IPO volume over time (gross proceeds in billion U.S. dollars)
Source: Statista (2021)

question in the IPO literature is the question of why firms decide to go public at all, which has been analyzed from different theoretical angles. While life cycle theory states that it becomes more profitable for firms to go public than using private equity financing if they have grown sufficiently large (Chemmanur and Fulghieri, 1999), market timing theory states that firms tend to go public in times of market overvaluation (see also Helbing, 2019; Lowry, 2003; Lucas and McDonald, 1990; Ritter and Welch, 2002). Another line of reasoning highlights the possibility of IPOs to facilitate acquisition activities, where the IPO firm might either be target or, more commonly (Brau and Fawcett, 2006), acquirer. In this regard, IPOs are associated with a possibility to maximize proceeds of the sale of the firm (Zingales, 1995), a reduction in valuation uncertainties and an increase in firm value (Celikyurt et al., 2010; Hsieh et al., 2011), the facilitation of cash and especially stock deals (Brau and Fawcett, 2006; Celikyurt et al., 2010), and an attractive exit strategy for early investors (Helbing, 2019). Some firms try to increase their payoffs by following a dual-track strategy by filing for an IPO and pursuing an M&A at the same time (Brau et al., 2010). Another factor that is discussed to influence a firm's decision to go public is the monitoring role of stock markets (Holmström and Tirole, 1993), in which the strict reporting standards of public firms can either reduce agency costs between managers and shareholders (Bancel and Mittoo, 2009; Helbing, 2019; Jensen and Meckling, 1976) or increase the costs of going public (Bancel and Mittoo, 2009; Helbing, 2019). Finally, there are also non-financial motives to go public. These include an improvement in a firm's visibility and reputation, as public firms tend to get more trust by investors and consumers than private firms (Bancel and Mittoo, 2009; Helbing, 2019; Maksimovic and Pichler, 2001; Pagano et al., 1998; Ritter and Welch, 2002). In sum, there are a lot of reasons why firms decide to go public and bear the costs associated with this process, like registration and underwriting fees, costs induced by underpricing or annual disclosing costs (Ritter, 1987; Zingales, 1995). Overall, firms probably do not only have one reason to go public but might decide to go public due to several advantages (Ritter and Welch, 2002).

Apart from the motivation to go public, further questions that are extensively discussed in the IPO literature reach from the analysis of the phenomenon of IPO waves (see Altı, 2005; Altı and Sulaeman, 2012; Benninga et al., 2005; Ibbotson and Jaffe, 1975; Lowry and Schwert, 2002; Lowry, 2003; Ritter, 1984, among others) over IPO underpricing (see Aggarwal et al., 2002; Bradley and Jordan, 2002; Ellul and Pagano, 2006; Loughran and Ritter, 2004; Ljungqvist, 2007; Quintana et al., 2017; Ritter, 1984; Rock, 1986, among others) to the long-run performance of firms after going public (see Aggarwal and Rivoli, 1990; Jain and Kini, 1994; Loughran and Ritter, 1995; Pástor et al., 2009; Ritter, 1991, among others). More recently, IPO research has widened the horizon and started to assess the phenomenon of IPO withdrawals as well as the effect of IPOs on industry rivals.

These two latter strands are in the focus of this dissertation and further illustrated in the following, starting with a short overview over the formal IPO process.

After a firm has decided to pursue an IPO, it has to meet numerous regulatory requirements and has to pass a formal IPO process, which can roughly be summarized in the following steps (see Figure 1.2).[1] The first step is to engage investment bankers (underwriters), auditors and lawyers who prepare the first version of the formal registration statement (Boeh and Dunbar, 2013). The registration statement is called Form S-1 and is examined by the U.S. Securities and Exchange Commission (SEC). It requires to include comprehensive information about the firm's business, its financial statements, management, risk factors, and detailed description of the security (SEC, 2017). While editing the Form S-1, the firm's investment bank sends a preliminary prospectus (also called "Red Herring") to potential investors and conducts road shows in order to attract investor's attention and promote their deal in face-to-face meetings in the second step. After promoting, the investment bank can better estimate the expected demand for the shares (Boeh and Dunbar, 2013). Based on the updated information obtained during this price discovery process, which is also called bookbuilding process[2], a price is set in the third step (Busaba et al., 2001). Busaba et al. (2001) and Busaba (2006) reason that for a successful IPO, investor valuations have to meet the firm's reservation value. If investors and the firm agree on a price, this price becomes the offer price (Busaba et al., 2001). From a legal perspective, the SEC has to declare the S-1 Form as effective and after that, firms are allowed to sell their shares and trading can start in the last step (SEC, 2019). The S-1 filing has to be published on the EDGAR database (Electronic Data Gathering, Analysis and Retrieval system) and thus the information about the firm is publicly available.

An important feature of the IPO process is that firms can withdraw their offering by filing the RW Form (Registration Withdrawal Request) with the SEC at any time during the IPO process (as regulated in rule 477 of the Securities Act of 1933) (Boeh and Dunbar, 2013). In the US, on average more than 25% of the firms that file for an IPO withdraw their registration later on. While the proportion of firms that withdraw fluctuates over

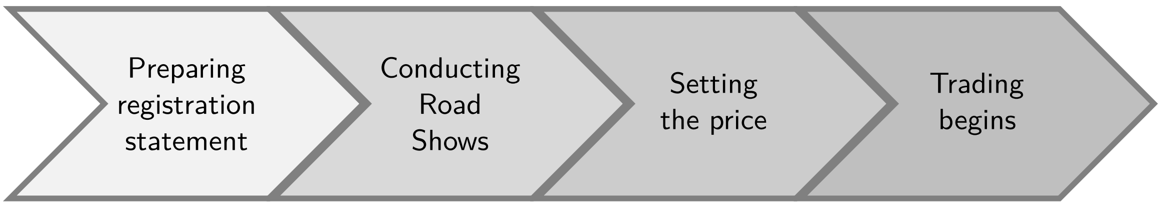


Figure 1.2: IPO process

[1]In the U.S., the rules that regulate the registration for an IPO are mainly defined by the Securities Act of 1933 and the Securities Exchange Act of 1934.

[2]For a theoretical model of the bookbuilding process see Benveniste and Spindt (1989).

time, IPO withdrawal seems to be a permanently existing phenomenon (see Figure 1.3). However, withdrawing an IPO can be very costly as the expenditures related to the formal IPO process are sunk in case of a withdrawal (Boeh and Dunbar, 2013). Therefore, it is rather surprising that such a high share of firms decides to withdraw their offer. In line with the model by Busaba et al. (2001) and Busaba (2006), a firm withdraws its registration if investor valuations fall below the firm's reservation value, which can occur if negative information is revealed during the price discovery process (see e.g. Helbing, 2019).

In the early 2000s, empirical literature on IPO withdrawal has started to explore different determinants that might influence investors' valuations and a firm's reservation value and thus the withdrawal probability (see e.g. Helbing, 2019). As one of the first, Busaba et al. (2001) point to the importance of issue-related characteristics (e.g. filing size), firm characteristics (e.g. firm age, debt ratio) and intermediary characteristics (e.g. Venture Capital (VC) backing). Dunbar and Foerster (2008) add the relevance of market characteristics both at the filing date and their change after the filing date (e.g. BAA-AAA yield spread at the filing date and its change after the filing). More recent studies point to the relevance of corporate governance characteristics (e.g. board size, board experience) (Boeh and Southam, 2011; Helbing et al., 2019). Further studies analyze the role of bank lending standards (Bergbrant et al., 2017), litigation risk (Hao, 2011), and network mechanisms in the context of IPO withdrawals (Owen-Smith et al., 2015) and started to consider post-IPO withdrawal outcomes (Boeh and Dunbar, 2013; Lian and Wang, 2009;

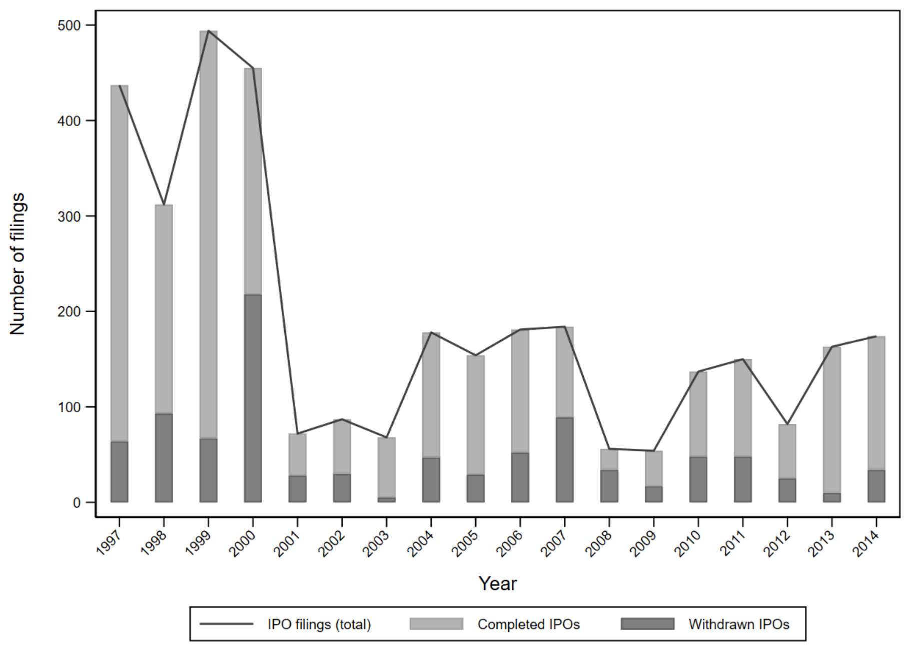


Figure 1.3: Number of completed and withdrawn first-time filings in the US over time
Source: Thomson Reuters Securities Data Company (SDC) Platinum

Lian and Wang, 2012).[3]

While the IPO withdrawal literature has grown over the last two decades, several open questions remain. One factor that is highly discussed is the ambiguous impact of VC backing, with some studies reporting a positive relationship (Boeh and Southam, 2011; Helbing et al., 2019), whereas others report a negative one (Busaba et al., 2001; Dunbar and Foerster, 2008). Similarly, no consensus has been reached with regard to the importance of the different determinants. While Dunbar and Foerster (2008) claim market conditions to be the most important determinants, Helbing et al. (2019) and Helbing (2019) lately call this finding into question and highlight the importance of corporate governance characteristics. At a more general level, the IPO withdrawal literature has been restricted to a backward-looking perspective which assesses the determinants of IPO withdrawal retrospectively. Unlike other areas in IPO literature (e.g. in case of the prediction of IPO underpricing (see Jain and Nag, 1995; Quintana et al., 2017; Reber et al., 2005, among others) or long-run IPO performance (see Jain and Nag, 1998, among others), no attention has been paid to the question whether these determinants can be used to predict future IPO withdrawal. A more forward-looking perspective would be of great relevance for the academic discussion as well as investors.

Apart from the withdrawal decision, another recently developed strand of IPO literature deals with intra-industry effects of IPOs. As outlined above, a lot of information about a firm is revealed during the IPO process. Regarding this high informational content of IPOs, it is likely to assume that there is also some information revealed that is industry-specific and thus an IPO might also influence the IPO firm's industry peers. From a theoretical perspective, this information effect can either be positive or negative. On the one hand, an IPO could signal positive growth prospects for the whole industry which would be in line with positive valuation effects on rival firms (see Akhigbe et al., 2003; Cotei and Farhat, 2013; Lee et al., 2011). On the other hand, an IPO could also reveal that the industry is overvalued (Slovin et al., 1995) or foreshadow future negative industry trends (Spiegel and Tookes, 2020) which would lead to negative valuation effects on industry rivals. However, there might not only be an information effect of IPOs on rivals but also a competition effect. The rationale of the competition effect is that by going public a firm gains some kind of competitive advantage over its rivals. This increases the competitive pressure in the industry which would in turn lead to negative valuation effects (Akhigbe et al., 2003). Empirical findings on the effects of IPOs on rivals are rather mixed. Most studies report overall negative valuation effects. However, some argue that this result is due to a negative information effect (Slovin et al., 1995), while others argue in favor of a negative competition effect (Hsu et al., 2010; McGilvery et al., 2012). Akhigbe et al. (2003) find

[3]For a recent overview over the withdrawal literature see also Helbing (2019).

overall insignificant valuation effects and argue that this is due to the fact that a negative and a positive information effect cancel each other out. Although there is an agreement in the literature that an IPO has an impact on industry rivals, the partly controversial results demonstrate that the direction and underlying processes of the impact have not been solved yet and further research is required on this issue.

The overarching goal of this dissertation is to contribute to the IPO literature by filling some of the research gaps in the strands of IPO withdrawal and intra-industry effects of IPOs. To this end, three independent but theoretically linked empirical studies are pursued. All studies are conducted using US data. The US IPO market is one of the largest IPO markets in terms of IPO volume and number of IPOs. Due to its high relevance, a great amount of research focuses on the US market (see e.g. Helbing, 2019) and so does this thesis.

This thesis is structured as follows. The study presented in chapter 2, **More than just reading tea leaves? Predicting IPO withdrawal using machine learning methods**, adds to the literature by taking a data-driven and forward-looking perspective on IPO withdrawals. It focuses on the prediction of IPO withdrawals, thereby comparing the performance of machine learning methods (lasso and random forest) to regression methods (logit) that are commonly applied in this research area. Results show that random forest outperforms lasso and logit with regard to in-sample prediction and cross-sectional out-of-sample prediction performance. In contrast, all models perform poorly when predicting future IPO withdrawal outcomes based on historical data. I identify concept drift - a change in the relationship between explanatory variables and IPO withdrawal over time - as a key explanation for this puzzling result. Further, the study uses a large set of possible explanatory variables and contributes to the clarification of the question of which variables are most important to predict IPO withdrawal by exploiting certain features of the machine learning methods. All models suggest that market characteristics at filing (like Nasdaq level or bank lending standard) are the most important variables for prediction, while corporate governance and intermediary characteristics seem to be least important. The importance of issuer and issue characteristics ranges between the two other categories.

The study presented in chapter 3, **IPO withdrawals: Are corporate governance and VC characteristics the guiding light in the rough sea of volatile markets?**[4], builds on the first chapter and takes a more theory-based perspective. Considering the high relevance of market characteristics, this study focuses on the question whether certain factors can serve as positive signal for investors and can thus reduce the withdrawal probability even in risky market environments. During the IPO process, there arise agency

[4]This chapter is joint work with Tereza Tykvová from University of St. Gallen

cost, e.g. as investors face issuers of unknown quality (Latham and Braun, 2010) or if the interests of firm insiders and investors diverge (Brav and Gompers, 2003). In these situations, quality signals (Benveniste et al., 2002; Owen-Smith et al., 2015) can possibly alleviate the aforementioned problems by reducing investors uncertainty and thus the withdrawal probability. Further, we argue that signals are more important in risky than in stable market environments, as investors tend to be more careful in these situations. In particular, we focus on the possible signaling effect of corporate governance characteristics (like board experience and board size) and VC backing (see e.g. Boeh and Southam, 2011; Cumming et al., 2016; Helbing et al., 2019). Concerning VC backing, the effect on the withdrawal probability does not necessarily have to be negative from a theoretical point of view. In contrast, the effect could also be positive if VCs follow a dual-track strategy, which would make them less dependent on the IPO success (Dunbar and Foerster, 2008) or if they decide to postpone the IPO in deteriorating market environments (Fan and Yamada, 2020). In order to get deeper insights on the effect of VC backing, this study goes beyond the consideration of average VC effects and additionally distinguishes three VC characteristics: syndicated vs. stand-alone VCs, domestic vs. foreign VCs, and VCs with high vs. low reputation. The results from an interaction term analysis reveal that corporate governance characteristics (especially larger and more experienced boards) seem to serve as positive signals in highly volatile markets, while they do not have a significant effect in more stable market environments. While no evidence can be found for an average effect of VC backing, local VCs and VC syndication tend to reduce the withdrawal probability (particularly in highly volatile markets) which supports the signaling explanation. This is not the case for reputable VCs, for which results suggest that they rather follow the dual-track strategy or postpone the IPO more likely in highly volatile markets than in less volatile markets.

An IPO filing might not only influence the firm that goes public but also its industry rivals. The study presented in chapter 4, **Lucky coincidence or ominous threat: New evidence on the causal mechanisms behind the effect of IPOs on industry rivals**, tries to shed new light on the underlying mechanisms behind the effect of IPOs on rivals. Specifically, it applies a new methodology that allows to separately test for the existence of the competition and information effect. The framework is based on the crucial assumption that the strength of the competition and information effect vary with the starting level of competition and information, respectively. The idea is that the higher the starting level of competition and information in the industry, the lesser the influence of an additional IPO and consequently the less strong the competition and information effect. For example, if the information level in an industry is already high, an additional IPO is likely to bring about less substantial information than an IPO in an industry with a rather low starting level of information. The same reasoning applies for the competition effect. If this condition holds, a change in rival firms' reaction to IPO filings due to an exogenous change in the competition and information level can be considered as evidence for the existence of the respective effects. Based on this assumption, I apply a methodology that consists of two steps: In the first step, I estimate short-term valuation effects of IPO filings on rivals using an event study approach. In the second step, I perform a Difference-in-Difference (DiD) analysis that compares the reaction of rival firms around two exogenous events between a treatment group that is affected by the event and an unaffected control group. In this respect, the exogenous events either only influence the competition level in the industry (import tariff reductions) or the information level in the industry (broker closures and mergers) and therefore allows to consider both effects separately. Results point to the existence of the competition effect, suggesting that IPO filings harm industry rivals to a certain extent. In contrast, results do not provide sufficient evidence for the existence of the information effect. However, this result could also be explained by an offsetting positive and negative information effect and does not necessarily imply that there is no information effect at all.

Finally, chapter 5 concludes with a summary and a discussion of the results and points to directions of potential future IPO research.

Chapter 2

More than just reading tea leaves? Predicting IPO withdrawal using machine learning methods[1]

2.1 Introduction

During the last years, a growing number of literature has drawn the attention to IPO withdrawal and has attributed a high economic significance to this topic (see Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019, among others). IPO withdrawal describes the phenomenon of filing for an IPO but then withdrawing the IPO before listing.[2] This is not an unusual phenomenon, since in the US almost one quarter of firms that file for registration withdraw their IPO later on (Dunbar and Foerster, 2008). Previous literature mainly focuses on the identification of determinants for IPO withdrawal, thereby taking different theories into account. Stated otherwise, they focus on hypothesis testing and take a rather retrospective view.

In recent years, a more data-driven approach has been evolved and successfully applied to different scientific areas including areas in finance (see Barboza et al., 2017; Butler et al., 2014; Chen et al., 2016; Henrique et al., 2019; Tian et al., 2015, among others), which is referred to as machine learning. In this research field, the focus is on prediction rather than hypothesis testing, meaning that the primary goal is to correctly predict the outcome as such. Prediction has so far received less attention in the IPO withdrawal literature but is also of high relevance in two regards. From a practically oriented view, it might be relevant for investors to predict which firms will issue and which firms will withdraw their IPO after filing. By doing so, they are better able to manage their investments. In addition,

[1]This chapter is a single authored manuscript by the candidate.
[2]See e.g. Helbing (2019) for a detailed description of the IPO withdrawal process.

it is also relevant from a scientific point of view. The IPO withdrawal literature reports rather low (in-sample) explanatory power of their models, especially when considering that they include a large set of explanatory variables (e.g. Dunbar and Foerster (2008) report a pseudo R-squared of 0.138 or Helbing et al. (2019) report a pseudo R-squared of 0.275, which corresponds to a rather low prediction accuracy). Therefore, more research is needed to identify approaches which are suited to improve model performance. Furthermore, previous literature gets contradictory findings regarding the importance of the determinants. Firms themselves often claim that changes in market characteristics after the filing are the most important drivers of IPO withdrawal (Boeh and Southam, 2011), which is supported by Dunbar and Foerster (2008) based on US data. In contrast, Helbing et al. (2019) argue that corporate governance characteristics are the most important variables in Europe. However, the models that are applied by these studies do not explicitly enable to identify the most important variables for prediction. Therefore, other approaches are needed in order to clarify this question, which is of high relevance regarding the vast amount of predictor variables considered in this literature.

The first aim of this study is to contribute to the literature by assessing whether machine learning methods can improve the prediction performance compared to parametric models. This includes the assessment of both in-sample prediction performance as well as out-of-sample prediction performance, where the outcome of observations not included in the model (test set) are predicted based on a model trained on a different set of observations (training set). In case of out-of-sample prediction, I further distinguish two different ways to split the sample into training and test set: Firstly, I impose a random spilt where both training and test set contain observations out of the whole sample period (cross-sectional out-of-sample prediction). Secondly, I try to predict future withdrawal outcomes based on a model that is trained on historical data (prediction over time). The second aim is to contribute to the clarification of the question of which variables are most important to predict IPO withdrawal by exploiting different features of the machine learning methods.

In order to pursue these aims, I use a non-missing data set on 2,444 US first-time filings between 1997 and 2014 and a vast selection of explanatory variables. The selection is based on previous IPO withdrawal literature and contains variables of five different categories, namely *issuer and issue characteristics*, *intermediary characteristics*, *corporate governance characteristics*, *market characteristics at filing* and in some additional analyses *changes in market characteristics after filing* (based on Boeh and Southam, 2011; Dunbar and Foerster, 2008; Helbing et al., 2019).

There are two criteria the machine learning method has to fulfill in order to be suitable to analyze the research questions. First, the method has to be able to handle classification problems, as the outcome variable – IPO withdrawal – is binary. Second, the method has to provide some kind of variable selection mechanism or variable importance measure in order to be able to identify the most important variables. I apply two machine learning methods that fulfill these criteria, namely lasso and random forest. Lasso is a rather classical method for variable selection, whereas random forest is more comprehensive and flexible than lasso (e.g. James et al., 2013). The performance of both methods is compared to a logit model in order to analyze whether machine learning methods are able to outperform conventional statistical models.

The results reveal a nuanced picture. Concerning in-sample performance, I find that random forest performs quite well in predicting IPO withdrawal and clearly outperforms lasso and logit. The same pattern can be observed for cross-sectional out-of-sample prediction. However, when trying to predict future IPO withdrawal outcomes based on historical data, the performance of all models is very poor and sometimes even worse than assigning outcomes randomly. I identify the presence of concept drift - a change in the relationship between explanatory variables and IPO withdrawal between training and test set - as one explanation for this puzzling result. Concept drift appears at different points in the sample period and is not limited to single variables. Concerning variable importance, both random forest and lasso suggest variables out of the category *market characteristics at filing* to be the most important for prediction.

The remainder of this paper is organized as follows. The following section reviews the related literature. Section 2.3 describes the data and variables that are used in the analysis. Subsequently, section 2.4 explains the methodologies and measures that are used to compare their performance. Section 2.5 presents results including robustness checks, section 2.6 concludes.

2.2 Related literature

This paper is related to two research areas. From a content-based perspective, it relates to the IPO withdrawal literature which started to emerge in the early 2000s (Helbing, 2019). From a methodological point of view, it is related to the recently growing machine learning literature that is applied to different problems in empirical finance.

Previous IPO withdrawal literature mainly focusses on the explanation of the IPO withdrawal phenomenon from a retrospective view (see e.g. Helbing (2019) for a recent review over the IPO withdrawal literature). The central aim of this strand of literature is to identify determinants that lead to the withdrawal decision. There are different theoretical frameworks that try to shed light on the withdrawal phenomenon. One of the most influential theoretical models is the one by Busaba et al. (2001) and Busaba (2006). In their framework, the price of an IPO is typically determined through a process called bookbuilding, originally modelled by Benveniste and Spindt (1989). The critical point for a successful IPO is that the issuing firm and investors agree on a price. However, if the minimum price demanded by the issuing firm is higher than the maximum price the investors are willing to pay, the IPO is withdrawn. Thus, the withdrawal decision is determined by factors influencing investors' and firm's valuations. In an early empirical analysis on US data, Busaba et al. (2001) find that issue characteristics (e.g. filing size), firm characteristics (e.g. debt ratio), and intermediary characteristics (e.g. whether the firm is backed by venture capital) influence a firm's withdrawal decision. Dunbar and Foerster (2008) extend their analysis by showing that also market characteristics at filing (e.g. the number of filings in the market or the BAA-AAA yield spread) and the change of market characteristics after the filing (e.g. the change in the BAA-AAA yield spread) contribute to the explanation of the IPO withdrawal phenomenon. In a subsequent work, Boeh and Southam (2011) find that also corporate governance characteristics (e.g. board size or board experience) play a role in explaining IPO withdrawal. Lately, Helbing et al. (2019) claim that corporate governance characteristics are even the most important determinants of IPO withdrawal in Europe.

The goal of this strand of literature is the identification of causal relationships between different determinants and IPO withdrawal by means of parametric modelling. A recently growing literature strand focusses stronger on predicting outcomes per se while paying less attention to the actual hypothesis tests (James et al., 2013). This perspective often relies on machine learning methods to improve prediction performance and identify variables that matter most in predicting outcomes, regardless of their theoretical underpinning.

Machine learning methods have also been emerged in different areas in financial research during the last years. One exemplary research area in which these methods have been extensively applied is financial market prediction. This area mainly uses support vector machines and neural networks to forecast financial time series data (for a review see e.g. Henrique et al. (2019)). Another strand of literature applies machine learning methods to predict bankruptcy and credit scoring. From a methodological point of view, this field is closer related to IPO withdrawal as the aim is to predict a binary and often unevenly distributed (imbalanced) outcome (for a review over this strand of literature see e.g. Lin et al. (2011) and Chen et al. (2016)). While these studies report that none of the numerous machine learning methods clearly outperforms the other ones, machine learning methods tend to outperform traditional models substantially (see e.g. Barboza et al. (2017) for a comparison of random forest and logistic regression). In addition to improving prediction performance, machine learning methods have also proven to be a useful instrument to identify the most important variables in bankruptcy prediction (see e.g. Tian et al. (2015)). However, there are also several problems that can arise when applying machine learning methods to real data. Chen et al. (2016) for example point to the problem of handling imbalanced data and the choice of the right measure to compare different classifiers.

In the IPO literature, the use of machine learning methods is still scarce and mainly deals with the prediction of IPO underpricing (see Jain and Nag, 1995; Quintana et al., 2017; Reber et al., 2005, among others) and long-run IPO performance (see e.g. Jain and Nag (1998) for early work). However, early studies on this topic show that machine learning models deliver promising results and also tend to outperform traditional statistical models in this strand of literature. The same holds true for the application of machine learning methods to identify the most important variables for prediction (with regard to underpricing see e.g. Butler et al. (2014) for the US and Bastı et al. (2015) for Turkey). To the best of my knowledge, only Esfahanipour et al. (2016) mention the prediction of IPO withdrawal in a sub-analysis of their study on the effect IPO withdrawal on underpricing. Using data on the Teheran stock exchange, they predict IPO withdrawal using neural networks and conclude that their model performs quite well. However, they use a rather small selection of variables for prediction and don't provide a sufficient discussion of their results with regard to IPO withdrawal. In addition, results on the Teheran stock exchange probably differ to an analysis using US or European data.

To sum up, the IPO withdrawal literature has so far mainly focused on estimating causal relationships using conventional statistical models. At the same time, there is a growing literature that have shown that machine learning methods can outperform conventional models with regard to prediction performance in related areas of financial research. This paper contributes to the literature by bringing both sides together. Firstly, this paper applies different machine learning methods to predict IPO withdrawal and compares their performance to a logit model. Secondly, it tries to identify the most important variables for prediction.

2.3 Data and variables

The sample contains information about US filings between 1997 and 2014. All filings in the considered time period were drawn from Thomson Reuters Eikon, thereby excluding American depositary receipts, convertible issues, unit offerings, closed-end funds, REITs, limited partnerships, small best effort offers, SPACs, issuers that are not seeking a listing on NYSE, NASDAQ or any other American exchange and companies from the financial industry (see Busaba et al., 2001; Dunbar and Foerster, 2008; Ritter and Welch, 2002, among others). Following previous IPO literature, the sample was reduced to first-time IPOs, because features of first-time IPOs might differ from the circumstances under which firms return for a second time (see Boeh and Dunbar, 2013; Chen et al., 2017; Dunbar and Foerster, 2008, among others). This leads to a sample containing 4,939 first-time filings. As filing dates, IPO dates and withdrawal dates are not always correctly reported in Thomson Reuters Eikon, all these dates were checked against EDGAR filings and only those firms were kept for which dates correspond to each other. This validity check reduced the sample to 3,438 filings. A vast selection of variables from different sources is merged to this data. This leads to a reduction in observations due to missing values that are asymmetrically distributed over all variables. When keeping only firms for which all information is available, the final non-missing data set consists of 2,444 observations.[3]

Table 2.1 shows how these observations are distributed over time. The average withdrawal probability over the whole sample period is 22.30%, indicating that the data is rather imbalanced. This means that the proportion of one class (majority class) is larger than the other one (minority class) (Bekkar et al., 2013). As discussed later on, this has to be taken into account when comparing model performance. In addition, the withdrawal rate changes considerably over time (with a minimum withdrawal rate of 6.03% in 2013 and a maximum withdrawal rate of 58.97% in 2008). This fact should be kept in mind when predicting IPO withdrawal over time.

The outcome is a binary variable that equals one if the IPO is withdrawn and zero otherwise. In order to identify the most important variables for prediction, this paper uses a vast selection of explanatory variables that have been included in previous studies that try to explain IPO withdrawal (see Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019, among others). As prediction is the main focus of this paper, the theoretical background and the direction of the effects of the variables that are included in the analysis are not discussed here. Stated otherwise, the focus of this analysis is to test the predictive power of the variables that show a significant influence in explaining IPO withdrawal in previous research.

Table A.1 provides an overview over all variables and describes their sources. The last column lists previous studies on IPO withdrawal, which find a relationship between the respective variables and the withdrawal probability. Following Dunbar and Foerster (2008) the variables are divided into different categories.[4] Table 2.2 displays summary statistics of all variables. The first category *issuer and issue characteristics* comprises filing size, firm size measured by total assets, firm age, a dummy variable indicating whether the firm belongs to a high-tech industry, debt ratio, the book-to-market ratio, asset turnover, calculated as the ratio of total revenues to total assets and a negative news dummy, indicating whether negative business news have been reported about the firm one

[3]To stay as close as possible at the IPO withdrawal literature, I decided to forgo all possible methods of missing value imputation.

[4]Note that their original classification include four classes: Issuer and issue characteristics, Investment bank characteristics, market conditions at time of filing and market conditions after the filing.

Table 2.1: Withdrawn and completed IPOs by year

Filing Year	Number of filings	Number completed	Number withdrawn	Percentage completed	Percentage withdrawn
1997	297	276	21	92.93	7.07
1998	214	158	56	73.83	26.17
1999	376	339	37	90.16	9.84
2000	322	190	132	59.01	40.99
2001	51	37	14	72.55	27.45
2002	63	45	18	71.43	28.57
2003	54	50	4	92.59	7.41
2004	134	106	28	79.10	20.90
2005	117	102	15	87.18	12.82
2006	121	93	28	76.86	23.14
2007	134	72	62	53.73	46.27
2008	39	16	23	41.03	58.97
2009	39	28	11	71.79	28.21
2010	94	65	29	69.15	30.85
2011	106	73	33	68.87	31.13
2012	50	38	12	76.00	24.00
2013	116	109	7	93.97	6.03
2014	117	102	15	87.18	12.82
Total	2,444	1,899	545	77.70	22.30

Note: Number of filings, completed issues and withdrawn issues from 1997-2014.

year prior to the filing (see Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019). The second category *intermediary characteristics* includes a dummy variable indicating whether the firm is backed by venture capital, the number of underwriters involved in the issue and the Carter-Manaster (CM) Rank of the lead underwriter (see Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019). The third category, *corporate governance characteristics*, is added to the classification by Dunbar and Foerster (2008) and was suggested by Boeh and Southam (2011) and Helbing et al. (2019). In this category, I include board size, board experience measured as the average age of all board members, the ratio of female board members, a dummy variable indicating whether the role of the CEO and the role of the chairman are hold by the same person (CEO duality) and finally a dummy variable that indicates whether the lock up period reported in EDGAR filings is larger than 180 days (see Boeh and Southam, 2011; Busaba et al., 2001; Helbing et al., 2019). The fourth category *market characteristics at filing* adds the Nasdaq composite index at the time of filing, a measure for the expected volatility over the next 30 days, the BAA-AAA yield spread at filing (difference between the Moody's BAA-rated corporate bonds and the

AAA-rated bonds) and the change in bank lending standard over the past three month before the filing date (see Bergbrant et al., 2017; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019).

One big difference between prediction and identifying causal relationships in a retrospective way is that in the former case only variables can be included that are known at the time of the prediction. For the main analysis, I predict IPO withdrawal at the time of the filing date. However, as firms themselves often claim that the change in market conditions after the filing are the most important determinant for their withdrawal decision (Boeh and Dunbar, 2013), I perform an additional analysis by predicting IPO withdrawal 30 days and 60 days after the filing. In these analyses, *changes in market characteristics after filing* are added as fifth category. In particular, I include daily Nasdaq returns, the standard deviation of Nasdaq returns to capture the actual volatility and the change in BAA-AAA yield spread. These analyses aim to provide deeper insights into the question whether changes in market characteristics after filing are able to improve prediction performance and how important these variables are. I choose the periods of 30 days and 60 days after filing because on the one hand the period has to be sufficiently long to reflect changes in market characteristics after the filing date (see Dunbar and Foerster, 2008).[5] On the other hand, the period should be kept as short as possible because it is desirable to predict IPO withdrawal as early as possible and the number of firms that have already reached an outcome at the time of prediction and thus have to be excluded from the analysis should be kept at a minimum. When predicting 30 days after filing, only 1.06% of observations are lost. When predicting 60 days after filing, this number already increases to 13.22%.

The fact that variables are summarized to five categories might rise the concern that variables are correlated and thus cause multicollinearity issues. Table A.2 in the appendix depicts Pearson correlation coefficients between all variables. Overall, only very few variables reach a correlation higher than 0.5 suggesting that multicollinearity concerns can be refuted. The only exceptions are variables measured 30 and 60 days after filing which are only included in the supplementary analyses. In addition to the correlation matrix, I test for multicollinearity by calculating the variance inflation factor (VIF) after an OLS regression. The mean VIF for the variables included in the prediction at the filing date is 1.46, while it is 1,59 for the variables included in the prediction 30 days after filing and 1.72 for the variables included in the prediction 60 days after filing. All VIF values are far below the critical value of 10, at which correlations might cause trouble (Wooldridge, 2009) and therefore there is no need to account for correlated variables in the following.

[5]They find a significant effect of the change in market characteristics between filing and two month after filing on the withdrawal probability.

Table 2.2: Descriptive statistics

	obs	mean	std. dev.	min	max
Issuer and issue characteristics					
Filing size	2,444	0.12	0.17	0.00	4.20
Firm size	2,444	0.48	1.67	0.00	13.86
Firm age	2,444	13.62	19.00	0.00	100.00
High-tech	2,444	0.60	0.49	0.00	1.00
Debt ratio	2,444	0.82	0.78	0.04	6.30
BM ratio	2,444	0.34	0.20	0.07	1.24
Asset turnover	2,444	1.07	1.13	0.00	6.16
Negative news	2,444	0.29	0.45	0.00	1.00
Intermediary characteristics					
VC backing	2,444	0.41	0.49	0.00	1.00
Number of underwriters	2,444	1.60	1.18	1.00	13.00
CM Rank	2,444	7.87	1.78	0.00	9.00
Corporate governance characteristics					
Board size	2,444	6.62	2.21	2.00	14.00
Board experience	2,444	50.60	5.94	36.75	64.62
Female board members	2,444	0.04	0.09	0.00	0.67
CEO duality	2,444	0.53	0.50	0.00	1.00
Lock up	2,444	0.03	0.17	0.00	1.00
Market characteristics at filing					
Nasdaq	2,444	21.48	9.73	7.84	47.05
Volatility (filing date)	2,444	21.71	6.70	9.06	64.24
Yield spread	2,444	0.82	0.23	0.50	3.43
Number of filings	2,444	0.15	0.10	0.00	0.45
Bank lending standard	2,444	2.02	17.56	-24.10	83.60
Market characteristics at/after filing (30 days)					
Volatility (30 days)	2,418	22.28	7.61	8.79	92.98
Nasdaq return (30 days)	2,418	0.05	0.39	-1.76	1.15
Nasdaq std. dev. (30 days)	2,418	0.69	0.71	0.10	4.86
Yield spread change (30 days)	2,418	0.01	0.10	-0.29	0.79
Market characteristics at/after filing (60 days)					
Volatility (60 days)	2,121	22.66	8.14	9.21	92.98
Nasdaq return (60 days)	2,121	0.05	0.28	-1.47	0.92
Nasdaq std. dev. (60 days)	2,121	1.02	0.99	0.15	4.62
Yield spread change (60 days)	2,121	0.03	0.16	-0.42	0.96

Note: Descriptive statistics of predictor variables. Variable descriptions and data sources are described in Table A.1.

2.4 Methodology

This section presents different methodologies that are applied to predict IPO withdrawal. In order to be suitable for the purpose of this study, the method has to fulfill two criteria: It should be suitable for classification problems and has to be able to identify the most important variables for prediction. The first machine learning method that fulfills these criteria is lasso. Lasso can be considered as a classical method for variable selection. However, the application of lasso can also increase prediction accuracy and it has been proved to be stable in the presence of imbalanced data (Tian et al., 2015). The second applied machine learning method is random forest, which is more flexible and comprehensive than lasso, as it avoids functional form assumptions and implicitly allows for complex interactions. Its advantage over other comprehensive methods that also perform well in classification problems, like support vector machine or neural networks, is that it also provides variable importance measures. In contrast, other models are often a black-box and their scope for identifying important variables is very limited (see Archer and Kimes, 2008). Finally, the performances of random forest and lasso are compared to a logit model, which is a statistical model that is commonly applied in this research area.[6]

2.4.1 Lasso model

Lasso stands for *least absolute shrinkage and selection operator* and was introduced by Tibshirani (1996). Originally, it is based on a linear regression model, but it is also applicable for logistic regressions. This version is applied here and is also referred to as *l_1 penalized logistic regression* (Friedman et al., 2009). Lasso performs variable selection by shrinking coefficient estimates towards zero, thereby shrinking some coefficients exactly to zero. Thus, the lasso model only contains a subset of variables. On the one hand, this shrinking procedure selects the important variables. On the other hand, it can lead to a reduction in variance and thus increase prediction accuracy (Tibshirani, 1996). Formally, the maximization problem can be depicted as:

$$\max_{\beta_0,\beta} \left\{ \sum_{i=1}^{N} \left[y_i(\beta_0 + \beta^T x_i) - log(1 + e^{\beta_0 + \beta^T x_i}) \right] - \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{2.1}$$

(Friedman et al., 2009, p.125). The important part in equation (3) is the second term, which is the penalty that shrinks the coefficient estimates towards zero.[7] In this regard, λ can be considered as tuning parameter. The larger λ, the stronger the penalty and

[6] As the logit model is better comparable to lasso, logit is used for comparison rather than probit.

[7] For a more detailed discussion and derivation of regularized logistic regression see e.g. Friedman et al. (2010).

thus the stronger the shrinking of the coefficients towards zero. When increasing λ, that is, when further shrinking coefficients towards zero, one variable after the other drops out of the model. The order in which the variables drop out of the model, which is referred to as lasso solution path, can be used to create a ranking of the importance of the variables. Most commonly, the optimal level of λ that minimizes the prediction error (λ_{min}) is determined by using cross-validation.[8] Alternatively, some studies report the λ that delivers the simplest model but also lies within one standard error of the optimal level of λ (λ_{1se}). However, it should be noted that the prediction accuracy is always the highest when using λ_{min} (James et al., 2013) and thus this will be the preferred model when comparing performances of the different models.

Zou (2006) claims that it is possible that the lasso model that maximizes the prediction performance might lead to an inconsistent variable selection. He introduces adaptive lasso, a version of lasso in which coefficients are weighted. He shows that this weighting procedure leads to consistent variable selection. In order to test whether the applied lasso model performs consistent variable selection, results are compared to an adaptive lasso model as a robustness check.

2.4.2 Random forest

Random forest is more flexible and complex than lasso and it focuses stronger on prediction than variable selection. Random forest is based on decision trees and was first introduced by Ho (1995) and then further developed by Breiman (2001). The idea of decision trees is to split the observations into regions that contain observations that are quite similar to each other with regard to covariates. Within these regions, each observation is assigned to the most commonly occurring class (James et al., 2013). For example, in the analysis of this paper, observations would be assigned to the two classes *withdrawal* or *issue*. However, decision trees are very non-robust and small changes in the data can lead to different results. The idea of random forest is to grow an ensemble of trees on bootstrapped training samples with replacement which increases the robustness compared to single trees. After drawing a sufficiently large number of bootstrapped samples and growing a tree in each one, the trees are averaged and majority votes are taken for all observations. The distinction of random forest compared to other decision tree methods, especially bagging, is that here only a subset of all predictors is allowed at each split (m). This procedure leads to a decorrelation of the trees and a reduction of variance. The reason is that by only allowing for a few predictors at each split, the trees differ stronger from each other (James et al., 2013).[9] The choice of the number of predictors that are allowed

[8] For a description of cross-validation see James et al. (2013).

[9] For a more technical description of random forests see Breiman (2001).

at each split (m) plays an important role for the performance of the model and can be considered as tuning parameter. Typically used values for this parameter (m) are the square root of the number of predictors, twice this number and half this number (see e.g. Liaw, 2002). Another important tuning parameter is the number of grown trees ($ntree$), which needs to be sufficiently high in order to get a stable model. As random forests cannot overfit (Breiman, 2001) a large number of trees should not cause any problems beyond additional computation time. The different models with the varying tuning parameters can be compared by their out-of-bag error (OOB). The OOB reflects the percent of incorrectly predicted observations from the out-of-bag sample, which contains observations that are not used for growing the tree in the bootstrap sample (Breiman, 2001; James et al., 2013).

The variable importance measure for random forest has also already been introduced by Breiman (2001). There are two main approaches to calculate variable importance: permutation importance (Breiman, 2001) and Gini importance (Breiman and Cutler, 2003). Intuitively, the permutation importance calculates the mean decrease in accuracy attributed to each variable by calculating the difference in accuracies of a model that includes the true values of this variable and a model that randomly permutes the values of the variable in the OOB sample (Breiman, 2001). In contrast, the Gini importance measure calculates the improvement in the Gini splitting index that is attributed to each variable (James et al., 2013). Strobl et al. (2007) argue that permutation importance is the preferred measure when including variables of different types (as in this analysis) as the Gini importance measure tends to rank continuous variables higher than categorical variables. I check the classical permutation importance against the AUC based permutation importance, as Janitza et al. (2013) shows that it performs better if data is imbalanced. This measure simply replaces the accuracy by the area under the ROC curve (AUC) to evaluate the decrease in performance attributed to each variable.

2.4.3 Performance measures

There exists an extensive number of measures to compare the performance of different models. The probably most commonly used measure is accuracy which depicts the percentage of correctly classified observations out of all observations (Ferri et al., 2009). However, accuracy can be very misleading in the presence of imbalanced data. For example, if one class contains 90% of observations, a high accuracy of 90% will already be achieved if predicting the same outcome for all observations (Bekkar et al., 2013).

Although no consensus has been achieved in the literature about the usage of the right performance measures in presence of imbalanced data (see Bekkar et al., 2013; Ferri et al., 2009; Huang and Ling, 2005, among others) there are some measures that are able to moderate this problem. Often applied and considered as the preferred measure in the following analysis is the AUC measure, which is a quantitative measure to evaluate the Receiver Operating Curve (ROC) curve. The ROC curve plots the true positive rate against the false positive rate for different threshold values (Fawcett, 2006). The independence of a threshold value is a property that makes the ROC curve and thus the AUC especially useful in the presence of imbalanced data (Bradley, 1997). In addition, the AUC measure puts the same weight on both classes, whereas other measures like accuracy often favor the majority class (Janitza et al., 2013). An AUC value of one, for which the area under the ROC curve is largest, would indicate perfect prediction. In contrast, an AUC value of 0.5, for which the ROC curve corresponds to the bisector angle, would be obtained under random prediction (Hanley and McNeil, 1982). As a rough rule of thumb, an AUC value between 0.5 and 0.6 can be considered as poor, a value between 0.6 and 0.7 as fair, a value between 0.7 and 0.8 as good, a value between 0.8 and 0.9 as very good and a value between 0.9 and 1 as excellent (see e.g. Bekkar et al., 2013).

The second applied measure is Cohen's Kappa (Cohen, 1960). Originally, it measures the agreement of two classifiers (Ferri et al., 2009). More intuitively, it measures how much better the classifier performs than a classifier that randomly assigns observations to classes according to the class distributions. Formally:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \tag{2.2}$$

Here, $P(A)$ is the relative observed agreement among classifiers, which corresponds to the accuracy in this case. $P(E)$ is the agreement obtained by random assignment (Ferri et al., 2009). By comparing the classifier to random assignment, the problem is moderated that accuracy might be overvalued due to imbalances in data. As a rough benchmark, a Kappa smaller than zero indicates poor agreement, values between 0.00 and 0.20 slight agreement, values between 0.21 and 0.40 fair agreement, values between 0.41 and 0.60 moderate agreement, values between 0.61 and 0.80 substantial agreement and values between 0.81 and 1.00 almost perfect agreement (Landis and Koch, 1977).

The third measure that is considered to compare the model performances is the F_1Score. It is the harmonic mean of precision and recall. In presence of imbalanced data, it might be better to use a weighted average of both measures than only considering one of them, as the single measures rather provide information about single classes (Chawla, 2009). Formally, the F_1Score is:

$$F_1score = \frac{2 * recall * precision}{precision + recall} \tag{2.3}$$

Here, precision is the ratio of observations correctly classified as positives (True Positives) and the overall number of observations predicted as positives (True Positives + False Positives). Recall is defined as the fraction of the observations correctly classified as positives (True Positives) and the total number of positives (True Positive + False Negative) (Ferri et al., 2009). The $F_1 Score$ also ranges between zero and one.

2.5 Results

2.5.1 In-sample prediction

This section presents the results of in-sample prediction and variable importance measures based on the whole sample, beginning with random forest, followed by lasso and logit.

Random forest

Table 2.3 shows the OOB error for different random forest models with varying number of trees (*ntree*) and different numbers of variables allowed at each split (m). Following previous literature, I apply the square root of the number of predictors (4), twice this number (8) and half of this number (2) for parameter m (see e.g. Liaw, 2002). The OOB error does not change much when changing the number of variables allowed at each split (m). Graph A.1 in the appendix shows a plot of the OOB error against the number of trees. It shows that the OOB decreases substantially when only slightly increasing the number of trees but then soon reaches a constant low level at about 2,000. This result is also supported by the OOB errors depicted in Table 2.3, as the OOB error does not change much for 2,500, 5,000 or 10,000 trees. As there are no big differences when varying both tuning parameters, I choose the model which allows for four variables at each split and builds 5,000 number of trees as baseline model, as both numbers seem to be sufficiently high and the model remains computationally feasible for these parameters.

Table 2.4, column (1) reports different performance measures for the baseline random forest model. These measures are calculated based on the whole sample (in-sample

Table 2.3: Tuning the random forest models

	Ntree=2,500	Ntree=5,000	Ntree=10,000
m=2	21.28%	21.28%	21.24%
m=4 (default)	20.74%	21.03%	20.91%
m=8	20.79%	20.87%	20.87%

Note: Performance of different random forest models with a varying number of trees (ntree) and a varying number of variables used at each split (m). OOB errors are reported.

prediction performance). The AUC measure of 0.74 indicates good in-sample performance, while the other measures also point to the result that random forest is able to predict IPO withdrawal quite well.

Figure 2.1 depicts the variable importance ranking delivered by the permutation importance measure based on the mean decrease in AUC. Four out of the five most important variables indicated by the ranking are out of the category *market characteristics at filing*, namely the Nasdaq level at filing date, the measure for bank lending standard, number of filings and expected volatility. With yield spread being at rank six, market characteristics seem to be the most important variables for (in-sample) prediction. Variables that seem to be least important in predicting IPO withdrawal are variables out of the categories *corporate governance characteristics* and *intermediary characteristics*. This result remains robust when measuring the permutation importance by mean decrease in accuracy instead of mean decrease in AUC, as displayed in Figure A.3 in the appendix.

Considering the high importance of market characteristics at filing, the question arises whether adding variables after filing further improves the prediction performance. In order to analyze this question, performance measures and variable importance measures are reported for the prediction 30 and 60 days after filing, respectively. Table A.3, column (1) in the appendix shows the measures for in-sample prediction. Observations that already have an outcome 30 or 60 days after filing are excluded from the analysis. For a better comparison, the same settings are applied as in the baseline model. Results show that adding market characteristics after filing neither improves prediction at 30 days after filing nor prediction at 60 days after filing.[10] Figure A.4 and Figure A.5 in the appendix depict variable importance measures for the prediction 30 and 60 days after filing. Some

Table 2.4: Performance of in-sample prediction

	Random forest	Lasso ($\lambda.min$)	Lasso ($\lambda.1se$)	Adaptive lasso	Logit	Logit (manually re-stricted)
AUC	0.7408	0.6443	0.6218	0.6492	0.6521	0.6442
Cohen's Kappa	0.1873	0.0182	0.0000	0.0360	0.0373	0.0510
F1 Score	0.2651	0.0253	-	0.0497	0.0561	0.0730
Accuracy	0.7913	0.7791	0.7770	0.7811	0.7799	0.7819

Note: Performance of in-sample prediction of the models reported in Table 2.5 and the preferred random forest model. The F1 Score of the lasso model with λ_{1se} cannot be calculated here, as this model predicts one class (zero) for all outcomes. For the performance measures that depend on a threshold, a threshold of 0.5 is applied.

[10]Note that the slight decrease in the measures lies within ranges of fluctuation that can be due to technical issues of the method.

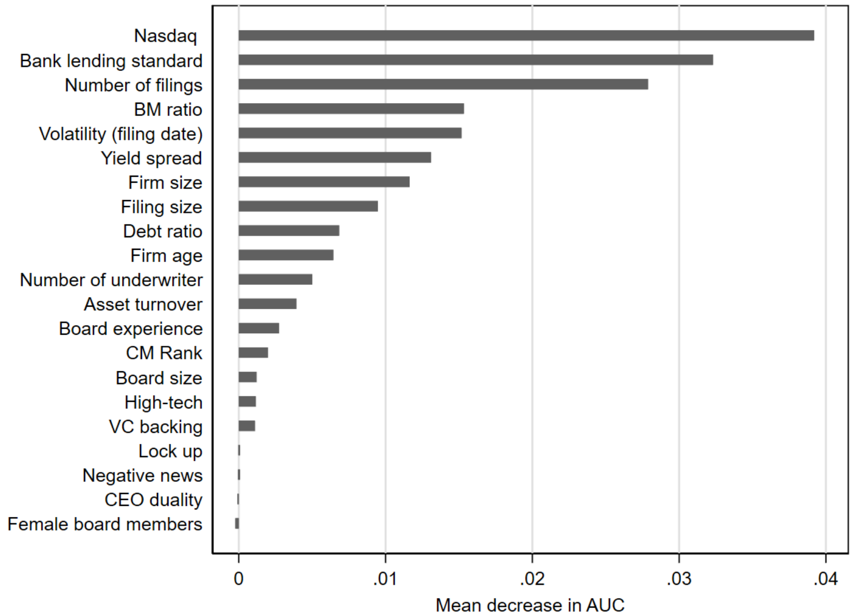


Figure 2.1: Variable importance of random forest based on AUC

of the added market characteristics after filing belong to the more important variables for both points in time. One explanation for this rather contradictory finding might be that the added variables absorb predictive power of the variables included in the baseline specification (e.g. the Nasdaq level at filing date is highly correlated with Nasdaq st. dev.).

Lasso and logit

Random forest results are compared to results obtained by lasso and logit. Table 2.5 reports the coefficients of the baseline lasso models in columns (1) and (2). The optimal level of lambda is obtained by 10 fold-cross-validation. Figure A.2 plots the binomial deviance against different values of the log of lambda. The left dotted line indicates the optimal level of lambda for which the error is smallest (λ_{min}). Here, λ_{min} is equal to 0.0069 (model depicted in column (1)). The right dotted line displays λ_{1se}, the value of lambda which delivers the simplest model but also lies within one standard error of the optimal level of λ, which corresponds to 0.0231 in the case at hand (model depicted in column (2)). Column (3) shows the results of the adaptive lasso model (with the optimal level of λ). The baseline lasso model in column (1) shrinks 9 out of the 21 variables exactly to zero. Most of these variables belong to the *intermediary* and *corporate governance characteristics* (VC backing, CM Rank, Board size, board experience, female board members), while view variables belong to the *issuer and issue characteristics* (filing size, firm size, high-tech, negative news). The simpler lasso model in column (2) additionally shrinks the two remaining variables out of the category *corporate governance characteristics* to zero (CEO duality, lock up). The adaptive lasso model in column (3) has a lower optimal level of lambda and thus shrinks only four variables to zero. Three out of the four variables are also out of the category *corporate governance characteristics*.

In order to get deeper insights about the importance of each variable, Figure 2.2 shows the lasso solution path for the first model of Table 2.5. The variable that drops out of the model first when increasing lambda is ranked at the bottom. The variable that stays in the model the longest and can thus be considered as most important is ranked at the top. The results are quite similar to the results obtained by the variable importance measure by random forest. Again, *market characteristics at filing* seem to be most important to predict IPO withdrawal. *Intermediary characteristics* drop out of the model first and can thus be considered to be least important, while *corporate governance characteristics* are ranked above the aforementioned category. This is a slight difference to the results obtained by random forest variable importance, in which *corporate governance characteristics* seem to be least important. Figure A.6 in the appendix shows the lasso solution path for the adaptive lasso model, which delivers quite the same picture. Thus, variable selection seems to be robust over the different lasso models.

The baseline logit results are depicted in column (4) in Table 2.5. While lasso performs variable selection by shrinking some coefficients to zero, logit does not provide a comparable variable selection procedure. One way to select important variables in logit would be to follow some kind of predefined statistical criterion, such as statistical significant thresholds. In this context, the crucial question is whether this manual procedure offers different

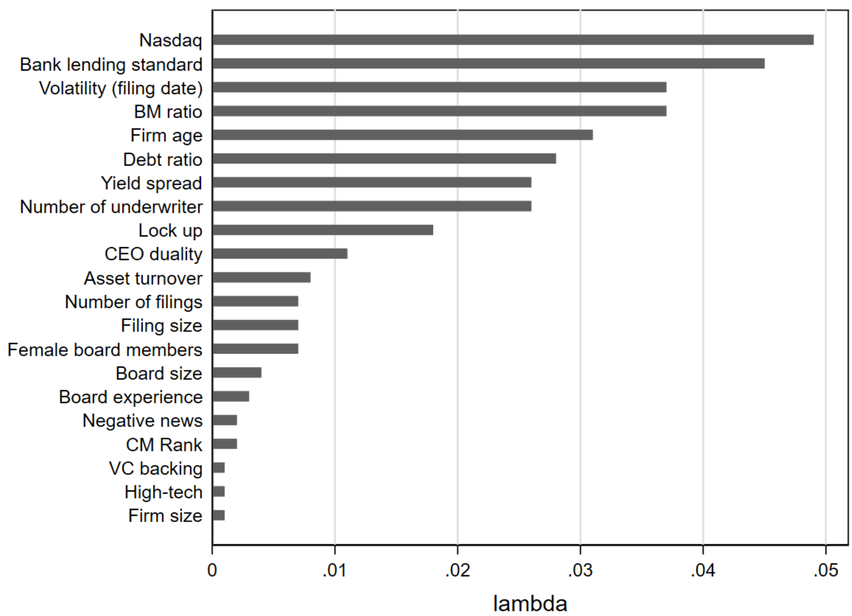


Figure 2.2: Lasso solution path

Table 2.5: Results of lasso and logit

	Lasso		Adaptive lasso	Logit	Logit (manually restricted)
	(1)	(2)	(3)	(4)	(5)
	lambda.min= 0.0069	lambda.1se= 0.0231	lambda.min= 0.00261		
Filing size	0.0000	0.0000	-0.3662	-0.2924	
				[0.4292]	
Firm size	0.0000	0.0000	-0.0099	0.0109	
				[0.0409]	
Firm age	-0.0067	-0.0016	-0.0071	-0.0098***	-0.0094***
				[0.0036]	[0.0032]
High-tech	0.0000	0.0000	0.0000	-0.0195	
				[0.1336]	
Debt ratio	0.1379	0.0240	0.1534	0.1698***	0.1891***
				[0.0604]	[0.0578]
BM ratio	1.0353	0.4077	1.2773	1.3117***	1.3109***
				[0.3060]	[0.2606]
Asset turnover	0.0026	0.0000	0.0446	0.0482	
				[0.0466]	
Negative news	0.0000	0.0000	0.0121	0.0217	
				[0.1133]	
VC backing	0.0000	0.0000	0.0060	0.0061	
				[0.1167]	
Number of underwriter	-0.1534	-0.0116	-0.1362	-0.2505***	-0.2390***
				[0.0677]	[0.0598]
CM Rank	0.0000	0.0000	-0.0132	-0.0083	
				[0.0330]	
Board size	0.0000	0.0000	0.0000	0.0140	
				[0.0245]	
Board experience	0.0000	0.0000	0.0000	-0.0042	
				[0.0093]	
Female board members	0.0000	0.0000	0.0000	-0.5667	
				[0.5951]	
CEO duality	-0.0293	0.0000	-0.1099	-0.1121	
				[0.1024]	
Lock up	-0.4363	0.0000	-0.8570	-0.8997**	-0.8510**
				[0.3997]	[0.3830]
Nasdaq	0.0276	0.0149	0.0354	0.0388***	0.0337***
				[0.0064]	[0.0052]
Volatility (filing date)	0.0186	0.0092	0.0288	0.0300***	0.0240***
				[0.0098]	[0.0079]
Yield spread	0.4612	0.0248	0.3379	0.5094*	0.7376***
				[0.2938]	[0.2388]
Number of filings	0.0000	0.0000	-0.7729	-11545	
				[0.7161]	
Bank lending standard	0.0023	0.0034	0.0038	0.0010	
				[0.0036]	
Constant	-2.7808	-1.9227	-2.9343	-2.8384***	-3.2445***
				[0.6458]	[0.2990]
Likelihood-ratio test on joint significance on excluded variables					
Chi2					7.04
P-value					0.9003

Note: Results of Lasso, adaptive lasso and logit. In specifications (4) and (5) standard errors are reported in parentheses and statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***. The model in specification (5) reports coefficients of a logit model in which the variables that are insignificant in the logit model in specification (4) are excluded. A likelihood-ratio test on joint significance on the excluded variables is reported to test whether the explanatory power of the unrestriced model (4) is better than the one of the restricted model (5). Variable descriptions and data sources are described in Table A.1.

results to the variable selection method by lasso. Column (5) reports coefficients of a logit model in which the variables that are insignificant at the 10 percent level in the baseline logit model are excluded. A likelihood-ratio test on joint significance shows that the excluded variables are not jointly significant, suggesting that the exclusion of the variables does not lead to a significant change in the predictive power. Indeed, variables that are shrunken to zero in the lasso model are insignificant in the logit model. Thus, selecting significant variables manually would have led to the same set of selected variables in this particular case. One explanation might be that variables in the logit model are either highly significant or not significant at all, which makes it relatively easy to pick the important ones. However, in contrast to lasso, there is no reasonable way in logit to draw up a ranking list of the important variables.

The corresponding performance measures of the models are depicted in Table 2.4. There is no big difference between the performance of the different lasso and logit models and their overall performance is only moderate. While the AUC value for the lasso models range between 0.62 and 0.65, the logit model reaches an AUC of 0.65. Cohen's Kappa also indicates only weak prediction performance of all lasso and logit models. Thus, lasso does not outperform logit with regard to in-sample prediction, neither the full logit model nor the logit model that excludes insignificant variables.

When comparing the performance of lasso and logit to random forest, the AUC value of random forest is approximately 10 percentage points higher than the one of the other two models and thus clearly outperforms them. The superiority of random forest over the two other methods in in-sample prediction is further illustrated by the plot of the three ROC curves in Figure 2.3. The ROC curve of random forest lies considerably more to the upper left than the ROC curves of the other models.

In sum, the results from in-sample prediction show that random forest is able to predict IPO withdrawal quite well, while lasso and logit performance is at most moderate. Thus, random forest clearly outperforms lasso and logit. Lasso can neither outperform logit in prediction nor does it chose different variables in variable selection. Both random forest variable importance and lasso solution path suggest that variables out of the category *market characteristics at filing* to be most important for prediction, while variables out of the category *corporate governance characteristics* and *intermediary characteristics* seem to be least important. It should be noted that, in general, it does not make much sense to consider variable importance if overall prediction is poor, as in this case obviously all variables will fail in prediction (Strobl et al., 2008). Thus, variable importance of random forest is preferred over the variable importance ranking obtained by lasso in this analysis.

2.5.2 Out-of-sample prediction

After assessing in-sample performance, it is important to evaluate out-of-sample performance in order to test the external validity of the model. This procedure is characteristic of machine learning methods and generally not performed in classical statistical hypothesis testing. For the purpose of out-of-sample prediction, a certain proportion of the overall sample is set aside as test set, then the model is trained on the remaining training set and after that tested against the test set. For the data structure at hand, there are two different ways to test out-of-sample performance. Firstly, the data set can be split randomly in a cross-sectional way, where both training and test set contain observations out of the whole sample period. This is relevant from a scientific point of view, because it delivers insights on how good the models perform on an unseen test set which is drawn from the same overall data set with same variable distributions as the training set. The second way is to predict future outcomes with a model trained on historical data (prediction over time). This might be more relevant for practitioners, as investors probably want to know whether an IPO will be withdrawn in the future.

When predicting future outcomes based on historical data, I exclude one year between training and test set. The rationale behind the exclusion is that the training set should only include firms, for which an outcome is already observable. To illustrate this, consider an example in which you want to test the performance of a model trained on data comprising the years 1997 till 2005 on future observations (e.g. 2006-2014). At the end of 2005 (at the time of the prediction), there are some firms that have already filed for an IPO but whose outcome will only be observable in 2006 and can thus not be included in the training set. However, simply excluding those firms from the training set might bias the results as the time between filing and outcome is on average higher for withdrawn issues than for completed issues. The average duration of the days between filing and outcome of the firms in the sample is approximately 156 days, while 90% of firms have an outcome after one year. Thus, excluding one year between training and test set should account for the aforementioned problem. In the baseline model, I split training and test set in the proportion of 70:30. For the prediction over time, a 70:30 split approximately corresponds to including observations from 1997 till 2005 in the training set and observations from 2007 to 2014 in the test set. Later on, I test the robustness of the results for different splitting ratios.

Table 2.6 depicts the results of both cross-sectional out-of-sample prediction and prediction over time for the baseline models. Performance measures from cross-sectional out-of-sample prediction show a quite similar pattern like measures of in-sample-performance. The AUC of the random forest model is 0.73, while it is 0.62 for the optimal lasso model and 0.61 for the logit model. The superiority of the random forest model is confirmed by the other performance measures, as Cohen's Kappa is much higher for random forest (0.18) than for lasso (0.00) and logit (0.02). The same is true for the F_1Score (0.24 for random forest, 0.01 for lasso, 0.05 for logit). It also reveals that lasso and logit perform hardly better than random prediction. Thus, again random forest clearly outperforms lasso and logit, while the performance of lasso and logit is quite similar and rather moderate. This is again also illustrated in the plots of the ROC curves in Figure 2.4, where the ROC curve of random forest lies more to the upper left than the ones of the two other models.

These findings change when trying to predict future outcomes based on historical data. The performance measures of all models drop considerably. The AUC of the random forest model decreases from 0.73 in cross-sectional out-of-sample prediction to 0.42, while Cohen's Kappa decreases from 0.18 to 0.01 and F_1Score from 0.24 to 0.01. Although the performance of lasso and logit is already low in cross-sectional out-of-sample prediction, their performance decrease even further. The AUC measures of both models drop to approximately 0.41, while their Cohen's Kappa even becomes negative. With an AUC value below 0.5 and Cohen's Kappa being negative, all models perform even worse than predicting outcomes randomly. Column (5) of Table 2.6 adds year dummies of the dotcom bubble (1999/2000) and the financial crisis (2008/2009) to the logit model in order to check whether controlling for crisis years could moderate this problem but the performance remains still poor.

Table A.4 in the appendix shows results for prediction at 30 days and 60 days after filing, respectively. However, including variables which account for the change in market characteristics after filing do not change the observed pattern. The AUC measures of all models drop below or close to 0.5, while Cohen's Kappa becomes nearly zero or negative when predicting future IPO withdrawal outcomes based on past data. Regarding cross-

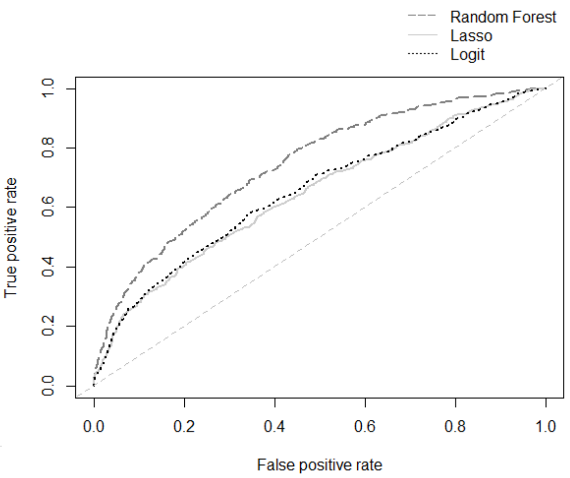


Figure 2.3: ROC curves of in-sample prediction

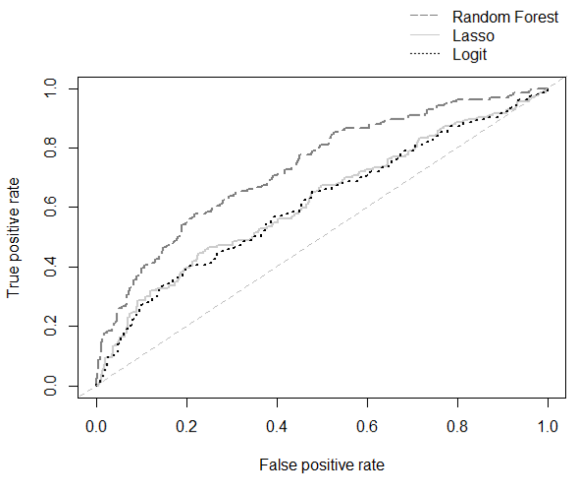


Figure 2.4: ROC curves of out-of-sample prediction

Table 2.6: Performance of out-of-sample prediction

	Random forest	Lasso $\lambda.min$	Logit	Logit (including crisis dummy)
Cross-sectional prediction (Split 70:30) *Withdrawal probability training set: 22.34%* *Withdrawal probability test set: 22.21%*				
AUC	0.7343	0.6194	0.6124	
Cohen's Kappa	0.1850	0.0014	0.0191	
F1 Score	0.2447	0.0124	0.0476	
Accuracy	0.7992	0.7751	0.7737	
Prediction over time (Split 70:30) *Training set: 1997-2005; withdrawal probability: 19.96%* *Test set: 2007 - 2014; withdrawal probability: 27.67%*				
AUC	0.4235	0.4190	0.4060	0.3962
Cohen's Kappa	0.0075	-0.0226	-0.1329	-0.1744
F1 Score	0.0104	-	0.0308	0.1827
Accuracy	0.7252	0.7122	0.6374	0.4978

Note: Performance of out-of-sample prediction for both the cross-sectional split into training and test set and for prediction over time. Splitting proportion is (approx.) 70:30 in both cases. Note that the F1 Score of the lasso model cannot be calculated here, as precision and recall are both zero in this prediction. For the performance measures that depend on a threshold, a threshold of 0.5 is applied.

sectional out-of-sample performance, findings from the in-sample prediction also hold for the prediction 30 and 60 days after the filing. Random forest performs quite well (AUC equals 0.72 in both models), while lasso and logit predict IPO withdrawal only moderately (AUC equals 0.64 for prediction 30 days after filing and 0.66 for prediction 60 days after filing).

Obtaining performances that are even worse than random prediction is quite puzzling and deserves closer consideration. This holds true in particular as the cross-sectional split has revealed that especially random forest seems to be suitable to predict IPO withdrawal even for out-of-sample cases. Thus, there has to be some mechanism that explains why prediction performance deteriorates when IPO withdrawal is predicted for future observations based on past outcomes. One possible and intuitive explanation here is the presence of dataset shift[11], in time series literature extensively discussed under the terminology structural change (see Aue and Horváth, 2013; Perron, 2006, among others).

[11] As terminologies often differ in literature, I base the terminology on the one by Moreno-Torres et al. (2012) but adapt it for the time dimension.

Data set shift describes the problem of the violation of the common assumption that training and test set follow the same distribution. It can occur in multiple forms and affects prediction in different ways. As I consider it in the case of prediction over time, I discuss this problem with regard to changes in distributions between two points in time (t_0 and t_1). The three most common forms are the following. The first one considers changes in the distribution of the input variables, also referred to as covariate shift ($P_{t0}(Y) \neq P_{t1}(Y)$, but $P_{t0}(y|X) = P_{t1}(y|X)$). The second one deals with changes in the distribution of the outcome, also referred to as probability shift ($P_{t0}(Y) \neq P_{t1}(Y)$, but $P_{t0}(X|y) = P_{t1}(X|y)$), e.g. different probabilities of withdrawal between training and test data. The third one deals with changes in the relationship between input and outcome variables, also referred to as concept drift or concept shift ($P_{t0}(X, y) \neq P_{t1}(X, y)$) (Moreno-Torres et al., 2012; Gama et al., 2014).

The one that immediately catches the eye in the context of IPO withdrawal is probability shift, as the withdrawal probability increases from 19.96% in the training set to 27.65% in the test set when predicting future withdrawal outcomes based on historical data. The fact that the withdrawal probability partly changes considerably over time is also displayed in Table 2.1. Thus, the change in the distribution in the outcome variable is always an issue when predicting over time. However, this problem only occurs, if the changes in distribution in the outcome variable cannot be explained by changes in the input variables.

Analyzing the presence of concept drift is probably the most important one, as the negative consequences for prediction over time are most severe in this case and it is most difficult to handle (Moreno-Torres et al., 2012). Intuitively, prediction has to fail if e.g. a highly influential variable has a positive relationship with the outcome variable in the training data but a negative one in the test data. Thus, in the case at hand, concept drift would cause bad prediction performance if a variable increases the probability of IPO withdrawal in the training set but decreases the probability of withdrawal in the test set, as the algorithm would erroneously predict IPO withdrawal more often for high values of the respective variable in the test set.

Testing and correcting for concept drift has mostly been applied in the context of high-frequent stream data (for recent work see e.g. Almeida et al. (2018), Gama et al. (2014), Gao et al. (2007), Lobo et al. (2018), and Yan (2020), for early work on sequential probability tests see Page (1954) and Wald (1945)). As the data structure at hand is very different to the one in these studies, the tests developed in the literature are not applicable for the purpose of this study. Therefore, I suggest a formal regression-based test on concept drift that tests for systematic differences in the relation between the input variables and IPO withdrawal between training and test set. As the test refers to changes in the distribution between explanatory variables and the outcome, its finding should generalize to all methods applied (even though the test itself is based on logit). The idea for the test is as follows. In the first step, I perform an interaction effects model for the whole sample in which all input variables are interacted with a variable that indicates whether the observation belongs to the training or test set. In the second step, I perform a test for joint significance on the interaction effects. It should be noted that joint significance of the interaction effects rather than the one of the single interaction terms matters. If the number of included interactions is sufficiently high, it is possible to get statistically significant results for some variables due to random statistical fluctuations. I perform this test for the time split, for which the interaction dummy equals one for the time period between 2007 and 2014 and zero for the period between 1997 and 2005, as well as for the random split (proportion 70:30). Table A.5 in the appendix shows that the test of joint significance for the time split (column (1) and (2)) is highly significant ($\chi^2 = 141.60$, $p - value < 0.0000$), indicating that the relationship between input and output variables indeed differs between the two samples. In contrast, the test is (as expected) insignificant for the random split (column (3) and (4)). These findings could explain why the different models (especially random forest) perform quite well in the cross-sectional prediction but fail when predicting over time. As interaction effects in a logit model can lead to misleading results (Ai and Norton, 2003), I repeat the tests in a linear probability model. Table A.6 in the appendix confirms the results obtained by the logit model.

In order to illustrate the problem of concept drift further, I plot the predicted probabilities of three variables that proved to be important (for in-sample prediction) in lasso and random forest, namely Nasdaq, bank lending standard and BM ratio, for both the training and test set and the time split and random split in Figures A.7 - A.9. The plot of Nasdaq most clearly illustrates the problem of concept drift. The relationship between Nasdaq and the outcome variable is exactly reversed in the training and test set for the time split. While the probability of withdrawal strongly increases with higher values of the Nasdaq in the training set, the opposite is true in the test set. Consequently, the trained algorithm will predict IPO withdrawal in the test set more often with higher values of the Nasdaq, while the probability actually decreases. This pattern explains why prediction performance is even worse compared to random prediction. The same applies for lending standard and BM ratio though to a lesser extent. In contrast, when splitting the sample randomly into training and test set, the relationship between each of the three variables and the outcome variable is quite similar in both training and test set. In order to assess whether the problem remains in the random forest model, I additionally show partial dependence plots for the three variables (time and random split) in Figures A.10 - A.12. The plots reveal that the problem of concept drift does not seem to vanish in the random forest model.

This finding raises the question whether the prediction performance can be improved by further analyzing the reasons for concept drift. For example, if the concept drift occurs at specific points in time, it might be possible to restrict the training set to periods in which the predictors have a rather constant effect on the withdrawal probability (as often applied in steam data algorithms, see e.g. Almeida et al. (2018), Gama et al. (2014), Gao et al. (2007), Lobo et al. (2018), and Yan (2020)). As the formal tests for structural breaks from the time series literature (e.g. Chow test, CUSUM test (see Muthuramu and Maheswari (2019) for a recent review over tests for structural breaks in time series)) are not applicable in this context, I inspect the temporal variation and development of the effects by means of year-by-year interactions. In particular, I interact the three most important variables (for in-sample prediction) with each year in order to estimate the effect on the withdrawal probability in this respective year. I exemplarily chose the three most important variables as they have the greatest influence on the prediction performance and thus cannot be easily excluded from the model to moderate the problem of concept drift. Figure A.13 displays the marginal effects of Nasdaq, BM ratio and bank lending standard for each year. First, it becomes apparent that there is no obvious or single point at which the effect of the variables changes. For example, the effect of Nasdaq is (sometimes significantly) positive between 1997 and 2001, turns negative in 2002 and 2003, gets strongly positive between 2004 and 2008 and then changes repeatedly from strongly positive to negative from 2009 on. Second, the drift patterns appear to differ between variables. While the effect of Nasdaq seems to be at least consistently positive between 1997 and 2000, the effect of lending standard and BM ratio turns from negative to strongly and significantly positive. It should be noted that the year-by-year interactions have to rely on a very small number of observations in some years. Moreover, the procedure applied here is no formal test in the statistical sense of the word. Nevertheless, these analyses show that concept drift does not seem to be an issue that can be fixed easily by limiting the training set to certain years or that only occurs in extraordinary circumstances. In addition, it is apparent that there is no single variable that entirely causes the concept drift and it is present for variables that proved to be important for prediction, so excluding a variable cannot moderate the problem.

2.5.3 Robustness checks

Table A.7 in the appendix reports some robustness checks of the results by using different ratios to split the sample into training and test set. It displays the out-of-sample AUC, which is the preferred performance measure, as well as the results of the joint significance test in order to formally test whether the concept drift remains for different time splits (as suggested by the analysis in the last section).

I choose the splitting ratios in the following ways. First, just the splitting ratio is varied from a 70:30 split to a 60:40 split and a 75:25 split in order to check whether the results are robust to slight increases and decreases of the training and test set. Overall, the results are robust to these variations. Random forest outperforms lasso and logit in all cross-sectional out-of-sample predictions and all tests for concept drift are insignificant in these models. For the time splits, the AUC measures for all three models are very poor and most of the time the performance remains worse than results obtained from random prediction. Similarly, the tests for joint significance are highly significant for all splitting ratios leading to the conclusion that concept drift is an issue when predicting over time, regardless of the splitting ratio.

When predicting future outcomes based on historical data, analysts have the possibility to consciously choose their training set. Therefore, some additional variations are performed for the time split. First, I check whether results change when varying the prediction horizon. Possibly, outcomes that are closer to the training set can be predicted more accurately. For this purpose, I choose a training set of 1997 to 2008 (to get a sufficiently large set and to avoid biases due to short-term fluctuations) and vary the test set from 2010 to 2012 (short horizon) to 2010 to 2014 (long horizon). Again, there are no big differences. The performance of all models remain very poor, although not worse than random prediction in the short horizon and the test of joint significance on the interaction effects still indicates concept drift. However, although staying overall significant, this test slightly loses significance when reducing the prediction horizon. Thus, reducing the prediction horizon leads to slight improvements but remains far away from delivering good prediction results. As a last test, I exclude crisis years (the dotcom bubble (1999/2000) and the financial crisis (2008/2009) years) from the training set in order to check whether eliminating extreme years improves the prediction performance.[12] Results clearly show that excluding crisis years does not help to improve the prediction performance. This finding is in line with the visual analysis of the year-by-year interactions reported in the previous section, which suggests that different variables drift at different points in time and not all variables solely drift in crisis years.

In sum, results remain robust when varying split proportions either randomly or according to rational considerations. In case of cross-sectional out-of-sample performance, random forest clearly outperforms lasso and logit. When predicting over time, all models perform very poor. One reason for this bad performance is the presence of concept drift, which seems to occur at different points in time and remains for different splitting proportions in the time period covered by the data set.

[12]Note that 2009 is already excluded from the model to avoid selection bias as described above.

2.6 Conclusion

This paper contributes to the IPO withdrawal literature by taking a step away from hypothesis testing and points to a more forward-looking and data-driven perspective by applying different machine learning methods to predict IPO withdrawal. In addition, it sheds new light on the discussion about the importance of the determinants for this prediction. Since machine learning methods have never been applied in the context of IPO withdrawal, it is relevant from a scientific point of view to compare the model performances to a conventional statistical method. In general, research based on conventional statistical models just consider in-sample performance. Results show that random forest in-sample performance is much better than logit in-sample performance, while lasso and logit in-sample performance is quite similar. Variables that seem to be most relevant when predicting IPO withdrawal are variables out of the category *market characteristics at filing*. This finding is consistent between machine learning and regression methods.

In addition to in-sample performance, I compare the model performances on an unseen test set that is drawn randomly from the whole sample in order to test their external validity. Once again, random forest clearly outperforms lasso and logit. In a third analysis, I try to predict future IPO withdrawal based on historical data, which might be more relevant for investors from a more practically oriented view. Surprisingly, all models perform very poorly when predicting IPO withdrawal over time. I identify concept drift as key explanation for this bad performance, which seems to occur at different points in time and is not limited to single variables.

This paper only provides a start of applying machine learning methods in the context of IPO withdrawal. Obviously, there are a lot of questions left open to future research. First of all, regarding its relevance for investors, there is a need for handling concept drift when predicting future IPO withdrawal outcomes. Among others, my analysis has revealed that the strength and direction of possible determinants (e.g. Nasdaq or further market characteristics) sometimes differs between time periods. From a theoretical point of view, this suggests that there are unobserved underlying interaction effects, which deserve closer consideration. Further assessing these interactions would contribute to the IPO literature both at the methodological (in terms of better prediction performance) as well as the theoretical level. Moreover, methodological research might look for possibilities to adapt algorithms from the forecasting literature to data structures given in the IPO withdrawal literature.

Secondly, there are issues regarding the data that deserve consideration. It was mentioned that prediction becomes challenging when data is imbalanced. This problem becomes even more severe with regard to the fact that withdrawal probabilities change over time. I tried to account for imbalanced data by choosing appropriate performance measures when comparing the models. However, there are other methods that can be applied to handle imbalanced data like under-sampling or over-sampling methods that are extensively discussed in the literature (see Barua et al., 2012; Batista et al., 2004; Chawla et al., 2002; Chawla, 2009; Japkowicz, 2001; Lin et al., 2017; Weiss and Provost, 2003; Yen and Lee, 2009, among others). It would also be interesting to investigate whether these methods can be applied in the context of IPO withdrawal prediction and whether results change if doing so. With regard to the large amount of asymmetrically missing values and as this causes a substantial reduction in sample size, it would make sense to check the scope of different imputation methods for missing values.

Thirdly, limitations concerning the external validity deserve some consideration. Although the data comprises a large time horizon, results are limited to this horizon. Related to that, the data used for this analysis are limited to the US context. Consequently, it would be interesting to test whether results differ when using e.g. European or Asian data or data from different time spans. This accounts to both the findings concerning prediction performance or variable importance, as well as problems related to prediction over time.

Lastly, for the scope of this paper, it is left open whether IPO withdrawal is actually good or bad for firms and investors. Brau et al. (2010) e.g. claim that firms often follow a dual-track strategy, that is, they file for an IPO and pursuit an M&A at the same time. In this case, firms and investors might be even better off when withdrawing their IPO and pursuing the M&A (see e.g. Aktas et al., 2018). In addition, there are first works that highlight the importance of the analysis of post-IPO withdrawal outcomes (see Boeh and Dunbar, 2013; Boeh and Dunbar, 2016; Cooney et al., 2009; Hao, 2011; Lian and Wang, 2012, among others). With regard to the possibility that they might be better off with an alternative strategy, it might be also important for investors to know what actually happens to firms after they withdraw their IPO. Taking this into account, it would be interesting to predict post-IPO withdrawal outcomes rather than IPO withdrawal in a next step.

A Appendix

A.1 Tables

Table A.1: Variable descriptions and data sources

Variable	Description	Source	Variable taken into account by e.g.*
Dependent variable			
Withdrawal	Dummy variable that equals one if the filing is withdrawn and zero if the filing is completed.	Thomson Reuters Eikon	
Independent variables			
Issuer and issue characteristics			
Filing size (bn USD)	Filing size in bn USD. Calculated as the product of original global shares filed and original mid filing price.	Thomson Reuters Eikon	Busaba et al. (2001), Dunbar and Foerster (2008), Helbing et al. (2019)
Firm size (bn USD)	Total assets in bn USD in most recent financial period before filing.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	Boeh and Southam (2011), Busaba et al. (2001), Helbing et al. (2019)
Firm age	Firm age at filing date in years (filing date – founding date)	Founding and filing date primarily taken from Thomson Reuters Eikon. Supplemented with and checked against data stemming from Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`) and S&P Capital IQ.	Boeh and Southam (2011), Helbing et al. (2019)

High-tech	Dummy indicating high tech industry. Equals one if primary industry of the firm is high-tech industry according to the definition in Thomson Reuters Eikon and zero otherwise.	Thomson Reuters Eikon	Boeh and Southam (2011), Busaba et al. (2001), Dunbar and Foerster (2008), Helbing et al. (2019)
Debt ratio	Ratio of total liabilities to total assets in most recent financial period before filing.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	Boeh and Southam (2011), Busaba et al. (2001), Helbing et al. (2019)
BM ratio	Book-to-market ratio for firms in the issuer's Fama-French industry (48) in the year before filing.	Kenneth R. French's Data Library (Kenneth R. French's website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`).	Dunbar and Foerster (2008)
Asset turnover	Ratio of total revenues to total assets in most recent financial period before filing.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	Boeh and Southam (2011), Busaba et al. (2001)
Negative news	Dummy that equals one if the LexisNexis Negative News Search delivers that negative business news were reported about the firm one year prior to the filing date and zero otherwise.	LexisNexis (hand-collected)	Helbing et al. (2019)

Intermediary characteristics			
VC backing	Dummy indicating whether the firm is backed by a venture capital investor at the time of the filing. Equals one if firm is backed by venture capital and zero otherwise.	S&P Capital IQ	Boeh and Southam (2011), Busaba et al. (2001), Dunbar and Foerster (2008), Helbing et al. (2019)
Number of underwriters	Number of underwriters involved in the issue.	Thomson Reuters Eikon supplemented with data from S&P Capital IQ	Boeh and Southam (2011)
CM Rank	Carter - Manaster Rank of lead underwriter in filing year. It ranges from 0 to 9.	Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`). Originally developed by Carter and Manaster (1990), updated by Carter et al. (1998) and by Loughran and Ritter (2004).	Boeh and Southam (2011), Busaba et al. (2001), Dunbar and Foerster (2008)
Corporate governance characteristics			
Board size	Absolute number of board members	EDGAR filings (hand-collected)	Boeh and Southam (2011), Helbing et al. (2019)

Board experience	Average age of all board members	EDGAR filings (hand-collected)	Boeh and Southam (2011)
Female board members	Ratio of number of female board members to total number of board members.	EDGAR filings (hand-collected)	Helbing et al. (2019)
CEO duality	Dummy that equals one if both the role of CEO and the role of chairman reside with the CEO of the company and zero otherwise.	EDGAR filings (hand-collected)	Boeh and Southam (2011), Helbing et al. (2019)
Lock up	Dummy that equals one if the lock up period that is reported in EDGAR filings is larger than 180 and zero otherwise.	EDGAR filings (hand-collected)	Helbing et al. (2019)
Market characteristics at filing			
Nasdaq (index/100)	NASDAQ composite index (scaled by 100)	FRED, website of Federal Reserve Bank of St. Louis (`https://research.stlouisfed.org/`)	Busaba et al. (2001)
Volatility (filing date)	Expected volatility at filing date: Based on CBOE S&P 100 volatility index (VXO). It measures the expected volatility over the next 30 days. VXO instead of VIX is chosen because it is continuously available over the whole sample period. The calculation of VIX changes in 2003.	CBOE (`http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`)	Helbing et al. (2019)

Volatility (30 days)	Expected volatility 30 days after filing: Based on CBOE S&P 100 volatility index (VXO). It measures the expected volatility over the next 30 days. VXO instead of VIX is chosen because it is continuously available over the whole sample period. The calculation of VIX changes in 2003.	CBOE (`http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`)	Helbing et al. (2019)
Volatility (60 days)	Expected volatility 60 days after filing: Based on CBOE S&P 100 volatility index (VXO). It measures the expected volatility over the next 30 days. VXO instead of VIX is chosen because it is continuously available over the whole sample period. The calculation of VIX changes in 2003.	CBOE (`http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`)	Helbing et al. (2019)
Yield spread	BAA-AAA yield spread at filing date. Difference between the Moody's BAA-rated corporate bonds and the AAA-rated bond.	FRED, webiste of Federal Reserve Bank of St. Louis (`https://research.stlouisfed.org/`)	Dunbar and Foerster (2008), Helbing et al. (2019)
Number of filings	Number of new filings 60 days prior the filing date (per day)	Own calculations based on data from Thomson Reuters Eikon	Busaba et al. (2001), Dunbar and Foerster (2008), Helbing et al. (2019)

Bank lending standard	Bank lending standard published in the most recent report before filing. Based on the quarterly Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices. The particular question was "over the past three months, how have your bank's credit standards for approving applications for C&I loans or credit lines - other than those to be used to finance mergers and acquisitions - to large and middle-market firms and to small firms changed". The response options range from (1) "Tightened considerable" to (5) "Eased considerable". The lending standard measure is then calculated as the difference between the number of loan officers reporting tightening lending standards and the number of officers reporting easing standards divided by the total number of reporting times 100. Thus, a positive value implies that lending standards have tightened and a negative value implies that lending standards have eased.	Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS) (`https://www.federalreserve.gov/data/sloos.htm`)	Bergbrant et al. (2017)
Market characteristics after filing			
Nasdaq return (30 days)	Daily NASDAQ return in percent from filing date to 30 days after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Busaba et al. (2001), Dunbar and Foerster (2008)

Nasdaq return (60 days)	Daily NASDAQ return in percent from filing date to 60 days after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Busaba et al. (2001), Dunbar and Foerster (2008)
Nasdaq std. dev. (30 days)	Standard deviation of NASDAQ returns between filing and 30 days after filing (scaled by 100).	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Busaba et al. (2001), Dunbar and Foerster (2008)
Nasdaq std. dev. (60 days)	Standard deviation of NASDAQ returns between filing and 60 days after filing (scaled by 100).	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Busaba et al. (2001), Dunbar and Foerster (2008)
Yield spread change (30 days)	Change in BAA-AAA yield spread between filing and 30 days after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Dunbar and Foerster (2008)
Yield spread change (60 days)	Change in BAA-AAA yield spread between filing and 60 days after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	Dunbar and Foerster (2008)

Note: Firm age, debt ratio, firm size, asset turnover, filing size, board size and board experience are winsorized at the 1% level in each tail

* The exact measure to account for the respective effect might vary slightly between the paper cited here and the measure used in this paper. However, they include a variable that aims to capture the same effect.

Table A.2: Correlation table

	Filing size	Firm size	Firm age	High-tech	Debt ratio	BM ratio	Asset turnover	Negative news	VC backing	Number of underwriter
Filing size	1.00									
Firm size	0.44***	1.00								
Firm age	0.20***	0.36***	1.00							
High-tech	−0.17 ***	−0.22 ***	−0.31 ***	1.00						
Debt ratio	−0.04	−0.01	0.01	0.01	1.00					
BM ratio	0.24***	0.26***	0.16***	−0.50 ***	0.00	1.00				
Asset turnover	−0.06 **	−0.11 ***	0.07***	−0.15 ***	0.17***	−0.03	1.00			
Negative news	0.08***	0.07***	0.00	0.11***	−0.04 *	−0.03	−0.07 ***	1.00		
VC backing	−0.12 ***	−0.16 ***	−0.27 ***	0.36***	−0.08 ***	−0.17 ***	−0.22 ***	0.10***	1.00	
Number of underwriter	0.34***	0.38***	0.16***	−0.19 ***	−0.01	0.28***	−0.06 **	0.05**	−0.06 **	1.00
CM Rank	0.24***	0.13***	0.08***	0.07***	−0.21 ***	0.03	−0.10 ***	0.16***	0.17***	0.19***
Board size	0.11***	0.17***	0.14***	0.01	−0.02	0.08***	−0.14 ***	0.10***	0.11***	0.11***
Board experience	0.10***	0.14***	0.20***	−0.16 ***	0.04*	0.22***	0.00	0.01	−0.08 ***	0.14***
Female board members	−0.03	0.08***	0.05*	−0.01	−0.01	0.02	−0.05 *	0.01	0.02	0.01
CEO duality	−0.04 *	−0.03	−0.03	0.01	−0.05 **	−0.02	0.02	0.00	−0.03	−0.08 ***
Lock up	−0.09 ***	−0.05 *	−0.03	−0.04	0.12***	−0.02	0.12***	−0.09 ***	−0.11 ***	−0.08 ***
Nasdaq	0.01	0.01	−0.12 ***	0.23***	−0.03	−0.05 *	−0.19 ***	0.05*	0.25***	0.16***
Volatility (filing date)	−0.11 ***	−0.05 **	−0.08 ***	0.08***	−0.01	−0.11 ***	0.06**	−0.04	−0.01	−0.30 ***
Volatility (30 days)	−0.07 ***	−0.03	−0.03	0.06**	−0.01	−0.09 ***	0.06**	−0.04 *	−0.04 *	−0.26 ***
Volatility (60 days)	−0.05 *	−0.04 *	−0.06 **	0.07***	−0.02	−0.06 **	0.03	−0.03	−0.04	−0.24 ***
Yield spread	0.15***	0.09***	0.05*	−0.06 **	0.06**	0.32***	0.01	0.06**	0.06**	0.25***
Number of filings	−0.17 ***	−0.14 ***	−0.14 ***	0.21***	−0.06 **	−0.31 ***	−0.05 *	−0.05 *	0.04*	−0.36 ***
Bank lending standard	−0.02	−0.01	−0.07 **	0.14***	−0.02	−0.02	−0.01	0.06**	0.07**	−0.21 ***
Nasdaq return (30 days)	−0.04 *	−0.04 *	−0.07 **	0.03	−0.02	−0.07 ***	0.00	0.03	0.03	−0.01
Nasdaq return (60 days)	−0.06 **	−0.02	−0.04 *	0.01	0.00	−0.10 ***	0.01	0.01	0.01	−0.02
Nasdaq std. dev. (30 days)	−0.05 *	−0.05 *	−0.13 ***	0.23***	−0.06 **	−0.12 ***	−0.11 ***	0.04	0.17***	−0.16 ***
Nasdaq std. dev. (60 days)	−0.04 *	−0.05 *	−0.14 ***	0.26***	−0.07 ***	−0.14 ***	−0.13 ***	0.04*	0.18***	−0.16 ***
Yield spread change (30 days)	0.01	−0.01	−0.02	0.05*	0.01	0.02	−0.02	−0.05 *	0.03	−0.02
Yield spread change (60 days)	0.01	−0.03	−0.01	0.05*	0.01	0.04	−0.01	−0.03	0.01	−0.02

Table A.2: Correlation table (continued)

	CM Rank	Firm size	Board experience	Female board members	CEO duality	Lock up	Nasdaq	Volatility (filing date)	Volatility (30 days)	Volatility (60 days)
CM Rank	1.00									
Board size	0.16***	1.00								
Board experience	−0.06 **	0.18***	1.00							
Female board members	0.01	0.07***	0.01	1.00						
CEO duality	−0.06 **	−0.09 ***	−0.12 ***	−0.02	1.00					
Lock up	−0.36 ***	−0.10 ***	−0.02	−0.02	0.04*	1.00				
Nasdaq	0.18***	0.08***	−0.06 **	0.01	−0.04	−0.09 ***	1.00			
Volatility (filing date)	−0.02	−0.08 ***	−0.22 ***	−0.01	0.10***	0.06**	−0.05 **	1.00		
Volatility (30 days)	0.01	−0.04	−0.17 ***	−0.01	0.08***	0.03	−0.02	0.74***	1.00	
Volatility (60 days)	0.03	−0.02	−0.17 ***	−0.01	0.07***	0.04	−0.01	0.60***	0.73***	1.00
Yield spread	0.09***	0.07***	0.15***	0.03	−0.06 **	−0.03	−0.08 ***	0.07***	0.03	0.05*
Number of filings	0.01	−0.08 ***	−0.29 ***	−0.07 **	0.10***	0.03	0.32***	0.42***	0.36***	0.29***
Bank lending standard	0.06**	0.02	−0.16 ***	0.00	0.06**	0.01	0.11***	0.48***	0.45***	0.50***
Nasdaq return (30 days)	−0.02	−0.04 *	−0.05 **	−0.01	−0.01	0.05*	−0.06 **	0.02	−0.36 ***	−0.15 ***
Nasdaq return (60 days)	−0.06 **	−0.05 *	−0.04	−0.02	0.02	0.05*	−0.18 ***	0.07***	−0.20 ***	−0.36 ***
Nasdaq std. dev. (30 days)	0.13***	0.01	−0.19 ***	−0.03	0.06**	−0.04 *	0.69***	0.33***	0.36***	0.31***
Nasdaq std. dev. (60 days)	0.14***	0.03	−0.20 ***	−0.03	0.05*	−0.05 *	0.73***	0.32***	0.33***	0.34***
Yield spread change (30 days)	0.03	−0.01	−0.02	−0.01	0.02	−0.04 *	0.15***	0.14***	0.26***	0.21***
Yield spread change (60 days)	0.03	0.00	−0.02	−0.02	0.00	−0.04 *	0.20***	0.09***	0.32***	0.35***

Table A.2: Correlation table (continued)

	Yield spread	Number of filings	Bank lending standard	Nasdaq return (30 days)	Nasdaq return (60 days)	Nasdaq std dev (30 days)	Nasdaq std dev (60 days)	Yield spread change (30 days)	Yield spread change (60 days)
Yield spread	1.00								
Number of filings	−0.49 ***	1.00							
Bank lending standard	0.36***	0.07**	1.00						
Nasdaq return (30 days)	−0.06 **	−0.06 **	−0.15 ***	1.00					
Nasdaq return (60 days)	−0.08 ***	−0.03	−0.23 ***	0.70***	1.00				
Nasdaq std. dev. (30 days)	−0.12 ***	0.58***	0.31***	−0.18 ***	−0.25 ***	1.00			
Nasdaq std. dev. (60 days)	−0.16 ***	0.57***	0.33***	−0.10 ***	−0.20 ***	0.89***	1.00		
Yield spread change (30 days)	−0.13 ***	0.13***	0.00	−0.24 ***	−0.26 ***	0.21***	0.23***	1.00	
Yield spread change (60 days)	−0.19 ***	0.12***	−0.01	−0.28 ***	−0.38 ***	0.28***	0.31***	0.74***	1.00

Table A.3: In-sample prediction 30 and 60 days after filing

	Random Forest	Lasso ($\lambda.min$)	Logit
30 days after filing			
AUC	0.7371	0.6749	0.6772
Cohen's Kappa	0.1703	0.0838	0.1072
F1 Score	0.2418	0.1235	0.1581
Accuracy	0.7899	0.7829	0.7841
60 days after filing			
	Random Forest	Lasso ($\lambda.min$)	Logit
AUC	0.7319	0.6841	0.6879
Cohen's Kappa	0.1646	0.0951	0.1172
F1 Score	0.2515	0.1536	0.1933
Accuracy	0.7671	0.7610	0.7600

Note: In-sample prediction of 30 and 60 days after filing. The same settings are applied as in the baseline model (prediction at filing date) reported in Table 2.4.

Table A.4: Out-of-sample prediction 30 and 60 days after filing

	Random Forest	Lasso ($\lambda.min$)	Logit
Prediction 30 days after filing			
Cross-sectional prediction (Split 70:30)			
AUC	0.7175	0.6473	0.6475
Cohen's Kappa	0.1006	0.0577	0.0782
F1 Score	0.1690	0.0833	0.1281
Accuracy	0.7457	0.7471	0.7457
Prediction over time (Split 70:30) *Training set: 1997-2005* *Test set: 2007-2014*			
AUC	0.4522	0.5047	0.4511
Cohen's Kappa	0.0075	0.0063	-0.0379
F1 Score	0.0104	0.0394	0.1158
Accuracy	0.7194	0.7120	0.6617
Prediction 60 days after filing			
Cross-sectional prediction (Split 70:30)			
AUC	0.7202	0.6648	0.6634
Cohen's Kappa	0.0989	0.0714	0.1141
F1 Score	0.1910	0.1311	0.1885
Accuracy	0.7326	0.7359	0.7425
Prediction over time (Split 70:30) *Training set: 1997-2005* *Test set: 2007-2014*			
AUC	0.5072	0.5375	0.4932
Cohen's Kappa	0.0345	0.0254	-0.0270
F1 Score	0.0518	0.0784	0.1388
Accuracy	0.6624	0.6531	0.6107

Note: Out-of-sample prediction of 30 and 60 days after filing. The same settings are applied as in the baseline model (prediction at filing date). Splitting proportion is (approx.) 70:30.

Table A.5: Test for concept drift (based on logit)

	Time split Training set: 1997-2005 Test set: 2007-2015		Random Split (70:30)	
	(1) restricted	(2) unrestricted	(3) restricted	(4) unrestricted
Filing size	-0.5873	-2.1338***	-0.2924	-0.2641
	[0.4680]	[0.7773]	[0.4292]	[0.4978]
Firm size	0.0221	0.0482	0.0109	0.0215
	[0.0412]	[0.0558]	[0.0409]	[0.0489]
Firm age	-0.0094**	-0.0107**	-0.0098***	-0.0156***
	[0.0037]	[0.0048]	[0.0036]	[0.0046]
Debt ratio	0.1769***	0.2800***	0.1698***	0.2170***
	[0.0619]	[0.0828]	[0.0604]	[0.0698]
BM ratio	1.3929***	1.6525***	1.3117***	1.0296***
	[0.3120]	[0.4231]	[0.3060]	[0.3575]
High-tech	0.0146	-0.0181	-0.0195	-0.2313
	[0.1384]	[0.1789]	[0.1336]	[0.1570]
Asset turnover	0.0431	0.0407	0.0482	0.0352
	[0.0480]	[0.0603]	[0.0466]	[0.0552]
Negative news	-0.0071	-0.3206**	0.0217	0.0104
	[0.1175]	[0.1555]	[0.1133]	[0.1347]
Board size	0.0108	0.0322	0.0140	0.0222
	[0.0254]	[0.0315]	[0.0245]	[0.0289]
Board experience	-0.0053	-0.0090	-0.0042	-0.0046
	[0.0096]	[0.0121]	[0.0093]	[0.0110]
Female board members	-0.4121	-0.5765	-0.5667	-1.7512**
	[0.6086]	[0.8052]	[0.5951]	[0.7579]
CEO duality	-0.0797	0.0060	-0.1121	-0.0423
	[0.1057]	[0.1346]	[0.1024]	[0.1222]
Lock up	-0.8396**	-0.6265	-0.8997**	-1.0956**
	[0.4018]	[0.4427]	[0.3997]	[0.4767]
VC backing	0.0169	-0.1266	0.0061	-0.0601
	[0.1198]	[0.1510]	[0.1167]	[0.1400]
Number of underwriters	-0.2740***	0.1354	-0.2505***	-0.2801***
	[0.0707]	[0.1839]	[0.0677]	[0.0818]
CM Rank	0.0035	0.0541	-0.0083	-0.0223
	[0.0345]	[0.0460]	[0.0330]	[0.0414]
Nasdaq	0.0438***	0.0640***	0.0388***	0.0434***
	[0.0065]	[0.0087]	[0.0064]	[0.0077]
Volatility (filing date)	0.0385***	0.0198	0.0300***	0.0286**
	[0.0097]	[0.0139]	[0.0098]	[0.0117]
Yield spread	0.2900	-1.1758**	0.5094*	0.4719
	[0.3020]	[0.5076]	[0.2938]	[0.3437]
Number of filings	-2.9259***	-3.7420***	-1.1545	-1.7710**
	[0.9483]	[1.1506]	[0.7161]	[0.8738]
Bank lending standard	0.0015	0.0040	0.0010	0.0012
	[0.0036]	[0.0054]	[0.0036]	[0.0042]
test set		5.9154***		-2.1313
		[1.7862]		[1.4764]
test set * Filing size		1.9234*		0.0203
		[1.0829]		[1.0145]
test set * Firm size		0.0697		-0.0625

		[0.0926]		[0.0929]
test set * Firm age		-0.0009		0.0171**
		[0.0082]		[0.0074]
test set * Debt ratio		-0.2432*		-0.2333
		[0.1301]		[0.1491]
test set * BM ratio		-2.1109***		1.2923*
		[0.7034]		[0.7228]
test set * High-tech		-0.1553		0.7640**
		[0.3077]		[0.3107]
test set * Asset turnover		0.0444		0.0272
		[0.1087]		[0.1058]
test set * Negative news		0.5914**		0.0257
		[0.2598]		[0.2556]
test set * Board size		-0.0791		-0.0295
		[0.0598]		[0.0563]
test set * Board experience		0.0013		0.0069
		[0.0224]		[0.0215]
test set * Female board members		0.4368		3.4978***
		[1.3336]		[1.2391]
test set * CEO duality		-0.1927		-0.2703
		[0.2379]		[0.2294]
test set * Lock up		-1.6288		0.4026
		[1.2473]		[0.9218]
test set * VC backing		0.0597		0.2639
		[0.2781]		[0.2618]
test set * Number of underwriters		-0.6410***		0.1222
		[0.2147]		[0.1523]
test set * CM Rank		-0.1274*		0.0209
		[0.0753]		[0.0708]
test set * Nasdaq		-0.1145***		-0.0140
		[0.0219]		[0.0144]
test set * Volatility (filing date)		0.0733***		0.0100
		[0.0236]		[0.0220]
test set * Yield spread		-0.5298		0.3468
		[0.8054]		[0.6844]
test set * Number of filings		-4.1929		2.1038
		[5.4840]		[1.5507]
test set * Bank lending standard		0.0081		0.0002
		[0.0089]		[0.0082]
Constant	-2.8621***	-2.4820***	-2.8384***	-2.3509***
	[0.6683]	[0.9316]	[0.6458]	[0.7676]
Observations	2,323	2,323	2,444	2,444
Test for joint significance				
Chi2		141.6000		26.8900
P-value		0.0000		0.2150

Note: Test for joint significance of interaction terms in the logit model. Test set is a dummy variable that equals 1 if the observation belongs to the test set and zero otherwise. It is interacted with each predictor variable. Both the restriced models (excluding the interaction terms) and the unrestricted models (including the interaction terms) are reported. A test for joint significance is reported on the interaction terms. Variable descriptions and data sources are described in Table A.1. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table A.6: Test for concept drift (based on OLS)

	Time split Training set: 1997-2005 Test set: 2007-2015		Random Split (70:30)	
	(1) restricted	(2) unrestricted	(3) restricted	(4) unrestricted
Filing size	-0.5873	-2.1338***	-0.2924	-0.2641
	[0.4680]	[0.7773]	[0.4292]	[0.4978]
Firm size	0.0221	0.0482	0.0109	0.0215
	[0.0412]	[0.0558]	[0.0409]	[0.0489]
Firm age	-0.0094**	-0.0107**	-0.0098***	-0.0156***
	[0.0037]	[0.0048]	[0.0036]	[0.0046]
Debt ratio	0.1769***	0.2800***	0.1698***	0.2170***
	[0.0619]	[0.0828]	[0.0604]	[0.0698]
BM ratio	1.3929***	1.6525***	1.3117***	1.0296***
	[0.3120]	[0.4231]	[0.3060]	[0.3575]
High-tech	0.0146	-0.0181	-0.0195	-0.2313
	[0.1384]	[0.1789]	[0.1336]	[0.1570]
Asset turnover	0.0431	0.0407	0.0482	0.0352
	[0.0480]	[0.0603]	[0.0466]	[0.0552]
Negative news	-0.0071	-0.3206**	0.0217	0.0104
	[0.1175]	[0.1555]	[0.1133]	[0.1347]
Board size	0.0108	0.0322	0.0140	0.0222
	[0.0254]	[0.0315]	[0.0245]	[0.0289]
Board experience	-0.0053	-0.0090	-0.0042	-0.0046
	[0.0096]	[0.0121]	[0.0093]	[0.0110]
Female board members	-0.4121	-0.5765	-0.5667	-1.7512**
	[0.6086]	[0.8052]	[0.5951]	[0.7579]
CEO duality	-0.0797	0.0060	-0.1121	-0.0423
	[0.1057]	[0.1346]	[0.1024]	[0.1222]
Lock up	-0.8396**	-0.6265	-0.8997**	-1.0956**
	[0.4018]	[0.4427]	[0.3997]	[0.4767]
VC backing	0.0169	-0.1266	0.0061	-0.0601
	[0.1198]	[0.1510]	[0.1167]	[0.1400]
Number of underwriters	-0.2740***	0.1354	-0.2505***	-0.2801***
	[0.0707]	[0.1839]	[0.0677]	[0.0818]
CM Rank	0.0035	0.0541	-0.0083	-0.0223
	[0.0345]	[0.0460]	[0.0330]	[0.0414]
Nasdaq	0.0438***	0.0640***	0.0388***	0.0434***
	[0.0065]	[0.0087]	[0.0064]	[0.0077]
Volatility (filing date)	0.0385***	0.0198	0.0300***	0.0286**
	[0.0097]	[0.0139]	[0.0098]	[0.0117]
Yield spread	0.2900	-1.1758**	0.5094*	0.4719
	[0.3020]	[0.5076]	[0.2938]	[0.3437]
Number of filings	-2.9259***	-3.7420***	-1.1545	-1.7710**
	[0.9483]	[1.1506]	[0.7161]	[0.8738]
Bank lending standard	0.0015	0.0040	0.0010	0.0012
	[0.0036]	[0.0054]	[0.0036]	[0.0042]
test set		5.9154***		-2.1313
		[1.7862]		[1.4764]
test set * Filing size		1.9234*		0.0203
		[1.0829]		[1.0145]
test set * Firm size		0.0697		-0.0625

		[0.0926]		[0.0929]
test set * Firm age		-0.0009		0.0171**
		[0.0082]		[0.0074]
test set * Debt ratio		-0.2432*		-0.2333
		[0.1301]		[0.1491]
test set * BM ratio		-2.1109***		1.2923*
		[0.7034]		[0.7228]
test set * High-tech		-0.1553		0.7640**
		[0.3077]		[0.3107]
test set * Asset turnover		0.0444		0.0272
		[0.1087]		[0.1058]
test set * Negative news		0.5914**		0.0257
		[0.2598]		[0.2556]
test set * Board size		-0.0791		-0.0295
		[0.0598]		[0.0563]
test set * Board experience		0.0013		0.0069
		[0.0224]		[0.0215]
test set * Female board members		0.4368		3.4978***
		[1.3336]		[1.2391]
test set * CEO duality		-0.1927		-0.2703
		[0.2379]		[0.2294]
test set * Lock up		-1.6288		0.4026
		[1.2473]		[0.9218]
test set * VC backing		0.0597		0.2639
		[0.2781]		[0.2618]
test set * Number of underwriters		-0.6410***		0.1222
		[0.2147]		[0.1523]
test set * CM Rank		-0.1274*		0.0209
		[0.0753]		[0.0708]
test set * Nasdaq		-0.1145***		-0.0140
		[0.0219]		[0.0144]
test set * Volatility (filing date)		0.0733***		0.0100
		[0.0236]		[0.0220]
test set * Yield spread		-0.5298		0.3468
		[0.8054]		[0.6844]
test set * Number of filings		-4.1929		2.1038
		[5.4840]		[1.5507]
test set * Bank lending standard		0.0081		0.0002
		[0.0089]		[0.0082]
Constant	-2.8621***	-2.4820***	-2.8384***	-2.3509***
	[0.6683]	[0.9316]	[0.6458]	[0.7676]
Observations	2,323	2,323	2,444	2,444
Test for joint significance				
Chi2		141.6000		26.8900
P-value		0.0000		0.2150

Note: Test for joint significance of interaction terms in the logit model. Test set is a dummy variable that equals 1 if the observation belongs to the test set and zero otherwise. It is interacted with each predictor variable. Both the restriced models (excluding the interaction terms) and the unrestricted models (including the interaction terms) are reported. A test for joint significance is reported on the interaction terms. Variable descriptions and data sources are described in Table A.1. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table A.7: Robustness checks

		Withdrawal probability		Out of sample prediction AUC			Test for joint significance	
Training set	Test set	Training set	Test set	Random forest	Lasso	Logit	Chi2	p-value
Time splits								
1997 - 2005 (70%)	2007 - 2014 (30%)	19.96%	27.63%	0.4235	0.4190	0.4060	141.60	0.0000
1997 - 2003 (60%)	2005 - 2014 (40%)	20.48%	25.19%	0.4025	0.5389	0.5286	144.30	0.0000
1997 - 2006 (75%)	2008 - 2014 (25%)	20.18%	23.17%	0.4557	0.4305	0.3869	83.77	0.0000
1997 - 2008	2010 - 2014	22.79%	19.88%	0.5145	0.4581	0.4531	83.63	0.0000
1997 - 2008	2010 - 2011	22.04%	31.00%	0.5178	0.5601	0.5501	31.37	0.0677
1997 - 2008 (exclude crisis years)	2010 - 2014 (exclude crisis years)	20.76%	19.88%	0.4809	0.3723	0.3805	119.00	0.0000
Random splits								
70%	30%	22.34%	22.21%	0.7343	0.6194	0.6124	26.89	0.2150
60%	40%	22.22%	22.42%	0.7062	0.6000	0.5967	29.75	0.1250
75%	25%	22.38%	22.05%	0.7665	0.6271	0.6213	24.57	0.3180

Note: Robustness checks on varying training and test sets. Performance of out-of-sample prediction of random forest, lasso and logit is reported in terms of AUC (which is the preferred measure). The last two colums report the results of the tests for joint significance, which are analogously performed to the tests in Table A.5.

A.2 Figures

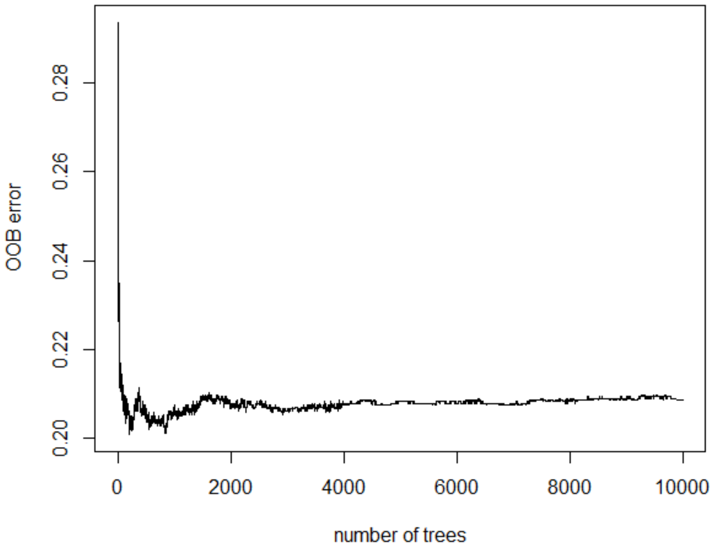


Figure A.1: OBB error against number of trees

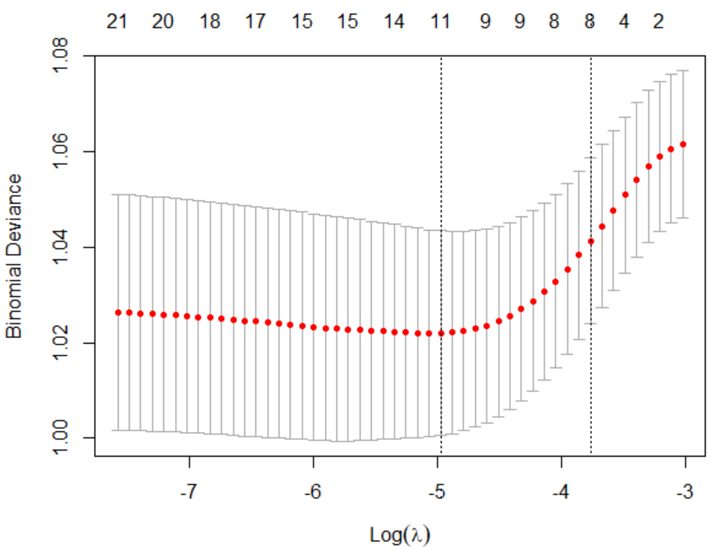


Figure A.2: Binomial deviance against different number of trees

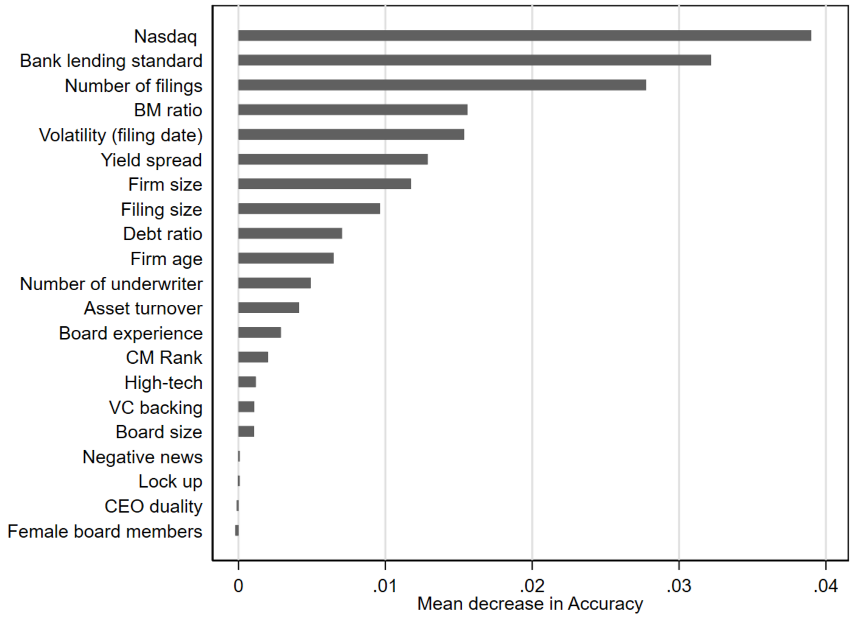


Figure A.3: Variable importance of random forest based on accuracy

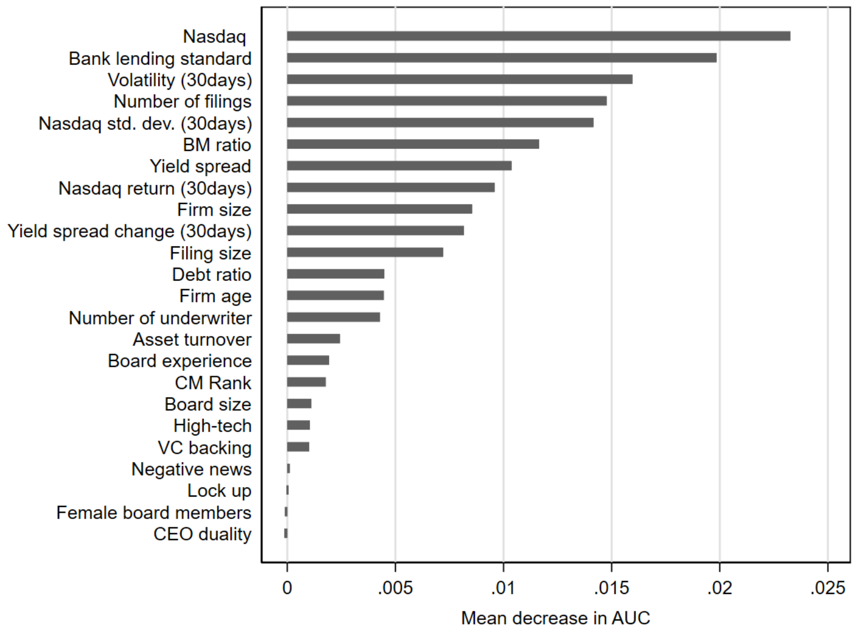


Figure A.4: Variable importance of random forest based on AUC - 30 days after filing

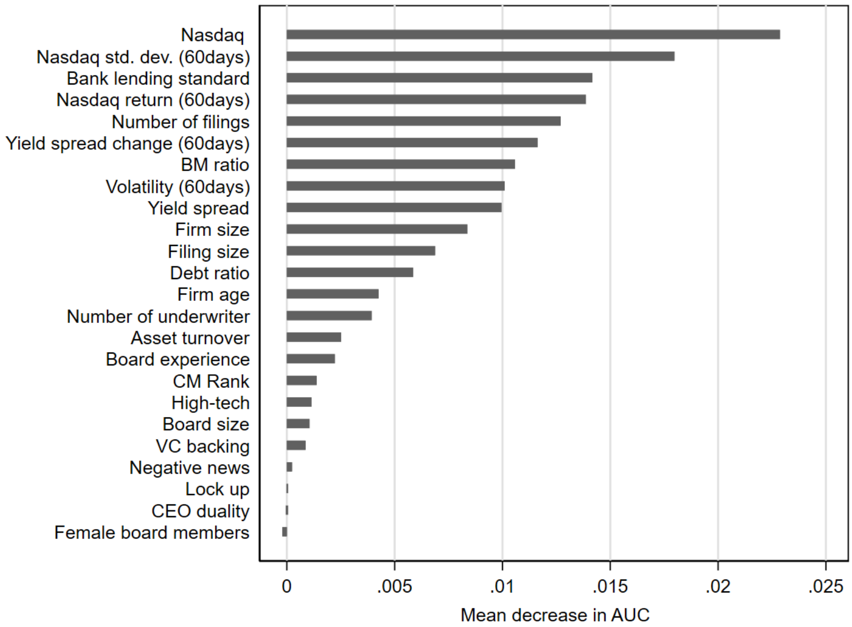


Figure A.5: Variable importance of random forest based on AUC - 60 days after filing

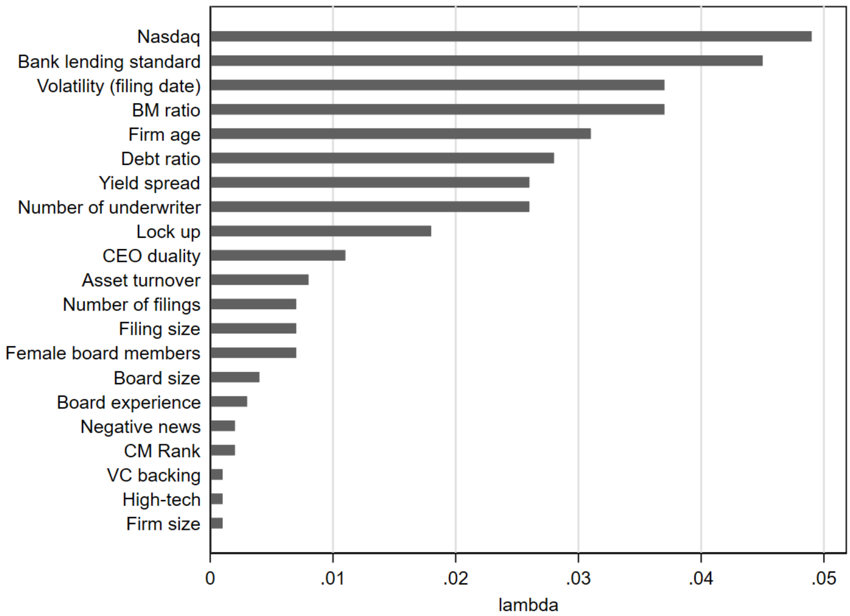


Figure A.6: Lasso solution path - adaptive lasso

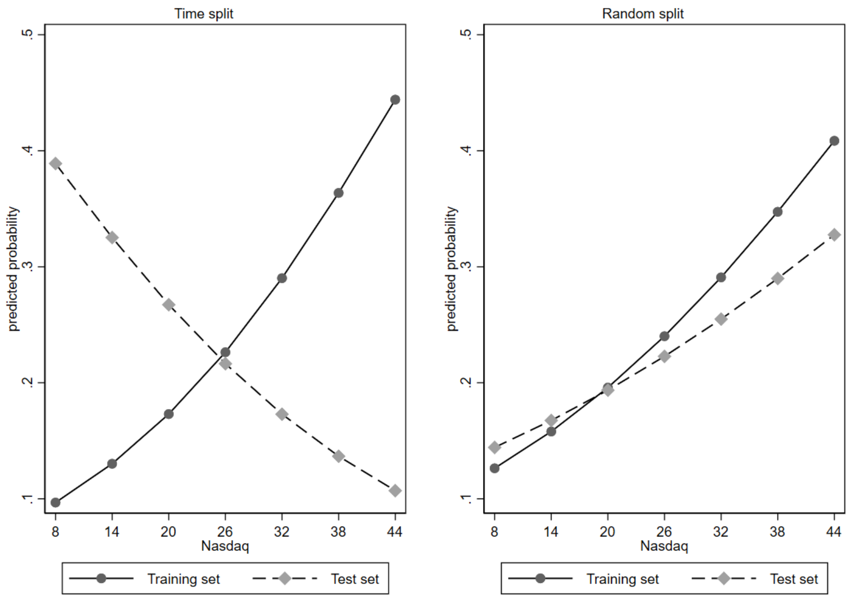


Figure A.7: Predicted probability against Nasdaq

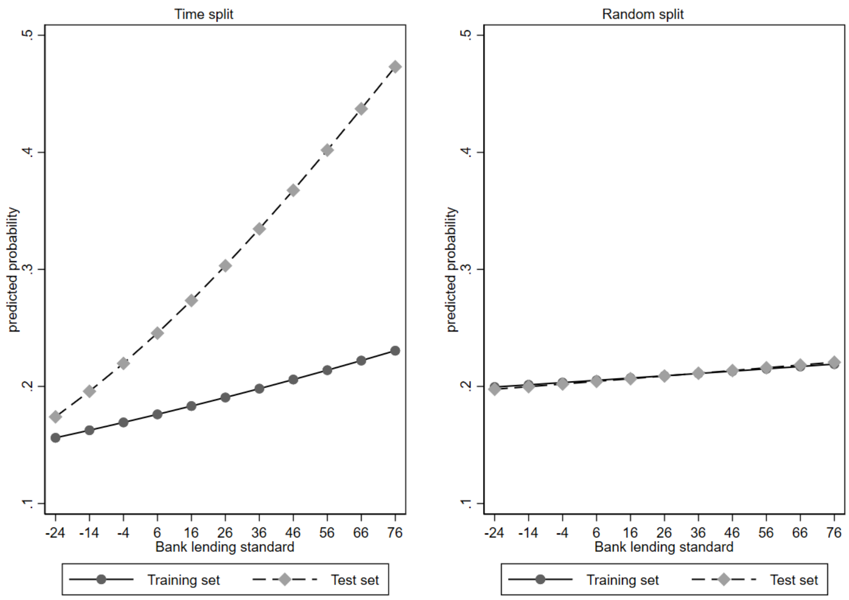


Figure A.8: Predicted probability against bank lending standard

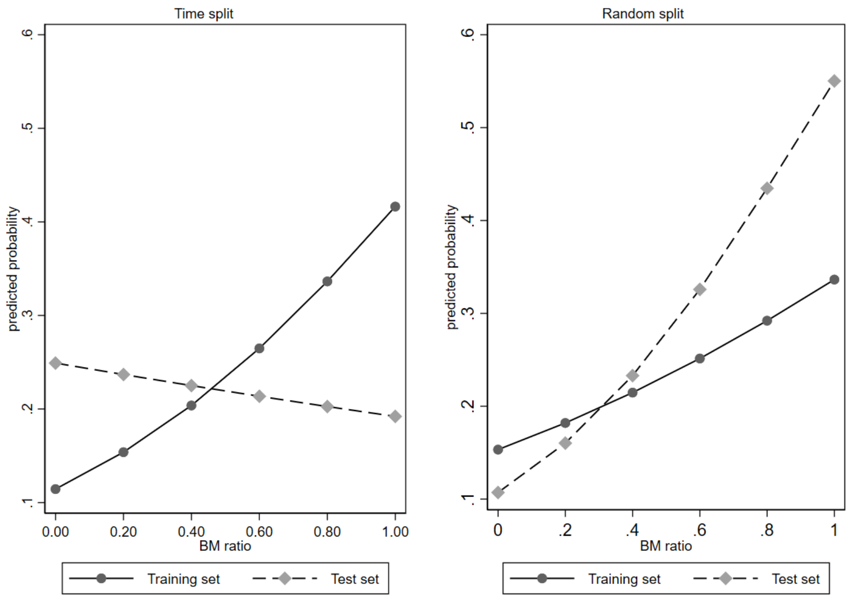


Figure A.9: Predicted probability against BM ratio

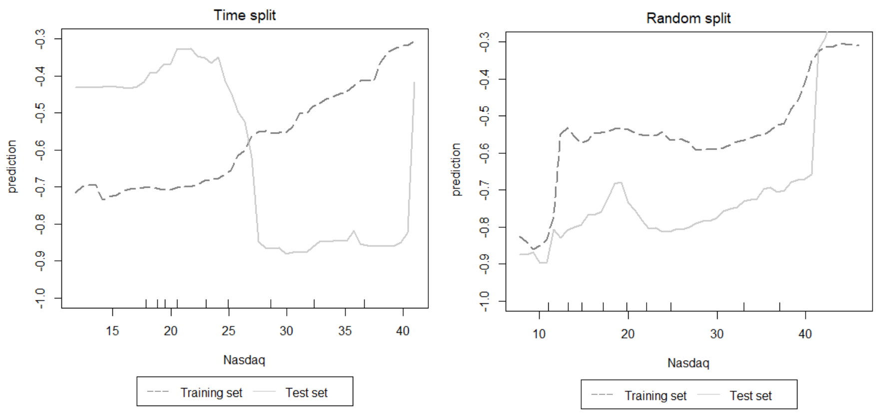


Figure A.10: Partial dependence plot of Nasdaq (obtained by random forest)

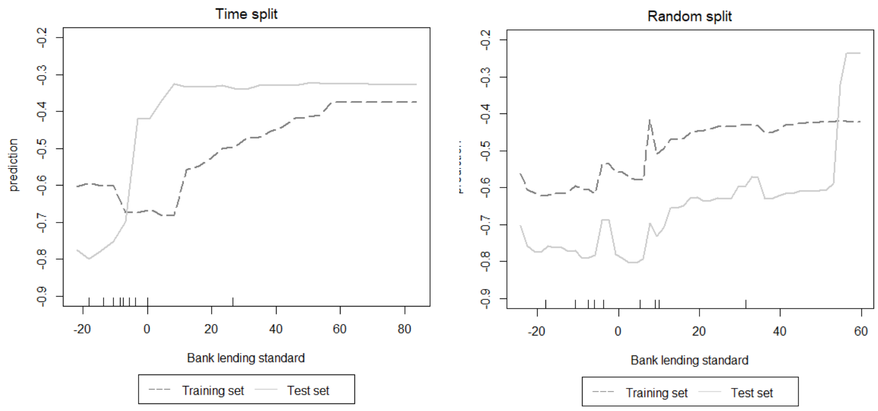


Figure A.11: Partial dependence plot of bank lending standard (obtained by random forest)

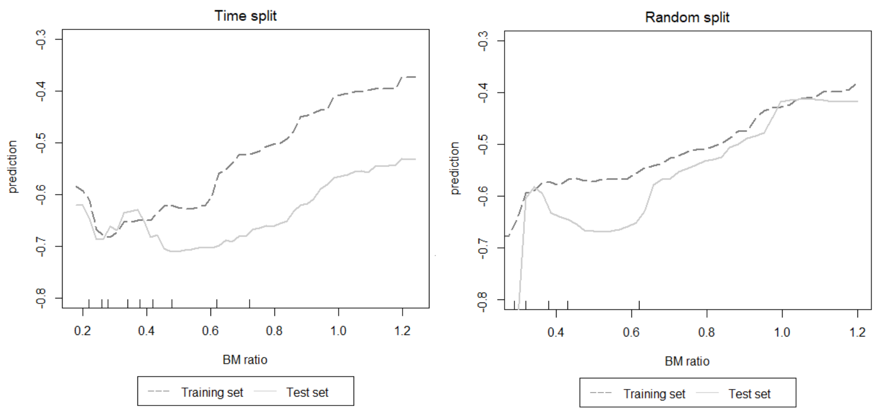


Figure A.12: Partial dependence plot of BM ratio (obtained by random forest)

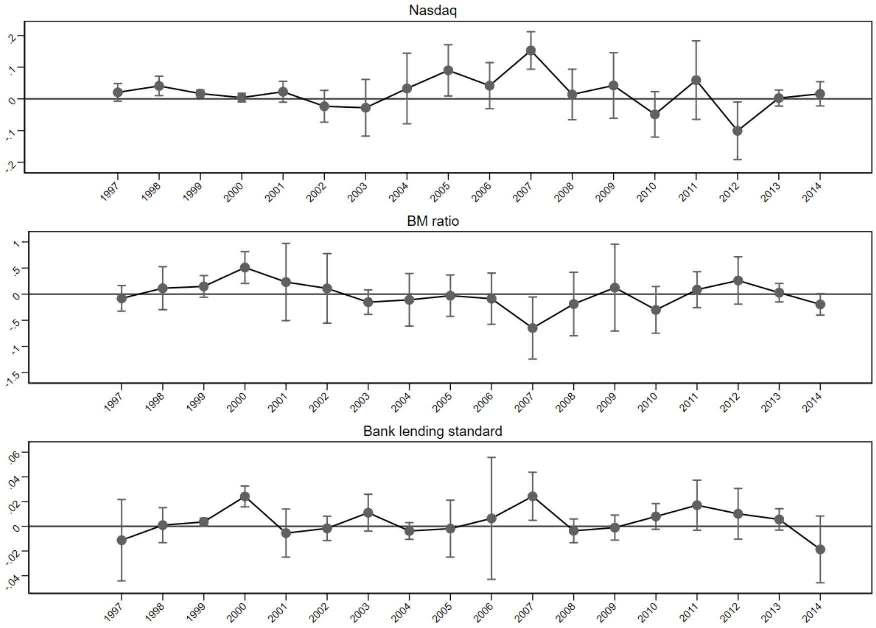


Figure A.13: Marginal effect of different variables on withdrawal probability for different years (models are estimated by OLS, 95% confidence intervals are depicted)

Chapter 3

IPO withdrawals: Are corporate governance and VC characteristics the guiding light in the rough sea of volatile markets?[1]

3.1 Introduction

One of the most important decisions for private firms is whether to stay private or to launch an initial public offering (IPO). Several studies (e.g. Bodnaruk et al., 2007; Chemmanur and Fulghieri, 1999; Chemmanur et al., 2009; Chemmanur and He, 2011; Kim and Weisbach, 2008; Pagano et al., 1998; Zingales, 1995) consider costs and benefits associated with this decision. This paper focuses on firms that decided to go public and launched the official registration process with the Securities and Exchange Commission (SEC) but withdrew from the IPO prior to the listing. In the US, more than 25% of all firms that file the security registration documents with the SEC do not complete the IPO process. In some years (e.g. 2008) even the majority of IPOs filed with the SEC were withdrawn.

Busaba et al. (2001 and 2015) and Busaba (2006) argue that firms receive a valuable option when they file for an IPO. The firm will exercise this option (and go public) when investor valuation exceeds firm's reservation value. When firm's reservation value is above investor valuation, the firm will withdraw its offering and the option expires worthless. At the time of the IPO filing, issuers do not know the exact investor valuations, so the

[1]This chapter is joint work with Tereza Tykvová from University of St. Gallen. It was originally published by Elsevier as:
Reiff, A. and Tykvová, T. (2021): "IPO withdrawals: Are corporate governance and VC characteristics the guiding light in the rough sea of volatile markets?", *Journal of Corporate Finance*, 67, 101908. Available online at `https://doi.org/10.1016/j.jcorpfin.2021.101908`.

offer price is uncertain. After the filing date, as this information gets revealed during the price discovery process, issuers may withdraw their IPO. In addition, new information may arrive after the filing date that may change investor valuation or issuer reservation value. Several studies (see Boeh and Southam, 2011; Busaba et al., 2001; Dunbar, 1998; Dunbar and Foerster, 2008; Helbing et al., 2019, among others) analyze various factors that may affect firm's reservation value and investor valuation and thus may be related to the probability of a withdrawal. We add to this literature by focusing on the role of signaling in the presence of agency costs that arise during the IPO process.

Two prominent issues of agency costs are adverse selection and moral hazard. Adverse selection results from the situation in which investors face issuers of unknown quality (Latham and Braun, 2010), which reduces their average valuations (Akerlof, 1970). In this setting, investors may rely on various signals that decrease their uncertainty regarding the value of the firm (Owen-Smith et al., 2015; Benveniste et al., 2002), such as venture capital (VC) backing. Moral hazard arises because the interests of firm insiders are not aligned with the interests of investors (Brav and Gompers, 2003). In this situation, investors benefit if effective monitoring mechanisms are in place, such as an experienced board (Boeh and Southam, 2011; Helbing et al., 2019). Insiders also may demonstrate their incentive alignment with the investors through signals, such as a high fraction of retained shares or a lock-up commitment (Boeh and Southam, 2011; Busaba et al., 2001; Helbing et al., 2019).

We examine the role of various corporate governance and VC financing signals for the withdrawal probability and argue that this role may differ in different market environments. More specifically, we expect that in more risky market environments when investors tend to be more careful, signals become more important than if market conditions are more stable. We test this assumption by interacting an indicator for high post-filing market volatility with the corporate governance and VC characteristics.

For VC financing, however, the effect on the withdrawal probability is not straightforward from a theoretical point of view. On the one hand, VCs may signal the quality of the firm and diminish investor uncertainty, which should reduce the withdrawal probability. On the other hand, VCs' goal is to maximize the exit proceeds, which may be associated with a higher withdrawal probability. To maximize proceeds, VCs may tend to follow a dual-track strategy by filing for an IPO but being involved in M&A negotiations at the same time (Brau et al., 2010). Their portfolio firms may thus be less dependent on the IPO success. VCs may also prefer to postpone the IPO of their portfolio firms if current investor valuations deteriorate (Fan and Yamada, 2020). Our goal is to understand whether the VC signaling effect or the dual-track / postponing strategy prevails and whether there are differences under different market conditions. We also examine how the outcome depends on VC characteristics. Hereby, we focus on three features: syndicated vs. stand-alone VCs, domestic vs. foreign VCs, and VCs with high vs. low reputation. For example, reputable VCs could provide a stronger quality signal than their less reputable counterparts. However, they might also provide their portfolio companies with a broader range of exit options.

Our study provides two main contributions. We first deliver new findings to the literature on signaling in the context of withdrawals. According to the signaling theory, signals of high quality, effective monitoring, and incentive alignment (such as high corporate governance standards or VC backing) may reduce investor uncertainty and increase their valuations (Signori, 2018). We add to the literature by distinguishing between the role of signals in more and less risky market environments. Typically, firms blame deteriorating market conditions after the filing, which lead to decreasing investor valuations, to be the reason for the withdrawal (Brau and Fawcett, 2006). Most studies based on older US data support this view (e.g Dunbar and Foerster, 2008).[2] Two studies consider explicitly the withdrawal decision during unfavorable market conditions. The first study by Fan and Yamada (2020) examines the relation between the level of retained shares of various types of owners (CEOs and VCs) and the withdrawal probability in Japan. The second study is Latham and Braun (2010) who focus on the link between the CEO ownership and the withdrawal probability by examining US filings after the burst of the high-tech bubble. Both studies focus on the relation between ownership structures and the withdrawal event, while we include various aspects of corporate governance and VC financing. Instead of signaling, which is at the core of our study, these studies rather focus on motivations of existing shareholders when they take a firm public. In addition, Fan and Yamada (2020) only provide descriptive statistics for a relatively small subsample of deals under weak market conditions (between 7 and 13 withdrawn deals) and compare them to the whole sample. Latham and Braun (2010) explicitly consider filings within weak market conditions only by including a time window of 2000-2002. In contrast, we perform a multivariate analysis and consider a long time period that covers varying market conditions.

As our second main contribution, we provide new insights to the controversial discussion on the role of VCs. From the theoretical perspective, arguments in both directions exist. Prior empirical findings on the VC effect on withdrawals are mixed too. On the one hand, VC investors could serve as a signaling device (e.g. Cumming et al., 2016) that reduces investors uncertainty and should thus be associated with a lower withdrawal probability (Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008). On the other

[2]While older literature focuses on the US exclusively, recently researchers started to analyze IPO withdrawals in Europe (Helbing et al., 2019) or in Asia (Fan and Yamada, 2020). Recent evidence from European countries (Helbing et al., 2019) suggests that market conditions are not particularly relevant there.

hand, Boeh and Southam (2011) or Helbing et al. (2019) demonstrate a positive relation between VC financing and withdrawal probability. VCs may tend to use a dual-track strategy, which may coincide with a higher withdrawal probability (Dunbar and Foerster, 2008). VCs also may decide to postpone the IPO if the prospects deteriorate (Fan and Yamada, 2020). So far, the literature has typically considered the average effect of VC backing. We add to this literature by focusing on VC heterogeneity and analyzing how different VC types differ in the withdrawal probabilities.

More generally, we also contribute to the literature on withdrawal determinants (Boeh and Southam, 2011; Busaba et al., 2001; Dunbar, 1998; Dunbar and Foerster, 2008; Helbing et al., 2019). We include more recent US filings compared to older studies and we cover a longer time period.[3] This is important since the intensity and patterns of withdrawals change over time (e.g. Busaba et al., 2001). Table 3.1 shows that the percentage of withdrawn deals varies between 6.1% in 2013 and 60.7% in 2008. In a recent literature review, Helbing (2019) claims that the findings from existing studies are not unanimous and that we still do not fully understand the drivers behind the withdrawal decision. Our study employs a comprehensive set of variables related to firm's reservation value and to investor valuation. We combine variables that have been used in prior studies and add new variables that have not been discussed in the literature so far.

We employ a sample of 3,438 US domestic first-time IPO filings between 1997 and 2014. We find that signals of good corporate governance (with the exception of a longer lock-up period) and VC backing are not related to a lower withdrawal probability on average. However, our results from the interaction term analysis confirm that signals through high-quality corporate governance are important in risky market environments, while they do not matter in less risky market environments. In addition, consistently with signaling theories, we find that firms backed by local VCs and by a VC syndicate face a lower withdrawal probability than firms backed by foreign and stand-alone VCs. The VC syndicate effect is particularly pronounced in highly volatile markets. Our results also give support to the conclusion that reputable VCs are more likely to pursue a dual-track strategy or postpone their IPO under riskier than under less risky market conditions.

The remainder of the paper is structured as follows. In Section 3.2, we discuss our theoretical background. Section 3.3 describes the data and provides descriptive statistics. We discuss issues related to the methodology in Section 3.4. In Section 3.5, we present the empirical results from our analysis. Section 3.6 concludes.

[3] Busaba et al. (2001) include data from 1984 to 1994, Dunbar and Foerster (2008) use 1985 to 2000, Boeh and Dunbar (2013) have a sample of filings between 1999 and 2004 and, finally, Boeh and Dunbar (2016) employ data between 1998 and 2011.

3.2 Theoretical background

3.2.1 IPO withdrawals

Firms that want to go public time their IPOs by exploiting periods in which they expect investor valuations to be relatively high compared to their reservation values (Allen and Faulhaber, 1989). However, at the time of the IPO filing, issuers are uncertain about investor valuations. Busaba et al. (2001 and 2015) and Busaba (2006) argue that filing thus represents a valuable option on a future IPO.[4] This option is costly. The filing process is associated not only with direct costs, i.e. expenditures related to the IPO preparation and to the registration process, but also with indirect costs such as management effort devoted to these activities (Ritter, 1987). These costs are sunk if the firm withdraws its offering. In addition, the withdrawal event may send negative information about the firm to the market participants (Bouis, 2009; Lian and Wang, 2009).

To understand why issuers withdraw their offerings, Busaba (2006) addresses the "two-perspective firm valuation" framework: The issuer compares investor valuations, which are revealed after the filing date, to the firm's reservation value. If investor valuations drop below the firm's reservation value, the issuer will not exercise the option to go public. Rather, the issuer will withdraw the offering and the firm stays private. This situation may arise either if investor valuations fall below expectations or if firm's reservation value increases (Busaba, 2006; Busaba et al., 2015).

Investor valuations may drop below firm's reservation value when negative information about the firm is revealed. This decrease is more likely when ex-ante uncertainty about the firm value is high (Helbing, 2019) or when agency costs are prevalent. The ex-ante uncertainty and the extent of agency costs depend on various characteristics related to the issuer, issue, corporate governance, and intermediaries. Investor valuations will also be affected by the overall market situation and its development after the filing date.

In the following subsections, we provide the theoretical motivation behind the variables we include in our analysis and discuss their expected effect on the withdrawal probability. Following Dunbar and Foerster (2008), we divide our variables into the following four categories: issuer and issue characteristics, intermediary characteristics, market characteristics at filing and, finally, market characteristics after filing. Later studies (e.g. Boeh and Southam, 2011; Helbing et al., 2019) point to issuer corporate governance characteristics that may also be related to the likelihood of an IPO withdrawal. We include them as a separate category. In each category, we select the relevant determinants by combining

[4] Note that we use Busaba et al.'s (2015) definition of the "option" here, that is, the firm's option to accept a high price and go public. They argue that this option is the complement of the option to reject a low price and withdraw in Busaba et al. (2001) and Busaba (2006).

variables that have been used in various prior studies on withdrawals (Bergbrant et al., 2017; Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008; Helbing et al., 2019). We also include additional variables that we believe might be relevant. Table B.1 summarize our predictions about the relation between each variable and the probability of a withdrawal.

Among the relevant characteristics, we focus on corporate governance and VC financing and argue that they may serve as potential signals. In addition, VCs (or certain VC types) may also increase firm's chances for an alternative exit to a current IPO. We expect that the role of signals for the withdrawal probability may differ in different market environments. In more risky market environments when investors tend to be more careful, quality signals become more important than if market conditions are more stable.

3.2.2 Corporate governance characteristics

High-quality corporate governance may provide a signal that reduces agency costs (Boeh and Southam, 2011; Helbing et al., 2019) and, thus, increases investor valuations. Agency costs arise for two main reasons. First, firm insiders may be better informed about the quality of the firm than the new investors. Such information asymmetry gives rise to adverse selection (Latham and Braun, 2010) where investors reduce their valuations because they fear that the issuer may be of a low quality (Akerlof, 1970). In this situation, high-quality corporate governance may signal the underlying quality of the firm to the new investors and thus lead to a higher valuation.

Second, firm insiders may follow their own goals, which do not coincide with the goals new investors have. Such misaligned incentives give rise to a moral hazard problem (Brav and Gompers, 2003). When investors expect such behavior, they reduce their valuations. High-quality corporate governance, which leads to a better monitoring of firm insiders, may be beneficial in reducing the moral hazard problem.

Boeh and Southam (2011) and Helbing et al. (2019) suggest that larger and more experienced boards tend to improve monitoring and mitigate the moral hazard problem. On the opposite, the monitoring is weaker and the potential for moral hazard increases when the CEO also occupies the chair position of the board of directors (CEO duality). Larger and more experienced boards and the separation between the CEO and the chair could also serve as signals of the underlying firm quality and thus reduce adverse selection. We therefore include board size, board experience, and CEO duality in this variable category. In addition, Helbing et al. (2019) argues that the ratio of female board members is another signal that reduces agency costs. More diversified boards combine various perspectives and experiences that may contribute to a better monitoring. We thus add the ratio of female board members. We expect that larger and more experienced boards, a higher ratio of female board members, as well as the separation of CEO and chair are associated with a higher investor valuation and, thus, a lower withdrawal probability.

When insiders commit to a longer lock-up period or when they retain a larger proportion of shares, investors may interpret it as a positive signal of insiders' confidence in the firm and of a better incentive alignment (Boeh and Southam, 2011; Busaba et al., 2001; Helbing et al., 2019). We therefore assume that a longer lock-up period and a greater fraction of retained shares are associated with higher investor valuations. They should thus be negatively related to the withdrawal probability.

We expect that corporate governance characteristics will be particularly relevant in highly volatile market environments when investors are exposed to greater risks and a higher uncertainty. In this situation, they benefit from reliable signals of firm quality, strong monitoring, and aligned incentives, which reduce the agency costs of adverse selection and moral hazard.

3.2.3 Intermediary characteristics

Helbing et al. (2019) argues that the final decision on the exercising of the withdrawal option resides with the CEO. However, intermediaries may play an important role as well. We account for intermediary characteristics by focusing on VC backing. As firm owners, VCs typically possess strong control rights that give them the option to decide about the exit strategy (e.g. Cumming, 2008).

From the theoretical perspective, the effect of the intermediary characteristics is unclear. On the one hand, high-quality intermediaries may be associated with a higher withdrawal probability for at least two reasons. Intermediaries are repeated players in the capital markets and high-quality intermediaries fear the loss of their reputation. Consequently, they only complete the IPO if the issue is of high quality. If during the price discovery process a firm turns out to be potentially a low-quality issue, they may force it to withdraw because they fear to lose their reputation if they took a low-quality firm public (Dunbar and Foerster, 2008). Another reason for why high-quality intermediaries may be associated with a higher likelihood of a withdrawal is that their firms have other options than a current IPO. Brau et al. (2010) find that VC backed and prestigious underwritten issues have a higher propensity to follow a dual-track strategy. These firms file for an IPO, but they also look for potential acquirers (Aktas et al., 2018). If they succeed in finding an acquirer, their reservation price increases. Even if they do not succeed in finding an acquirer, VCs may withdraw their IPOs more often because VC backed firms have better access to capital in general. Dunbar and Foerster (2008) argue that VC financing makes firms less dependent on the IPO success because it improves their chances to find alternative sources of financing, including another VC round. In addition, as investors that aim at maximizing their financial returns from exits, VCs may decide to postpone the IPO if they believe to achieve greater proceeds in a later exit (Fan and Yamada, 2020).

On the other hand, VCs and prestigious underwriters may provide a signal about the quality of their issues (Cumming et al., 2016; Krigman et al., 2001). This signal reduces adverse selection costs, increases valuations and leads to a lower likelihood of withdrawals. Boeh and Southam (2011), Busaba et al. (2001), and Dunbar and Foerster (2008) discuss this argument for VC investors. In addition, VC investors typically retain a large fraction of their shares and exhibit an active involvement in the monitoring of their firms beyond the IPO (Krishnan et al., 2011). This involvement, which reduces the moral hazard problem and aligns the incentives between VCs and new investors, should lead to a lower withdrawal probability.

Prior empirical literature that deals with the relationship between VC backing and the probability that firms withdraw their offerings provides contradictory findings. Dunbar and Foerster (2008) show a negative relation between withdrawal probability and VC backing, while Boeh and Southam (2011) or Helbing et al. (2019) demonstrate a positive relation. Therefore, we delve deeper into the relationship between the withdrawal probability and the VC financing by investigating whether different VC types are associated with differing withdrawal probabilities.

Besides the VC dummy, which reflects whether the issuer has obtained VC financing or not, we focus on three VC related characteristics. The first is VC syndication, which reflects whether the firm has obtained financing from a single stand-alone VC or from a VC syndicate. VC syndicates may provide a stronger signal than a stand-alone investor (Boeh and Southam, 2011; Brander et al., 2002), which may lead to a higher valuation and a lower withdrawal probability. However, as more investors have their reputational capital at stake and as syndicates may improve the chances to find alternative sources of capital (Tian, 2012), syndicated deals could also be associated with a higher withdrawal probability. Our second VC related variable is VC location. We distinguish between domestic and foreign VCs. Here, the direction of the effect is again uncertain as domestic investors may be associated negatively or positively with the withdrawal probability. On the negative side, local VC investors likely provide better monitoring than foreign investors as they are located close to their portfolio firms. On the positive side, local investors may be beneficial in finding alternative sources of capital in the local market. Our third VC related variable is reputation. All arguments discussed above for the VCs in general should be strengthened for reputable VCs: If VC backing provides a quality signal, the VC's level of reputation may amplify this effect (Boeh and Southam, 2011). In addition, Krishnan et al. (2011) demonstrate that reputable VCs are better monitors than less reputable VCs. If VC financing improves the chances to find alternative sources of financing, the chances should increase with VC reputation. Finally, highly reputable VCs are concerned more about reputational losses than less reputable VCs.

Whether VC backed firms behave differently in less risky and more risky markets is an empirical question, too. On the one hand, if VC financing is a positive signal that reduces the withdrawal probability, then this effect should be stronger in more risky markets. On the other hand, withdrawals of VC backed firms may be more likely in risky markets because the dual-track strategy may prove beneficial in particular when public markets deteriorate. In addition, VCs may be able to provide another round of financing in uncertain markets. These investors do not suffer immediately from a market deterioration because they have a guaranteed access to capital through the commitments of their limited partners. All these arguments that apply to the differences between VC backed and non-VC backed firms in more risky markets should be amplified for syndicated, local, and reputable VCs.

Additionally, we control for underwriter characteristics in this variable category. We use the number of underwriters to check whether larger syndicates provide better services and result in a stronger signal (Corwin and Schultz, 2005). To capture lead underwriter quality, we include lead underwriter market share and Carter-Manaster (CM) rank (Carter and Manaster, 1990; Carter et al., 1998; Loughran and Ritter, 2004).

3.2.4 Other factors

Issuer and issue characteristics

The first three variables in this category are filing size, firm size, and firm age. Firms are easier to value and thus investors face lower risk if filing and firm size are large and if firms are relatively old (Boeh and Southam, 2011; Busaba et al., 2001; Dunbar and Foerster, 2008). In addition, larger issues are associated with higher post-IPO liquidity and are more likely to attract institutional investors who generate information (Busaba et al., 2001). Therefore, greater issues, larger and older firms should be associated with a lower withdrawal probability. We further distinguish between high-tech and low-tech firms. High-tech firms, which typically produce specialized products and are more dependent on human capital specific skills, will be hurt more by negative publicity following a withdrawal decision. This should lead to a lower issuer reservation value and imply a lower probability to withdraw (Busaba et al., 2001; Dunbar and Foerster, 2008).

In order to measure the financial situation of the firm, we add three ratios. The first one is the debt ratio. From a theoretical point of view, the relation between debt ratio and the likelihood of a withdrawal is unclear. On the one hand, a withdrawal may be more likely for firms with higher debt because higher debt ratio is associated with more financial risk (Boeh and Southam, 2011), which may reduce investor valuation. In addition, a higher debt ratio expresses better access to alternative sources of financing (Busaba et al., 2001) and may thus result in a higher reservation price. On the other hand, firms that have high debt ratios may be close to their debt capacity (e.g. Lemmon and Zender, 2010). Due to lack of financing alternatives, their reservation price is lower. From the viewpoint of investors, higher debt level may increase investor valuations because it reduces Jensen's (1986) agency costs of free cash flow (Boeh and Southam, 2011). Our second ratio is the industry book-to-market ratio (BM ratio). Again, it is unclear how BM ratio is related to the withdrawal probability. If a low ratio captures high growth opportunities in the industry, we would expect higher valuations for IPO firms as well and, thus, a lower likelihood of a withdrawal (Dunbar and Foerster, 2008). In contrast, if a low ratio rather captures overvaluation, investors could discover it during the bookbuilding process and bid less, leading to a higher withdrawal probability (Dunbar and Foerster, 2008). As our third ratio, we include asset turnover. We assume that a higher asset turnover is associated with more viable firms, which are easier to value (Busaba et al., 2001). Consequently, it should lead to a lower withdrawal probability.

Finally, we use the binary variable negative news (Helbing et al., 2019). Shi et al. (2016) argue that negative news affect stock market returns. In our context, negative business news could reduce investor valuation. We therefore expect that a withdrawal will be more likely if negative news about the issuer are released.

Market characteristics at and after the filing

Market timing theories consider market characteristics as important determinants of the filing and the withdrawal decisions (Pástor and Veronesi, 2005).

At the time of filing, we consider equity and debt market conditions as well as the competitive situation at the IPO market. Boeh and Dunbar (2014) argue that high equity valuations (we use the current Nasdaq level) reflect high investor demand and that this relation is consistent with irrational and rational explanations: high equity valuations may be related to overoptimistic investors, but also to good investment opportunities (Pástor and Veronesi, 2005). Both explanations suggest that higher valuations lead to a lower withdrawal probability (Boeh and Southam, 2011). We further consider the expected pre-filing stock market volatility, which can be considered as a proxy for the equity risk premium (Pástor and Veronesi, 2005). Increasing risk premium should be associated with a higher withdrawal probability. In addition, Busaba et al. (2015) claim that with increasing volatility the option on the future uncertain public price increases in value. Consequently, we should observe an increase in IPO attempts (see also Boeh and Dunbar, 2014), but an average filing would be of a lower quality, which increases the withdrawal probability. To capture debt market conditions, we employ corporate debt yield spread (BAA-AAA). In periods when corporate debt is expensive, i.e. the yield spread is large, aggregate default risk is high as well so we would assume that withdrawals are more likely due to lower investor valuations (Dunbar and Foerster, 2008). However, at the same time, firms attempting to raise capital in high-spread environments have fewer alternatives so their reservation value decreases. Consequently, they want to pursue an IPO and prevent a withdrawal (Dunbar and Foerster, 2008). Finally, we assume that the likelihood of a withdrawal increases if more filings compete for investor funds around the filing date (Busaba et al., 2001; Dunbar and Foerster, 2008).

We finally add the development of market characteristics after filing. Our first two variables in this category are the Nasdaq return and its standard deviation within the post-filing period. We expect that withdrawals are more common during periods of lower returns (consistent with Boeh and Southam, 2011; Bouis, 2009; Busaba et al., 2001; Edelen and Kadlec, 2005) and higher volatility (Bouis, 2009; Schill, 2004). Besides the changes in equity markets, we consider changes in debt markets. The increase in the corporate yield spread and more severe lending standards (Bergbrant et al., 2017) are associated with higher aggregate default risks, which could lead to lower investor valuations and thus a higher withdrawal probability. However, when debt becomes more expensive and more difficult to obtain, firms may want to prevent withdrawals as they have fewer alternatives to raise capital (Dunbar and Foerster, 2008).

3.3 Data

3.3.1 Filings, completed IPOs, and withdrawals

The sample consists of US domestic IPO filings between 1997 and 2014 drawn from Thomson Reuters Securities Data Company (SDC) Platinum. Following IPO literature (e.g. Busaba et al., 2001; Dunbar and Foerster, 2008; Ritter and Welch, 2002) we remove American depositary receipts, convertible issues, unit offerings, closed-end funds, REITs, limited partnerships, small best effort offers, SPACs, and issuers that are not seeking a listing on NYSE, NASDAQ or any other American exchange. We also exclude firms from the financial industry. After these steps, we end up with 5,095 filings. We only consider the first attempt of firms to go public. The reason is that firms might have different aims with the first and the second attempts and investors may have different perceptions as well (Boeh and Dunbar, 2013; Chen et al., 2017; Dunbar and Foerster, 2008). Consequently, the probability of a withdrawal will be different for the first-time and the second-time IPOs. This reduces the sample size to 4,939 firms. We then check filing dates, IPO dates and withdrawal dates against EDGAR filings and remove observations for which one of these dates differ. Our final sample then consists of 3,438 firms. The withdrawal rates and the characteristics of our sample firms are within the range of prior studies, so we believe that that our sample is representative.

Table 3.1 shows the number of filings in our sample by year. The highest number of filings was reached at the peak of the dotcom bubble in 1999 and 2000 with 494 and 455 filings, the lowest number occurred during the financial crisis with 56 filings in 2008 and 54 filings in 2009.[5] Table 3.2 also provides descriptive statistics about the number and percentage of filings in our sample that were completed or withdrawn, respectively. The withdrawal rate is 27.28% on average, varying between 6.13% in 2013 and 60.71% in 2008.

Table 3.1: Withdrawn and completed offerings by year

Filing year	Number of filings	Number completed	Number withddrawn	Percentage completed	Percentage withdrawn
1997	437	373	64	85.35	14.65
1998	312	219	93	70.19	29.81
1999	494	427	67	86.44	13.56
2000	455	237	218	52.09	47.91
2001	72	44	28	61.11	38.89
2002	87	57	30	65.52	34.48
2003	68	63	5	92.65	7.35
2004	178	131	47	73.60	26.40
2005	154	125	29	81.17	18.83
2006	181	129	52	71.27	28.73
2007	184	95	89	51.63	48.37
2008	56	22	34	39.29	60.71
2009	54	37	17	68.52	31.48
2010	137	89	48	64.96	35.04
2011	150	102	48	68.00	32.00
2012	82	57	25	69.51	30.49
2013	163	153	10	93.87	6.13
2014	174	140	34	80.46	19.54
Total	3,438	2,500	938	72.72	27.28

Note: Number of filings, completed issues and withdrawals from 1997 to 2014. The data stems from Thomson Reuters SDC Platinum.

[5] In the empirical analysis below, we provide a robustness check where we exclude these years to see whether issues in these periods drive our results.

3.3.2 Descriptive statistics

We provide a list of all variables, which contains their detailed descriptions and sources, in Table B.1. Furthermore, we show descriptive statistics for the full sample and for the subsamples of completed and withdrawn offerings in Table 3.2 and perform a univariate t-test of our predictions. Table B.2 shows the correlations among all pairs of variables.

We first focus on corporate governance characteristics. We hand-collect board characteristics (size, experience, share of female members, CEO duality) and lock-up period information from EDGAR filings. As most firms in the data set have a lock-up period of 180 days, we construct a dummy that equals one if the lock-up period is larger than 180 days and zero otherwise. Data on retained shares is from EDGAR filings, S&P Capital IQ, and Bloomberg. Unfortunately, for approximately half of the withdrawn issues, this variable is missing, mostly because the information on retained shares was not published in the preliminary prospectus (and also not available from S&P Capital IQ or Bloomberg). We therefore do not include this variable in our main regressions but consider its effect in a separate regression. In Table 3.2 we observe only minor differences in corporate governance characteristics between completed and withdrawn offerings. The t-test for the differences between these two subsamples shows that none of these differences is statistically significant. These results do not support the view that high-quality corporate governance provides a signal that reduces agency costs during the post-filing period.

In our next category, we consider intermediary characteristics. The information on VC backing and VC location is gathered from EDGAR filings, S&P Capital IQ, and Crunchbase. For the measure of VC reputation, we rely on the Lee-Pollock-Jin VC Reputation Index (Lee et al., 2011). We do not observe a significant difference between the share of VC backed firms in the subsample of completed and withdrawn offerings in Table 3.2. However, we find significant differences in the share of syndicated deals and deals by reputable VCs, which are more common in the subsample of completed than in the subsample of withdrawn deals. These differences are consistent with signaling effects through VC syndication and VC reputation. The data on underwriter market share stems from Thomson Reuters Securities Data Company (SDC) Platinum. Annual data on CM rank comes from Jay Ritter's website. Our descriptive statistics suggest that larger underwriter syndicates are associated with a higher probability of completion. This is consistent with the conjecture that larger syndicates tend to deliver more valuable services (Corwin and Schultz, 2005). Underwriter quality (CM rank and market share) does not differ in both subsamples.

To obtain issuer and issue characteristics, we rely on Compustat and Thomson Reuters Securities Data Company (SDC) Platinum. Because data on firms that withdraw their offerings is scarce there, we hand-collect further information from EDGAR filings. We

Table 3.2: Descriptive statistics of completed and withdrawn IPOs

	Full sample		Completed		Withdrawn		
	Mean	Obs	Mean	Obs	Mean	Obs	p-values
Corporate governance characteristics							
Board size	6.45	3,436	6.48	2,499	6.35	937	0.13
Board experience	50.51	3,426	50.61	2,493	50.24	933	0.11
Female board members	0.04	3,317	0.05	2,416	0.04	901	0.15
CEO duality	0.54	3,028	0.54	2,202	0.53	826	0.53
Lock-up	0.04	3,438	0.04	2,500	0.04	938	0.78
Retained shares	0.69	2,858	0.69	2,415	0.69	443	0.94
Intermediary characteristics							
VC backing	0.43	3,399	0.42	2,477	0.44	922	0.23
VC syndication$^+$	0.83	1,401	0.84	1,007	0.79	394	0.02**
VC US$^+$	0.69	1,410	0.70	1,013	0.68	397	0.63
VC reputation$^+$	0.43	1,419	0.44	1,019	0.39	400	0.07*
Number of underwriters	1.58	3,416	1.64	2,486	1.44	930	0.00***
Underwriter market share	0.06	3,289	0.06	2,382	0.06	907	0.25
CM Rank	7.79	3,368	7.82	2,460	7.71	908	0.11
Issuer and issue characteristics							
Filing size (bn USD)	0.12	3,319	0.12	2,393	0.12	926	0.34
Firm size (bn USD)	0.50	3,405	0.53	2,494	0.42	911	0.13
Firm age	13.50	3,233	14.39	2,441	10.76	792	0.00***
High-tech	0.58	3,416	0.58	2,486	0.58	930	0.95
Debt ratio	0.82	3,375	0.79	2,486	0.92	889	0.00***
BM ratio	0.34	3,402	0.33	2,481	0.37	921	0.00***
Asset turnover	1.06	3,318	1.06	2,468	1.08	850	0.58
Negative news	0.27	3,438	0.28	2,500	0.26	938	0.39
Market characteristics at filing							
Nasdaq	21.51	3,438	20.87	2,500	23.24	938	0.00***
Implied volatility	0.04	3,438	0.03	2,500	0.05	938	0.01**
Yield spread	0.83	3,438	0.82	2,500	0.85	938	0.00***
Number of filings	0.11	3,436	0.11	2,500	0.11	936	0.47
Market characteristics after filing							
Nasdaq return	0.04	3,438	0.07	2,500	-0.04	938	0.00***
Nasdaq std. dev.	0.69	3,436	0.62	2,500	0.88	936	0.00***
Yield spread change	0.01	3,436	0.00	2,500	0.03	936	0.00***
Lending standard change	6.11	3,438	2.85	2,500	14.77	938	0.00***

$^+$ VC-backed sample only

Note: Descriptive statistics for the full sample and the subsamples of completed and withdrawn offerings. The table reports sample means and number of observations. Additionally, it shows the p-values of a t-test for the difference between the subsamples of completed and withdrawn offerings. Variable descriptions and data sources are summarized in Table B.1. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

complement these data sources with S&P Capital IQ, Jay Ritter's website, Kenneth R. French's Data Library, and LexisNexis. Consistently with Boeh and Southam (2011), we find that older firms tend to withdraw their offerings less often than younger firms. Withdrawn issues have significantly higher debt ratios. This is consistent with the argumentation that higher debt leads to more financial risk and lower investor valuations (Boeh and Southam, 2011) and/or that these issues have a better access to alternative sources of financing so that the issuer reservation value is higher (Busaba et al., 2001). BM ratio is larger for withdrawn than for completed deals. This finding provides support for the conjecture that high BM ratio captures low growth opportunities in the industry (Dunbar and Foerster, 2008). We do not observe significant differences in firm and filing size, the proportion of high-tech firms, asset turnover or the frequency of negative news between withdrawn and completed issues.

As to the market characteristics at filing, we obtain the Nasdaq composite index from FRED. In contrast to our expectations, the descriptive statistics show that the Nasdaq composite index at the filing date is higher in the subsample of withdrawn offerings than in the subsample of completed IPOs. A possible explanation is that when the index is at a high level, it is likely that it will grow at a slower pace or even decrease. Table B.2 indeed shows a negative correlation between the Nasdaq level at the filing date and the post-filing Nasdaq return. The pre-filing increase in expected volatility is based on the VXO measure, which we collect from CBOE. In line with Boeh and Dunbar (2014) and Busaba et al. (2015), the expected volatility is positively associated with the likelihood of a withdrawal. The BAA-AAA yield spread, which is also taken from FRED, is higher in the subsample of withdrawn offerings supporting the view that higher yields increase the default risks and thus lead to lower investor valuations (Dunbar and Foerster, 2008). We retrieve the number of filings around the filing date from Thomson Reuters Securities Data Company (SDC) Platinum. We do not find any differences in the number of concurrent filings between completed and withdrawn issues.

The differences in characteristics capturing post-filing market developments between the withdrawn and completed offerings are all significant. To measure these characteristics, we consider the period between the filing date and 30 days after this date. We rely on FRED data to obtain the Nasdaq return and its standard deviation. The descriptive statistics confirm the findings from prior studies. Withdrawals are more likely when post-filing stock market returns are low (see e.g. Busaba et al., 2001). In addition, higher uncertainty, reflected in higher Nasdaq volatility, is associated with a higher probability of a withdrawal (e.g. Bouis, 2009; Schill, 2004). The yield spread change, based on FRED data, is significantly larger in the subsample of withdrawn offerings than in the subsample of completed IPOs. We proxy for the change in lending standards by using the

Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS). The results suggest that lending standards are tightened more in the subsample of withdrawn offerings than in the subsample of completed offerings. Our two debt-related variables support the conjecture that outside investors fear defaults and reduce their valuations when debt becomes more expensive (Dunbar and Foerster, 2008) and lending standards tighten.

3.4 Methodology

3.4.1 Timing

When assessing the market characteristics after filing, we face the challenge to choose an appropriate time horizon. The choice of this time horizon is important not only for the construction of the variables that capture post-filing market developments. It is also relevant when we create our indicator of high post-filing market volatility.

Ideally, we would consider the market development between the filing date and the date on which the firm decides whether to go public or to withdraw. For IPOs, we can proxy the latter date with the IPO date. The IPO date follows soon after the end of the bookbuilding process, when the firm has solicited investor demand for its shares and, based on this information, taken the final decision to complete the IPO (Busaba, 2006). For withdrawn issues, the situation is more complicated. From official sources (SEC, S&P Capital IQ, etc.), we can obtain the date when the firm submits the official Request to Withdraw Registration Statement to the SEC. However, firms that decide to withdraw their offerings may submit this document with a substantial time lag. Therefore, we have searched for press releases or news about the announcement of the withdrawal decision (using Factiva and Google) to obtain a more realistic date. Unfortunately, we have found such information for less than 50% of the withdrawn issues from our sample only. Among these cases, the date we have found almost always (more than 90%) corresponds to the official SEC date.

We consider three alternative approaches to deal with this issue. In the first approach, we employ a fixed period that has the same length for all sample firms (IPOs and withdrawals). We set the length at 30 days after the filing date, in line with Busaba et al. (2001). The second approach is to apply the period between the filing and the outcome for IPOs and use an estimate for the withdrawn issues. As an estimate for the period length between the filing and the withdrawal, we employ a fixed number of 118 days, which corresponds to the average period length between the filing and the IPO in our sample. In the third approach, we rely on the outcome dates reported by the SEC, i.e. the realized

IPO date and the date of the official Request to Withdraw Registration Statement to the SEC; this procedure has been applied, among others, by Boeh and Dunbar (2013). In our main specifications, we consider the first approach and we provide a robustness check with the two other approaches.

3.4.2 Baseline regressions

In our baseline regressions, we examine how each of the five variable categories described above is related to the withdrawal probability in a multivariate setting. As our dependent variable withdrawal is binary, we run a non-linear probit model, which was used in prior studies on withdrawals (Busaba et al., 2001; Helbing et al., 2019). Our dependent variable equals one for withdrawn issues and zero for completed issues.

We calculate marginal effects at the mean (MEM) because the coefficients in non-linear models do not have a meaningful interpretation. We run several specifications to check the robustness of our results. The main aim is to examine whether our variables of interest change when we add further variables or alter variable definitions. We start by five specifications, in which we include the five variable categories one after the other. We first include issuer and issue controls. We then add corporate governance characteristics to examine whether these characteristics are related to the withdrawal probability while controlling for other issuer and issue variables. In the next specification, we add intermediary characteristics to the second specification and focus on the effect of the VC dummy. As next, we additionally include our fourth category of variables: market characteristics at filing. Finally, we run a full specification, in which we examine all five categories of characteristics. We thus add variables that reflect the market development 30 days after the filing date. We then alter the definition of the variables from the last category by employing alternative time horizons for their measurement. In a separate subsection, we add VC related characteristics to the full specification.

In all specifications, we control for major regulatory changes, which may have affected the likelihood of withdrawals and may be correlated to some of our variables of interest. Without these controls, our results could suffer from an omitted variable bias and reflect spurious correlations. We include period-related dummies (for a detailed description see Table B.1) that consider the following regulatory events (in line with Busaba et al. (2015)):

- The amendment of Rule 477 and the adoption of the safe harbor Rule 155 in March 2001, which increase issuers' possibilities to withdraw an offering and switch to a private offering.
- Sarbanes-Oxley Act (SOX) in July 2002, which essentially increases the costs of being public, particularly for young firms.

- JOBS Act in April 2012, which allows emerging growth firms to solicit indications from institutional investors prior to the official filing (Section 105(c)). These firms also get the opportunity to submit draft registration statements to the SEC for confidential review, which is not subject to filing fees and does not require consent of auditors or other experts (Section 106(a)). "Testing the Waters" prior to official filing may reduce the withdrawal probability because some of the withdrawal candidates might decide against the filing when they are disappointed by the test results.

We also consider the effects of the recent financial crisis.

3.4.3 Regressions with interaction terms

To examine whether corporate governance and VC characteristics turn to be more relevant when markets become highly volatile, we interact corporate governance and VC characteristics with an indicator of highly volatile stock markets. To construct the indicator of highly volatile stock markets, we divide the sample into three subsamples, depending on the 30-day post-filing Nasdaq return volatility. We define a binary variable high volatility that equals one for those issues for which the Nasdaq return volatility in the post-filing period is in the top tertile. The indicator equals zero if the volatility is in the bottom or the medium tertile. For the interaction effect analysis, we rely on the full specification, which includes all five categories of variables. We add the interaction terms between the indicator of highly volatile markets and the corporate governance characteristics one-by-one. For our six corporate governance variables we thus run six different specifications. For the VC characteristics, we run four different specifications, by adding interaction terms between the high volatility indicator and the VC dummy as well as the dummies for VC syndication, US VCs and VCs with high reputation.

If we want to understand whether during highly volatile markets good corporate governance and VC financing (or some types thereof) become more important to prevent a withdrawal, we have to quantify the marginal effects on the interaction terms because the coefficients in non-linear models do not have a meaningful interpretation. Ai and Norton (2003) show that the interaction effect in non-linear models is conditional on the values of all independent variables. It is even possible that the sign of the interaction effect changes when we vary the value of other independent variables. This has also direct implications for testing the statistical significance of the interaction effect because it cannot be tested with a simple t-test on the estimated interaction effect as in a linear model. As the coefficients on the interaction cannot be interpreted in a meaningful way and their test statistics are not correct, we apply the method suggested by Ai and Norton (2003) and Norton et al. (2004) to calculate the magnitudes of the interaction effects and to test their significance. In their method the interaction effects are calculated by computing cross-partial derivatives.

As the marginal effect is different for each observation, the most straightforward way to report these effects and their corresponding z-statistic (for continuous variables) is to plot them against the predicted probabilities. Besides these figures, we also provide summary statistics of the marginal effects in a table to obtain an overview. For continuous variables, we exhibit the MEMs and their corresponding z-statistics. In order to check whether the binary variables are more significant in highly volatile than in less volatile markets, we evaluate their effects (of a change between zero and one) separately at both values of the high volatility indicator.

3.5 Results

3.5.1 Baseline results

We exhibit the outcomes from the multivariate analysis in Table 3.3. The results in specification (1) are mostly consistent with the findings from the univariate analysis. The only exception is the high-tech dummy, but this variable turns insignificant in later specifications. We add our corporate governance characteristics in specification (2) and observe that, with one exception, none of them is statistically significant. Only the lock-up dummy has a statistically significant effect. This effect is negative, suggesting that longer lock-up periods tend to serve as a signaling device to outside investors. This result is consistent with findings in Boeh and Southam (2011) and Helbing et al. (2019). However, it must be interpreted with caution because the variation in this variable is very low. Only 3.7% of the sample observations have a lock-up period that exceeds 180 days (for which the binary variable equals 1). In specification (3), we find a positive effect of the VC backing, suggesting that VCs are associated with a higher withdrawal probability. VC dummy is significant only in this specification. Once we control for market characteristics, VC dummy loses its significance. For the underwriter characteristics, the results are consistent with the univariate results. The marginal effects of the issue, issuer, and corporate governance characteristics do not change much compared to the first and second specifications.

In specification (4), in line with the descriptive statistics, the Nasdaq level and the expected volatility are positively associated with the likelihood of a withdrawal. Yield spread has a positive and statistically significant effect, too. The number of filings, which was not significantly different in the subsample of withdrawn and completed issues, has a negative and significant effect, which contradicts our expectations. This finding suggests that in hot markets, the completion rates are higher. The signs and statistical significance levels of the other effects do not change much compared to the first three specifications.

In specification (5), our full specification, we find a highly statistically significant negative effect of the post-filing Nasdaq return, which is consistent with the univariate results. The development on the debt market seems to be important as well. An increase in the yield spread is related to a higher withdrawal probability, which supports the view that increased default risk reduces investor appetite for new firms (Dunbar and Foerster, 2008). The tightening in lending standards is also related to the higher probability to withdraw. All in all, we conclude that negative equity and debt market developments 30 days after the filing, reflected in low levels of market returns, yield spread and tightening of lending standards, are associated with higher withdrawal rates. Withdrawals are more likely when market conditions become more challenging. This is consistent with what issuers themselves claim (Boeh and Dunbar, 2013).

In specifications (6) and (7), we perform regressions with an alternative definition of the post-filing period. For the IPOs, we consider the development in the period between the filing and the listing. For the withdrawals, we employ first a fixed period, common to all withdrawals (average of the filing-to-outcome period for IPOs) and, second, a period between the filing and the date of the official withdrawal letter for each individual observation. The results from columns (5) and (6) are similar, in column (7) the Nasdaq standard deviation turns positive suggesting that withdrawals are more likely in volatile environments. Lending standard change becomes insignificant. The choice of the post-filing period does not have any important effects on the variables from the other four categories. Importantly, the choice of the post-filing period does not affect the results on our variables of interest (corporate governance characteristics and VC backing). Finally, specification (8) adds the variable retained shares, which we have not included in the other specifications because it reduces the number of observations, particularly for the withdrawn issues, due to its missing values. The effect of this corporate governance variable is not statistically significant.

We perform two robustness checks for our main results. First, in order to examine whether individual industries drive our results, we perform the full specification while leaving out observations from one industry.[6] We run this regression twelve times for each of the Fama French 12 industries. As another robustness check, we exclude the filings during the dotcom bubble and the recent financial crisis. We display the results from these robustness checks in Table B.3. We see only minor differences to the main results. Most of the variables keep their sign and significance.

[6] It was not possible to perform an industry specific analysis since our models do not converge for several individual industries due to the small number of observations.

Table 3.3: Probit analysis of the decision to withdraw an offering, baseline specifications

	(1) MEM	(2) MEM	(3) MEM	(4) MEM	(5) MEM	(6) MEM	(7) MEM	(8) MEM
Corporate governance characteristics								
Board size		0.0007	0.0000	0.0007	-0.0003	0.0008	0.0025	-0.0013
		[0.0039]	[0.0041]	[0.0039]	[0.0040]	[0.0041]	[0.0036]	[0.0027]
Board experience		0.0002	-0.0005	-0.0003	0.0001	-0.0014	-0.0015	0.0007
		[0.0012]	[0.0010]	[0.0009]	[0.0008]	[0.0014]	[0.0010]	[0.0008]
Female board members		-0.0009	0.0131	0.0311	0.0073	0.0075	0.0108	0.0918*
		[0.0782]	[0.0796]	[0.0792]	[0.0766]	[0.0735]	[0.0735]	[0.0472]
CEO duality		-0.0163	-0.0178	-0.011	-0.0137	-0.0111	-0.0142	-0.0081
		[0.0197]	[0.0214]	[0.0219]	[0.0246]	[0.0237]	[0.0240]	[0.0157]
Lock-up		-0.1895***	-0.1803***	-0.1670***	-0.1602***	-0.1382**	-0.1563***	-0.0942***
		[0.0353]	[0.0514]	[0.0503]	[0.0495]	[0.0558]	[0.0504]	[0.0304]
Retained shares								-0.0156
								[0.0261]
Intermediary characteristics								
VC backing			0.0277**	0.0023	0.0115	0.0072	-0.0055	0.0078
			[0.0133]	[0.0132]	[0.0126]	[0.0141]	[0.0191]	[0.0150]
Number of underwriters			-0.0166**	-0.0283***	-0.0134	-0.0193*	-0.0132*	0.0037
			[0.0071]	[0.0082]	[0.0108]	[0.0101]	[0.0079]	[0.0081]
Underwriter market share			-0.1126	-0.1944	-0.2593	-0.3124*	-0.1752	-0.3545*
			[0.2082]	[0.1919]	[0.1749]	[0.1772]	[0.2439]	[0.1839]
CM Rank			-0.0011	-0.0044	-0.0058	-0.0026	-0.0020	-0.0120**
			[0.0073]	[0.0069]	[0.0074]	[0.0071]	[0.0087]	[0.0057]
Issuer and issue characteristics								
Filing size (bn USD)	-0.0772	-0.1091	-0.0977	-0.1672	-0.1350	-0.1837	-0.1533	-0.1098
	[0.1008]	[0.0996]	[0.1051]	[0.1149]	[0.1127]	[0.1258]	[0.1373]	[0.0830]
Firm size (bn USD)	-0.0103	-0.0033	0.0128	0.0231	0.0076	0.0274	0.0010	-0.0228
	[0.0298]	[0.0286]	[0.0325]	[0.0335]	[0.0360]	[0.0344]	[0.0378]	[0.0234]
Firm age	-0.0284**	-0.0279***	-0.0258**	-0.0268***	-0.0253***	-0.0235**	-0.0274**	-0.0189
	[0.0115]	[0.0106]	[0.0114]	[0.0102]	[0.0097]	[0.0099]	[0.0132]	[0.0121]
High-tech	0.0398**	0.0432**	0.0343	-0.0087	-0.0145	-0.0109	0.0073	-0.0152
	[0.0161]	[0.0187]	[0.0234]	[0.0229]	[0.0211]	[0.0182]	[0.0165]	[0.0173]
Debt ratio	0.0255***	0.0313**	0.0365**	0.0332*	0.0351**	0.0380**	0.0373**	0.0208**
	[0.0097]	[0.0148]	[0.0168]	[0.0173]	[0.0161]	[0.0172]	[0.0178]	[0.0095]

BM ratio	0.2734***	0.2802***	0.3130***	0.2023**	0.2106**	0.1715**	0.1297	0.1188**
	[0.1014]	[0.1011]	[0.1048]	[0.0960]	[0.0880]	[0.0798]	[0.0863]	[0.0580]
Asset turnover	0.0051	0.0056	0.0072	0.0114	0.0100*	0.0063	-0.0007	0.0051
	[0.0062]	[0.0065]	[0.0064]	[0.0071]	[0.0055]	[0.0056]	[0.0049]	[0.0056]
Negative news	0.0150	0.0079	0.0084	0.0093	0.0136	0.0061	0.0152	0.0076
	[0.0129]	[0.0093]	[0.0112]	[0.0112]	[0.0110]	[0.0116]	[0.0156]	[0.0111]
Market characteristics at filing								
Nasdaq				0.0103***	0.0059***	0.0026	-0.0070***	0.0024**
				[0.0013]	[0.0017]	[0.0019]	[0.0017]	[0.0010]
Implied volatility				0.1309***	0.0716**	0.0896***	0.0808***	-0.0109
				[0.0210]	[0.0298]	[0.0271]	[0.0252]	[0.0303]
Yield spread				0.0397	0.0215	0.1852**	0.2344***	-0.0388
				[0.0345]	[0.0519]	[0.0731]	[0.0703]	[0.0287]
Number of filings				-0.7272***	-0.5731***	-0.2351**	-0.2832***	0.0128
				[0.1682]	[0.0986]	[0.1030]	[0.1073]	[0.1424]
Market characteristics after filing								
Nasdaq return					-0.0697**	-0.2623***	-0.3378***	-0.0164
					[0.0333]	[0.0453]	[0.0300]	[0.0213]
Nasdaq std. dev.					-0.0289	0.0006	0.1038***	-0.0180
					[0.0308]	[0.0110]	[0.0116]	[0.0253]
Yield spread change					0.3115***	0.2887***	0.3131***	0.1998***
					[0.1059]	[0.0826]	[0.0655]	[0.0645]
Lending standard change					0.0038***	0.0023***	-0.0003	0.0024***
					[0.0005]	[0.0006]	[0.0006]	[0.0004]
Regulation 1	-0.0328	-0.0386	-0.0447	-0.0213	-0.0287	-0.0573	-0.0428	-0.0070
	[0.0359]	[0.0390]	[0.0415]	[0.0556]	[0.0532]	[0.0552]	[0.0492]	[0.0371]
Regulation 2	0.0651	0.0181	0.0459	0.0117	-0.2997	-0.4559***	-0.3562**	-0.0756
	[0.1447]	[0.1567]	[0.1804]	[0.1582]	[0.1921]	[0.1266]	[0.1432]	[0.1380]
Regulation 3	-0.1284***	-0.1401***	-0.1389***	-0.2960***	-0.1711***	-0.1243***	0.0331	-0.1244***
	[0.0253]	[0.0219]	[0.0267]	[0.0275]	[0.0224]	[0.0269]	[0.0355]	[0.0209]
Obs.	2,996	2,588	2,470	2,468	2,468	2,461	2,460	2,145
Pseudo R-Squared	0.0254	0.0301	0.0346	0.0693	0.1029	0.1288	0.2238	0.0973

Note: Probit model in which the dependent variable equals one for withdrawn offerings and zero for completed offerings. Descriptions of the variables and data sources are depicted in Table B.1. Marginal effects at the mean (MEM) are reported. Variable categories are added one by one in specifications (1) – (5). Specifications (6) and (7) employ alternative definitions of the post-filing period. Specification (8) adds the variable retained shares. Pseudo R-squared corresponds to McFadden's pseudo R-squared. Standard errors clustered at industry level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

3.5.2 Corporate governance characteristics in risky market environments

Our results so far have not supported the view that corporate governance characteristics provide signals that affect the withdrawal probability. In this section, we examine whether signaling through corporate governance characteristics turns to be relevant when markets become highly volatile and investors thus face higher risks and uncertainties.

To this aim, we interact our high volatility indicator with corporate governance characteristics and show an overview of the marginal effects on these interaction terms in Table 3.4. (For a detailed overview, we exhibit the coefficients on all variables included in these models in Table B.4.) Panel A covers the continuous corporate governance variables: board size, board experience, female board members, and retained shares. The results suggest that larger and more experienced boards are associated with a lower withdrawal probability in highly volatile environments. Both interaction terms are highly statistically significant at the 1% level. These findings support the view that signaling through good corporate governance, which does not seem to be important on average, turns relevant when investors face greater risks and uncertainties. Larger and more experienced boards may signal good quality of the firms and mitigate agency costs resulting from adverse selection. They may also provide a signal of good monitoring that reduces agency costs of moral hazard.

Boards with a higher share of females are associated with a higher withdrawal probability in volatile markets. This effect is significant at the 10% level only. We thus do not find support for our prediction that more diverse teams would provide a better monitoring and thus deliver a positive signal towards the investors. A potential explanation for this finding could be that more heterogeneous boards are more difficult to manage and less likely to reach an agreement when markets become more volatile. Finally, the sign of the interaction term on retained shares is negative, in line with the expectation that incentive alignment signals are more important in highly volatile markets, but the effect is not statistically significant. Figures B.1 - B.4 in the appendix exhibit the marginal effects in detail.[7]

Table 3.4, Panel B exhibits the interaction effects between the high volatility indicator and the two corporate governance dummy variables CEO duality and lock-up. For the CEO duality, we observe a positive sign under volatile markets, which is in line with the

[7] All figures show a similar pattern, where the interaction effect is nearly zero and insignificant for predicted probabilities at y=0 and reaches its maximum or minimum value at about y=0.5. This is reasonable, since if the predicted probability of a withdrawal is y=0, which means that the outcome is clear, then the influence of the considered variable is negligible in volatile market conditions. In contrast, the considered variable is important volatile market conditions if the predicted probability is around y=0.5 and the outcome is not clear due to other covariates.

signaling effect, but this effect is not statistically significant. The only corporate governance variable, which has been significant in the baseline regressions, is the lock-up dummy. In line with the signaling theories, we have observed that longer lock-up periods are associated with lower withdrawal probabilities. The results in Table 3.4, Panel B suggest that, surprisingly, this effect is driven by observations in less volatile market environments. As indicated above, the lock-up results must be interpreted carefully because the variance of this variable is very low.

To sum up, we find that corporate governance characteristics, such as larger and more experienced boards, which were insignificant in the baseline regressions, are highly statistically significant in highly volatile environments. We thus conclude that signaling through high-quality corporate governance seems to matter in risky and uncertain market environments.

Table 3.4: Interactions of corporate governance characteristics with high volatility indicator

Panel A: Continuous variables						
		Obs	Mean	Std Dev	Min	Max
Board size	Interaction term	2,468	-0.02***	0.01	-0.03	0.00
	Z-value	2,468	-2.63	0.49	-4.78	-0.70
Board experience	Interaction term	2,468	-0.01***	0.00	-0.01	0.00
	Z-value	2,468	-3.60	0.73	-5.52	-0.77
Female board members	Interaction term	2,468	0.22*	0.08	0.04	0.34
	Z-value	2,468	1.74	0.11	0.71	2.04
Retained shares	Interaction term	2,145	-0.12	0.06	-0.27	0.00
	Z-value	2,145	-1.38	0.20	-2.25	-0.48

Panel B: Dummy variables		
	at high volatility=0	at high volatility=1
CEO duality	-0.0432	0.0281
	[0.0377]	[0.0190]
Lock-up	-0.2081***	0.0592
	[0.0243]	[0.0737]

Note: In Panel A, summary statistics of the marginal effects of the interaction terms from the models in Table B.4 are reported for all continuous variables. Marginal effects and their z-statistics are calculated using the method by Ai and Norton (2003) and Norton et al. (2004). Details on marginal effects and their z-statistics are depicted in Figures B.1 - B.4. In Panel B, we evaluate the effect of the dummy variables at both values of the high volatility variable. Descriptions of the variables and data sources are depicted in Table B.1. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

3.5.3 VCs and withdrawals

The role of VCs in IPO withdrawals has been discussed controversially. On the one hand, VCs could provide a signal of a higher quality and a better monitoring and thus increase investor valuations. On the other hand, they may raise firm's reservation value by following a dual-track strategy or by postponing the IPO. In our full specification from Table 3.3, the VC dummy was insignificant.

In a more detailed analysis, we want to delve into the effects of VC characteristics to understand whether a particular outcome is associated with a particular VC type more often. Table 3.5 exhibits the MEMs from baseline regressions, which are based on the full specification from Table 3.3, but consider VC characteristics instead of a VC dummy. In each of the three specifications, we include two VC related dummies. The first dummy equals one for non-VC backed firms and zero otherwise. The second dummy equals one for VC syndication (specification (1)), for US VCs (specification (2)), and for reputable VCs (specification (3)), respectively. We thus compare syndicated to non-syndicated deals, deals of US VCs to deals of foreign VCs and deals of reputable VCs to deals of less reputable VCs, respectively.

The negative marginal effect on the VC syndication dummy in specification (1) suggests that firms backed by a VC syndicate face a lower withdrawal probability than firms backed by stand-alone VCs. This result is in line with the signaling and value-adding effects through multiple VCs compared to stand-alone VCs (Boeh and Southam, 2011; Brander et al., 2002). From specification (2) we learn that US VCs tend to be associated with a lower withdrawal probability compared to non-US VCs. This finding is consistent with the signaling theories because domestic VCs may signal to the investors that the company will be more strongly monitored than with more distant foreign VCs. The effect of reputation is insignificant. A possible explanation is that two effects work in the opposite direction. On the one hand, firms backed by reputable VCs may more often postpone their IPOs. They also could be associated more frequently with a dual-track strategy, which increases their reservation value and, in turn, the probability of their withdrawal. On the other hand, reputable VCs may provide a stronger signal to potential investors, which raises investor valuations and should lead to a lower withdrawal probability.

Table 3.5: Probit analysis of the decision to withdraw an offering, full specification with VC characteristics

	(1) MEM	(2) MEM	(3) MEM
Corporate governance characteristics			
Board size	-0.0010	-0.0016	-0.0012
	[0.0042]	[0.0042]	[0.0042]
Board experience	0.0000	0.0000	0.0001
	[0.0009]	[0.0009]	[0.0009]
Female board members	0.0047	0.0063	0.0031
	[0.0779]	[0.0775]	[0.0777]
CEO duality	-0.0139	-0.0122	-0.0141
	[0.0234]	[0.0246]	[0.0245]
Lock up	-0.1548***	-0.1591***	-0.1541***
	[0.0510]	[0.0521]	[0.0512]
Intermediary characteristics			
Non-VC	-0.0555*	-0.0494**	-0.0175
	[0.0286]	[0.0220]	[0.0123]
VC syndication	-0.0520*		
	[0.0310]		
VC US		-0.0467*	
		[0.0260]	
VC reputation			-0.0114
			[0.0162]
Number of underwriters	-0.0128	-0.0131	-0.0126
	[0.0114]	[0.0113]	[0.0112]
Underwriter market share	-0.2493	-0.2553	-0.2613
	[0.1768]	[0.1784]	[0.1753]
CM Rank	-0.0062	-0.0062	-0.0054
	[0.0074]	[0.0074]	[0.0074]
Issuer and issue characteristics			
Filing size (bn USD)	-0.1385	-0.1339	-0.1314
	[0.1147]	[0.1131]	[0.1127]
Firm size (bn USD)	0.0110	0.0121	0.0087
	[0.0365]	[0.0364]	[0.0361]
Firm age	-0.0242**	-0.0245**	-0.0244**
	[0.0097]	[0.0097]	[0.0097]
High-tech	-0.0112	-0.0143	-0.0116
	[0.0206]	[0.0207]	[0.0211]
Debt ratio	0.0323*	0.0338*	0.0328*
	[0.0181]	[0.0186]	[0.0183]
BM ratio	0.2065**	0.2128**	0.2139**
	[0.0870]	[0.0876]	[0.0883]
Asset turnover	0.0104**	0.0115**	0.0103*
	[0.0053]	[0.0053]	[0.0053]
Negative news	0.0098	0.0079	0.0103
	[0.0117]	[0.0111]	[0.0109]
Market characteristics at filing			
Nasdaq	0.0058***	0.0055***	0.0056***
	[0.0016]	[0.0016]	[0.0016]
Implied volatility	0.0741**	0.0704**	0.0718**
	[0.0297]	[0.0330]	[0.0306]
Yield spread	0.0132	0.0111	0.0169
	[0.0512]	[0.0515]	[0.0515]

Number of filings	-0.6028***	-0.5991***	-0.5758***
	[0.1020]	[0.0950]	[0.0958]
Market characteristics after filing			
Nasdaq return	-0.0729**	-0.0721**	-0.0696**
	[0.0323]	[0.0304]	[0.0322]
Nasdaq std. dev.	-0.0249	-0.0228	-0.0245
	[0.0296]	[0.0304]	[0.0300]
Yield spread change	0.3044***	0.3152***	0.3135***
	[0.1010]	[0.1024]	[0.1033]
Lending standard change	0.0038***	0.0038***	0.0038***
	[0.0005]	[0.0005]	[0.0005]
Regulation 1	-0.0260	-0.0273	-0.0269
	[0.0526]	[0.0512]	[0.0529]
Regulation 2	-0.2919	-0.2949	-0.2972
	[0.1953]	[0.1955]	[0.1944]
Regulation 3	-0.1687***	-0.1633***	-0.1649***
	[0.0209]	[0.0210]	[0.0220]
Obs.	2,435	2,438	2,442
Pseudo R-Squared	0.1027	0.1038	0.1022

Note: Probit model in which the dependent variable equals one for withdrawn offerings and zero for completed offerings. Specification (5) of Table 3.3 is taken as baseline speicfication and VC dummy is replaced by different VC characteristics one-by-one. Descriptions of the variables and data sources are depicted in Table B.1. Marginal effects at the mean (MEM) are reported. Pseudo R-squared corresponds to McFadden's pseudo R-squared. Standard errors clustered at industry level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

In the next step, we want to examine the VC effects under more and less risky market environments. To this aim, we run regressions with interaction terms of VC variables and the indicator of highly volatile markets. To find out whether the VC (binary) variables are more or less significant in highly volatile markets, we evaluate their effects (of a change between zero and one) separately at both values of the high volatility variable. We show the results in Table 3.6.

The overall VC effect, which was insignificant in the baseline regressions, remains insignificant in highly volatile and in less volatile environments. From the analysis of VC characteristics, we conclude that the syndicate effect, which we have found in baseline regressions, is present in highly volatile, but not in less volatile environments. VC syndicates thus seem to provide a better signaling device than stand-alone VCs when investors face greater risks and uncertainties. In the baseline analysis, VC reputation was insignificant. When we distinguish between less volatile and highly volatile markets, we find that reputable VCs are associated with lower withdrawal rates only in less volatile and not in highly volatile environments. This result does not support the signaling explanation. Rather, it suggests that reputable VCs are more likely to follow a dual-track strategy or postpone their IPOs when market conditions get more risky.

Table 3.6: Interactions of VC characteristics with high volatility indicator

	at high volatility=0	at high volatility=1
VC backing	0.034	-0.017
	[0.024]	[0.021]
VC syndication	-0.024	-0.074**
	[0.042]	[0.035]
VC US	-0.052	-0.027
	[0.053]	[0.022]
VC reputation	-0.036**	0.014
	[0.015]	[0.020]

Note: We evaluate the effect of the dummy variables at both values of the high volatility variable. Descriptions of the variables and data sources are depicted in Table B.1. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

3.6 Conclusions

This paper contributes to a better understanding of the withdrawal phenomenon. We demonstrate that corporate governance characteristics may become important when markets get highly volatile. In risky environments, investors seem to rely on signals of high quality and strong monitoring that result from large and experienced boards. VCs seem to be an important force behind the withdrawal decision too, but the VC effect varies with the VC heterogeneity. While in volatile markets, VC syndication may serve as a positive signal that is associated with a lower withdrawal probability, reputable VCs may more likely be seeking for alternatives to a current IPO.

This study raises some interesting questions for further research. First, our paper and most other studies deal with US data. Recently, researchers have started to analyze withdrawals in other regions, such as Europe (Helbing et al., 2019) or Asia (Fan and Yamada, 2020). More evidence is needed because the withdrawal determinants seem to differ in different regions (Helbing et al., 2019). As corporate governance and VC financing are very heterogeneous across US, Europe and Asia, it would be interesting to see whether and how these heterogeneities relate to the differences in withdrawal probabilities. Second, our conjecture that the differences in the IPO outcome for reputable VCs in highly volatile and less volatile markets may be associated with the dual-track strategy deserves a deeper investigation. Further research might examine, for example, the characteristics of firms that follow the dual-track strategy and analyze its outcomes as well as implications for the various market participants (issuers, intermediaries, potential IPO investors, and potential acquirers).

B Appendix

B.1 Tables

Table B.1: Variable descriptions and data sources

Variable	**Description**	**Source**	**Expected effect**
Dependent variable			
Withdrawal	Dummy variable that equals one if the filing is withdrawn and zero if the filing is completed.	Thomson Reuters Securities Data Company (SDC) Platinum	
Independent variables			
Corporate governance characteristics			
Board size	Absolute number of board members.	EDGAR filings	–
Board experience	Average age of all board members.	EDGAR filings	–
Female board members	Ratio of number of female board members to total number of board members.	EDGAR filings	–
CEO duality	Dummy that equals one if both the role of CEO and the role of chairman reside with the CEO of the company and zero otherwise.	EDGAR filings	+
Lock-up	Dummy that equals one if the lock-up period that is reported in EDGAR filings is larger than 180 and zero otherwise.	EDGAR filings	–
Retained shares	Proportion of shares hold by insiders post offering. (Shares outstanding post offering - (primary shares + secondary shares))/shares outstanding post offering	Own calculation based on data stemming from S&P Capital IQ, Bloomberg and EDGAR filings	–

Intermediary characteristics

VC backing	Dummy indicating whether the firm is backed by a venture capital investor at the time of the filing. Equals one if firm is backed by venture capital and zero otherwise.	S&P Capital IQ, EDGAR filings, Crunchbase	+ / −
Non-VC	Dummy that equals one if firm is not backed by venture capital and zero otherwise.	S&P Capital IQ, EDGAR filings, Crunchbase	
VC syndication	Dummy that equals one if the firm is backed by venture capital and the VC deal is syndicated and zero otherwise.	S&P Capital IQ, EDGAR filings, Crunchbase	+ / −
VC US	Dummy that equals one if the firm is backed by venture capital and all VCs that are involved in the deal are located in the US and zero otherwise.	S&P Capital IQ, EDGAR filings, Crunchbase	+ / −
VC reputation	Dummy variable that equals one if the firm is backed by high reputable VCs and zero otherwise. A VC deal is considered to be reputable if the VC with the highest rank is among the top 100 according to the reputation index provided by Lee et al. (2011). The index ends in 2010. Thus, we use the ranking of 2010 for all years after 2010.	S&P Capital IQ, EDGAR filings, Crunchbase. VC reputation index provided by Lee et al. (2011) (`http://www.timothypollock.com/vc_reputation.htm`)	+ / −
Number of underwriters	Number of underwriters involved in the issue.	Thomson Reuters Securities Data Company (SDC) Platinum	+ / −

Underwriter market share	Market share of lead underwriter in filing year. The amount of a transaction is divided by the total number of bookrunners on the transaction. Each bookrunner will receive that amount as credit for the transaction.	Lead Tables of Thomson Reuters Securities Data Company (SDC) Platinum, EDGAR filings	+ / −
CM Rank	Carter - Manaster Rank of lead underwriter in filing year. It ranges from 0 to 9.	Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`). Originally developed by Carter and Manaster (1990), updated by Carter, Dark, and Singh (1998) and by Loughran and Ritter (2004).	+ / −
Issuer and issue characteristics			
Filing size (bn USD)	Filing size in bn USD. Calculated as the product of original global shares filed and original mid filing price.	Thomson Reuters Securities Data Company (SDC) Platinum	−
Firm size (bn USD)	Total assets in bn USD in most recent financial period before outcome.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	−
Firm age	Firm age at outcome date in years (outcome date − founding date)	Founding and outcome date primarily taken from Thomson Reuters Securities Data Company (SDC) Platinum. Supplemented with and checked against data stemming from Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`) and S&P Capital IQ.	−

High-tech	Dummy indicating high tech industry. Equals one if primary industry of the firm is high-tech industry according to the SDC definition and zero otherwise.	Thomson Reuters Securities Data Company (SDC) Platinum	–
Debt ratio	Ratio of total liabilities to total assets in most recent financial period before outcome.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	+ / –
BM ratio	Book-to-market ratio for firms in the issuer's Fama-French industry (49) in the year before filing.	Kenneth R. French's Data Library (Kenneth R. French's website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`).	+ / –
Asset turnover	Ratio of total revenues to total assets in most recent financial period before outcome.	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.	–
Negative news	Dummy that equals one if the LexisNexis Negative News Search delivers that negative business news were reported about the firm one year prior to the outcome date and zero otherwise.	LexisNexis	+
Market characteristics at filing			
Nasdaq (index/100)	Nasdaq composite index (scaled by 100)	FRED, webiste of Federal Reserve Bank of St. Louis (`https://research.stlouisfed.org/`)	–

Implied volatility	Based on CBOE S&P 100 volatility index (VXO). It measures the expected volatility over the next 30 days. We choose VXO instead of VIX because it is continuously available over the whole sample period. The calculation of VIX changes in 2003. For our implied volatility measure we calculate: VXO at filing date − VXO one month before filing date / VXO one month before filing date.	CBOE (`http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data`)	+
Yield spread	BAA-AAA yield spread at filing date. Difference between the Moody's BAA-rated corporate bonds and the AAA-rated bond.	FRED, webiste of Federal Reserve Bank of St. Louis (`https://research.stlouisfed.org/`)	−
Number of filings	Number of new filings +/− 21 days around the filing date (per day)	Own calculations based on data from Thomson Reuters Securities Data Company (SDC) Platinum	+
Market characteristics after filing			
Nasdaq return (daily)	Compounded daily Nasdaq returns after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	−
Nasdaq std. dev.	Standard deviation of Nasdaq returns after filing (scaled by 100).	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	+
Yield spread change	Change in BAA-AAA yield spread after filing.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	+ / −

Lending standard change	Change in bank lending standard around the filing date. Based on the quarterly Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices. The particular question was "over the past three months, how have your bank's credit standards for approving applications for C&I loans or credit lines - other than those to be used to finance mergers and acquisitions - to large and middle-market firms and to small firms changed". The response options range from (1) "Tightened considerable" to (5) "Eased considerable". The lending standard measure is then calculated as the difference between the number of loan officers reporting tightening lending standards and the number of officers reporting easing standards divided by the total number of reporting times 100. Thus, a positive value implies that lending standards have tightened and a negative value implies that the lending standards have eased. We employ the nearest quarter after the filing date.	Federal Reserve Board's Senior Loan Officer Opinion Survey on Bank Lending Practices (SLOOS) (`https://www.federalreserve.gov/data/sloos.htm`)	+ / −
High volatility	Dummy that equals one if the Nasdaq std. dev. is in the top tertile, zero otherwise.	own calculations based on on FRED data (`https://research.stlouisfed.org/`)	+
Time dummies			
Regulation 1	Dummy that equals one between April 2001 and September 2004 and zero otherwise. It starts the month following the amendment of Rule 477 and the adoption of Rule 155.	Cf. Busaba et al. (2015)	+

Regulation 2	Dummy that equals one during the financial crisis (between August 2008 and May 2009) and zero otherwise.	Cf. Busaba et al. (2015)	+
Regulation 3	Dummy that equals one after the Jumpstart our Business Startups (JOBS) Act (since May 2012) until the end of the sample period and zero otherwise.	Cf. Busaba et al. (2015)	−

Note: Firm age, debt ratio, firm size, asset turnover, filing size, board size and board experience are winsorized at the 1% level in each tail

Table B.2: Correlation table

	Filing size	Firm size	Firm age	High-tech	Debt ratio	BM ratio	Asset turnover	Negative news	Board size	Board experience
Filing size	1.00									
Firm size	0.47***	1.00								
Firm age	0.22***	0.34***	1.00							
High-tech	−0.18 ***	−0.21 ***	−0.28 ***	1.00						
Debt ratio	−0.05 ***	−0.02	0.01	0.01	1.00					
BM ratio	0.25***	0.24***	0.15***	−0.50 ***	0.00	1.00				
Asset turnover	−0.07 ***	−0.10 ***	0.09***	−0.13 ***	0.16***	−0.05 ***	1.00			
Negative news	0.07***	0.08***	0.01	0.11***	−0.04 **	−0.03 *	−0.07 ***	1.00		
Board size	0.05***	0.15***	0.11***	0.04**	−0.03 *	0.04**	−0.14 **	0.11***	1.00	
Board experience	0.06***	0.14***	0.18***	−0.16 ***	0.06***	0.21***	−0.02	0.00	0.17***	1.00
Female board members	−0.02	0.05***	0.04**	−0.02	0.00	0.00	−0.03	0.01	0.07***	−0.01
CEO duality	−0.04 *	−0.03 *	−0.04 **	−0.01	−0.03 *	−0.02	0.02	−0.01	−0.11 ***	−0.10 ***
Lock-up	−0.11 ***	−0.04 **	−0.04 **	−0.05 ***	0.12***	−0.03 *	0.11***	−0.07 ***	−0.10 ***	−0.01
Retained shares	−0.09 ***	0.01	−0.10 ***	0.21***	−0.09 ***	−0.17 ***	−0.08 ***	0.06***	0.11***	−0.09 ***
VC backing	−0.13 ***	−0.16 ***	−0.25 ***	0.36***	−0.07 ***	−0.20 ***	−0.21 ***	0.11***	0.13***	−0.10 ***
VC syndication	−0.06 **	−0.02	−0.06 **	0.17***	−0.06 **	−0.07 ***	−0.06 **	0.03	0.05*	−0.06 **
VC US	0.05*	0.08***	0.02	−0.10 ***	0.03	0.05*	0.11***	−0.05 *	−0.06 **	−0.05 *
VC reputation	−0.07 **	−0.04	−0.04	0.20***	−0.13 ***	−0.10 ***	−0.10 ***	0.01	0.00	−0.05 *
Number of underwriters	0.38***	0.35***	0.13***	−0.18 ***	−0.01	0.26***	−0.07 ***	0.06***	0.07***	0.13***
Underwriter market share	0.32***	0.20***	0.08***	0.00	−0.13 ***	0.10***	−0.08 ***	0.14***	0.08***	−0.07 ***
CM Rank	0.28***	0.14***	0.08***	0.08***	−0.22 ***	0.04**	−0.10 ***	0.16***	0.16***	−0.08 ***
Nasdaq	0.02	0.00	−0.12 ***	0.23***	−0.04 **	−0.03 **	−0.19 ***	0.06***	0.06***	−0.04 **
Implied volatility	0.01	0.03	−0.01	−0.01	−0.01	0.01	0.00	0.02	−0.03 *	0.00
Yield spread	0.18***	0.07***	0.04***	−0.08 ***	0.06***	0.32***	0.00**	0.04**	0.05***	0.15***
Number of filings	−0.20 ***	−0.14 ***	−0.13 ***	0.22***	−0.09 ***	−0.32 ***	−0.04 **	−0.01	−0.04 **	−0.29 ***
Nasdaq return	−0.02	−0.02	−0.06 ***	0.02	0.01	−0.06 ***	0.00	0.03*	−0.02	−0.03 *
Nasdaq std. dev.	−0.05 ***	−0.06 ***	−0.12 ***	0.23***	−0.06 ***	−0.12 ***	−0.11 ***	0.03*	0.01	−0.19 ***
Yield spread change	0.02	0.01	−0.01	0.04***	−0.01	0.03	−0.02	−0.04 **	0.00	−0.02
Lending standard change	−0.04 **	−0.04 **	−0.10 ***	0.18***	−0.05 ***	−0.08 ***	−0.06 ***	0.03*	0.01	−0.18 ***

Table B.2: Correlation table (continued)

	Female board members	CEO duality	Lock-up	Retained shares	VC backing	VC syndication	VC US	VC reputation	Number of underwriters	Underwriter market share
Female board members	1.00									
CEO duality	−0.02	1.00								
Lock-up	−0.03	0.04**	1.00							
Retained shares	0.01	0.00	−0.07 ***	1.00						
VC backing	0.02	−0.05 **	−0.11 ***	0.19***	1.00					
VC syndication	0.00	−0.04	−0.02	0.05*		1.00				
VC US	−0.01	0.11***	0.00	0.02		−0.17 ***	1.00			
VC reputation	0.05**	−0.02	−0.04	0.12***		0.34***	−0.03	1.00		
Number of underwriters	0.01	−0.07 ***	−0.09 ***	−0.03	−0.07 ***	−0.03	0.00	0.00	1.00	
Underwriter market share	−0.03	−0.03 *	−0.15 ***	0.08***	0.08***	0.05*	0.02	0.16***	0.13***	1.00
CM Rank	0.01	−0.07 ***	−0.40 ***	0.11***	0.19***	0.04	0.00	0.13***	0.20***	0.55***
Nasdaq	0.01	−0.05 ***	−0.10 ***	0.18***	0.23***	0.11***	−0.08 ***	0.09***	0.17***	0.08***
Implied volatility	0.02	−0.01	−0.01	−0.02	0.00	0.01	−0.02	0.00	0.00	0.00
Yield spread	0.02	−0.06 ***	−0.04 *	−0.04 *	0.01	0.00	0.02	0.00	0.22***	0.15***
Number of filings	−0.04 **	0.09***	0.01	0.15***	0.08***	0.02	0.00	0.05*	−0.35 ***	−0.04 **
Nasdaq return	0.01	−0.01	0.04**	0.02	0.02	−0.02	0.02	0.00	0.01	−0.02
Nasdaq std. dev.	−0.02	0.04**	−0.05 ***	0.18***	0.17***	0.07*	−0.04 *	0.04	−0.15 ***	0.07***
Yield spread change	−0.03 *	0.01	−0.04 **	0.01	0.03	0.03	0.00	−0.03	−0.01	0.03*
Lending standard change	−0.01	0.03	−0.03	0.17***	0.09***	0.06**	0.01	0.02	−0.24 ***	0.11***

Table B.2: Correlation table (continued)

	CM Rank	Nasdaq	Implied volatility	Yield spread	Number of filings	Nasdaq return	Nasdaq std dev	Yield spread change	Lending standard change
CM Rank	1.00								
Nasdaq	0.19***	1.00							
Implied volatility	0.02	−0.03	1.00						
Yield spread	0.08***	−0.07 ***	−0.03 **	1.00					
Number of filings	0.02	0.29***	0.05***	−0.50 ***	1.00				
Nasdaq return	−0.02	−0.07 ***	0.02	−0.04 **	−0.09 ***	1.00			
Nasdaq std. dev.	0.12***	0.68***	0.03*	−0.12 ***	0.52***	−0.19 ***	1.00		
Yield spread change	0.04**	0.16***	0.25***	−0.12 ***	0.14***	−0.23 ***	0.22***	1.00	
Lending standard change	0.11***	0.33***	0.08***	0.10***	0.26***	−0.16 ***	0.57***	0.19***	1.00

Table B.3: Industry- and time-specific analysis

Industry-specific analysis							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MEM	MEM	MEM	MEM	MEM	MEM	MEM
Corporate governance characteristics							
Board size	0.0001	-0.0013	-0.0004	0.0006	-0.0001	-0.0034	-0.0007
	[0.0043]	[0.0041]	[0.0042]	[0.0043]	[0.0041]	[0.0034]	[0.0043]
Board experience	0.0002	-0.0002	0.0000	-0.0002	0.0001	-0.0009	0.0003
	[0.0008]	[0.0008]	[0.0008]	[0.0009]	[0.0008]	[0.0008]	[0.0009]
Female board members	-0.0134	0.0106	0.0231	0.0095	0.0063	0.0172	0.0079
	[0.0884]	[0.0772]	[0.0752]	[0.0772]	[0.0788]	[0.0931]	[0.0789]
CEO duality	-0.0170	-0.0202	-0.0174	-0.0197	-0.0131	0.0188	-0.0162
	[0.0244]	[0.0226]	[0.0243]	[0.0230]	[0.0252]	[0.0200]	[0.0246]
Lock-up	-0.1673***	-0.1522***	-0.1559***	-0.1586***	-0.1853***	-0.1211***	-0.1480***
	[0.0618]	[0.0495]	[0.0546]	[0.0489]	[0.0482]	[0.0393]	[0.0501]
Intermediary characteristics							
VC backing	0.0112	0.0158	0.0055	0.0062	0.0114	0.0236	0.0094
	[0.0130]	[0.0127]	[0.0106]	[0.0106]	[0.0127]	[0.0165]	[0.0128]
Number of underwriters	-0.0146	-0.0120	-0.0113	-0.009	-0.0141	-0.0228**	-0.0146
	[0.0113]	[0.0109]	[0.0110]	[0.0102]	[0.0111]	[0.0102]	[0.0111]
Underwriter market share	-0.2150	-0.3034*	-0.2667	-0.2386	-0.2458	-0.0860	-0.3307**
	[0.1813]	[0.1627]	[0.1776]	[0.1856]	[0.1787]	[0.2047]	[0.1602]
CM Rank	-0.0063	-0.0052	-0.0060	-0.0051	-0.0059	-0.0060	-0.0038
	[0.0075]	[0.0076]	[0.0079]	[0.0079]	[0.0075]	[0.0098]	[0.0073]
Issuer and issue characteristics							
Filing size (bn USD)	-0.1452	-0.1328	-0.1221	-0.2308***	-0.1436	-0.1452	-0.1379
	[0.1222]	[0.1212]	[0.1136]	[0.0730]	[0.1168]	[0.1420]	[0.1225]
Firm size (bn USD)	0.0105	-0.0031	-0.0032	0.0175	0.0060	0.0156	0.0161
	[0.0377]	[0.0367]	[0.0341]	[0.0364]	[0.0367]	[0.0407]	[0.0384]
Firm age	-0.0233**	-0.0237**	-0.0251**	-0.0251**	-0.0264***	-0.0231*	-0.0242**
	[0.0100]	[0.0097]	[0.0106]	[0.0105]	[0.0101]	[0.0126]	[0.0101]
High-tech	-0.0118	-0.0041	-0.0155	-0.0130	-0.0111	-0.0204	-0.0121
	[0.0210]	[0.0198]	[0.0229]	[0.0222]	[0.0218]	[0.0275]	[0.0211]
Debt ratio	0.0359**	0.0311*	0.0344**	0.0343**	0.0367**	0.0238	0.0342**
	[0.0168]	[0.0160]	[0.0164]	[0.0161]	[0.0166]	[0.0154]	[0.0163]
BM ratio	0.2226**	0.2420***	0.2379***	0.2091**	0.2180**	0.1320	0.1837*
	[0.0897]	[0.0835]	[0.0911]	[0.0910]	[0.0877]	[0.0819]	[0.1003]
Asset turnover	0.0119**	0.0099*	0.0105*	0.0100	0.0118**	0.0070	0.0085
	[0.0053]	[0.0054]	[0.0056]	[0.0062]	[0.0052]	[0.0060]	[0.0055]
Negative news	0.0143	0.0109	0.0154	0.0138	0.0132	0.0073	0.0112
	[0.0118]	[0.0106]	[0.0116]	[0.0116]	[0.0111]	[0.0139]	[0.0114]
Market characteristics at filing							
Nasdaq	0.0062***	0.0053***	0.0058***	0.0060***	0.0062***	0.0077***	0.0063***
	[0.0017]	[0.0016]	[0.0018]	[0.0017]	[0.0017]	[0.0020]	[0.0017]
Implied volatility	0.0710**	0.0851***	0.0575*	0.0675**	0.0700**	0.0671	0.0730**
	[0.0306]	[0.0256]	[0.0300]	[0.0317]	[0.0306]	[0.0466]	[0.0300]
Yield spread	0.0163	0.0094	0.0183	0.0014	0.0142	0.0672	0.0415
	[0.0524]	[0.0523]	[0.0555]	[0.0485]	[0.0506]	[0.0494]	[0.0521]
Number of filings	-0.6138***	-0.5318***	-0.5982***	-0.5898***	-0.6024***	-0.5643***	-0.5272***
	[0.0904]	[0.0906]	[0.0957]	[0.1018]	[0.0953]	[0.1609]	[0.0948]
Market characteristics after filing							
Nasdaq return	-0.0805***	-0.0700**	-0.0712**	-0.0707**	-0.0714**	-0.0365	-0.0690*
	[0.0306]	[0.0330]	[0.0346]	[0.0337]	[0.0333]	[0.0389]	[0.0355]
Nasdaq std. dev.	-0.0308	-0.0286	-0.0242	-0.0291	-0.0295	-0.0599	-0.0406
	[0.0319]	[0.0306]	[0.0314]	[0.0317]	[0.0308]	[0.0448]	[0.0321]
Yield spread change	0.3124***	0.2969***	0.3317***	0.3175***	0.2996***	0.4265***	0.3002***
	[0.1095]	[0.1046]	[0.1095]	[0.1107]	[0.1066]	[0.1375]	[0.1046]
Lending standard change	0.0037***	0.0039***	0.0038***	0.0038***	0.0037***	0.0037***	0.0038***
	[0.0005]	[0.0004]	[0.0005]	[0.0005]	[0.0005]	[0.0007]	[0.0005]
Regulation 1	-0.0369	-0.0332	-0.0359	-0.0263	-0.0295	-0.076	-0.0292
	[0.0549]	[0.0565]	[0.0553]	[0.0566]	[0.0539]	[0.0496]	[0.0546]
Regulation 2	-0.3085	-0.2942	-0.3610*	-0.2808	-0.2893	-0.0847	-0.3016
	[0.1935]	[0.1932]	[0.2071]	[0.1986]	[0.1933]	[0.1986]	[0.1916]
Regulation 3	-0.1766***	-0.1588***	-0.1759***	-0.1781***	-0.1729***	-0.1790***	-0.1746***
	[0.0216]	[0.0206]	[0.0232]	[0.0225]	[0.0234]	[0.0338]	[0.0223]
Obs.	2,401	2,428	2,380	2,388	2,438	1,616	2,358
Pseudo R-Squared	0.1052	0.1046	0.1072	0.1046	0.105	0.0962	0.0978

Note: Robustness test of the probit analysis of the decision to withdraw an offering. Specification (5) of Table 3.3 is re-estimated. Specfications (1) - (12) report an industry-specific analysis. In each regression, one industry of the Fama French 12 industries is left out (industry 1 is excluded in specification (1), industry 2 is excluded in specification (2), etc.).

(table continued on next page)

Table B.3: Industry- and time-specific analysis (continued)

Industry-specific analysis					Time-specific analysis		
(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
MEM	MEM	MEM	MEM	MEM	MEM	MEM	MEM
Corporate governance characteristics							
-0.0012	0.0011	-0.0005	0.0034	-0.0004	-0.002	0.0007	-0.0004
[0.0041]	[0.0040]	[0.0042]	[0.0035]	[0.0047]	[0.0030]	[0.0041]	[0.0031]
0.0002	0.0002	0.0001	0.0004	0.0005	-0.0005	0.0003	-0.0002
[0.0008]	[0.0009]	[0.0012]	[0.0008]	[0.0008]	[0.0018]	[0.0009]	[0.0020]
0.0166	-0.0613	-0.0107	0.0480	0.0227	0.0486	0.0274	0.0678
[0.0783]	[0.0561]	[0.0890]	[0.0733]	[0.0878]	[0.0823]	[0.0744]	[0.0822]
-0.0160	-0.0102	-0.0185	-0.0072	-0.0162	-0.0281	-0.0177	-0.0334
[0.0245]	[0.0276]	[0.0266]	[0.0281]	[0.0274]	[0.0348]	[0.0236]	[0.0322]
-0.1578***	-0.1707***	-0.1783***	-0.1786***	-0.1538***	-0.1431**	-0.1529***	-0.1377**
[0.0488]	[0.0558]	[0.0612]	[0.0481]	[0.0542]	[0.0643]	[0.0529]	[0.0682]
Intermediary characteristics							
0.0108	0.0142	0.0094	0.0146	0.0090	0.0126	0.0044	0.0052
[0.0131]	[0.0135]	[0.0146]	[0.0141]	[0.0146]	[0.0166]	[0.0125]	[0.0191]
-0.0105	-0.0151	-0.0141	-0.0202*	-0.0084	-0.0387***	-0.0072	-0.0316***
[0.0111]	[0.0119]	[0.0127]	[0.0117]	[0.0101]	[0.0095]	[0.0111]	[0.0103]
-0.2680	-0.1858	-0.2909*	-0.3589**	-0.3078*	-0.1839	-0.2773	-0.1884
[0.1809]	[0.1855]	[0.1752]	[0.1684]	[0.1867]	[0.1864]	[0.1894]	[0.1929]
-0.0058	-0.0082	0.0020	-0.0096	-0.0067	-0.0081	-0.0025	-0.0051
[0.0075]	[0.0076]	[0.0060]	[0.0066]	[0.0084]	[0.0054]	[0.0075]	[0.0056]
Issuer and issue characteristics							
-0.1519	-0.1194	-0.0963	-0.1737	-0.0763	-0.0552	-0.1345	-0.0484
[0.1138]	[0.1283]	[0.1119]	[0.1407]	[0.1048]	[0.1268]	[0.1278]	[0.1468]
0.0121	-0.0196	0.0132	0.0440	0.0068	0.0265	0.0056	0.0242
[0.0383]	[0.0234]	[0.0352]	[0.0482]	[0.0388]	[0.0405]	[0.0371]	[0.0429]
-0.0278***	-0.0195**	-0.0298***	-0.0240**	-0.0308***	-0.0270***	-0.0245**	-0.0244**
[0.0096]	[0.0089]	[0.0096]	[0.0108]	[0.0100]	[0.0098]	[0.0104]	[0.0099]
-0.0134	-0.0085	-0.0316	-0.0209	-0.0115	-0.0329*	-0.0116	-0.0268
[0.0214]	[0.0243]	[0.0277]	[0.0212]	[0.0227]	[0.0179]	[0.0208]	[0.0197]
0.0342**	0.0339**	0.0532***	0.0368**	0.0342*	0.0292*	0.0322**	0.0257*
[0.0162]	[0.0170]	[0.0101]	[0.0184]	[0.0177]	[0.0154]	[0.0160]	[0.0149]
0.2225**	0.2163**	0.1857**	0.2700***	0.1824*	0.0144	0.2041**	0.0208
[0.0896]	[0.0890]	[0.0852]	[0.0938]	[0.1020]	[0.0436]	[0.0850]	[0.0386]
0.0102*	0.0114**	0.0110*	0.0056	0.0114*	0.0090	0.0120**	0.0114
[0.0057]	[0.0057]	[0.0063]	[0.0058]	[0.0067]	[0.0071]	[0.0053]	[0.0078]
0.0114	0.0119	0.0133	0.0155	0.0229***	0.0431**	0.0136	0.0440**
[0.0110]	[0.0116]	[0.0130]	[0.0121]	[0.0077]	[0.0185]	[0.0109]	[0.0191]
Market characteristics at filing							
0.0061***	0.0054***	0.0059***	0.0059***	0.0050***	0.0214***	0.0055***	0.0200***
[0.0017]	[0.0017]	[0.0019]	[0.0017]	[0.0014]	[0.0019]	[0.0016]	[0.0019]
0.0759**	0.0784**	0.0524	0.0799**	0.0831***	0.0899***	0.0695**	0.0806***
[0.0301]	[0.0305]	[0.0333]	[0.0328]	[0.0291]	[0.0262]	[0.0305]	[0.0242]
0.0236	0.0210	0.0329	0.0196	0.0028	0.1104***	0.0080	0.1265
[0.0554]	[0.0577]	[0.0657]	[0.0546]	[0.0571]	[0.0417]	[0.0824]	[0.0789]
-0.5622***	-0.6003***	-0.4955***	-0.5717***	-0.5820***	0.2299	-0.5359***	0.2997
[0.1035]	[0.1074]	[0.1194]	[0.1007]	[0.1116]	[0.1881]	[0.1048]	[0.2089]
Market characteristics after filing							
-0.0698**	-0.0861***	-0.0493	-0.0732**	-0.0758**	-0.0346	-0.0665*	-0.0294
[0.0335]	[0.0279]	[0.0395]	[0.0353]	[0.0371]	[0.0323]	[0.0354]	[0.0351]
-0.0286	-0.0279	-0.0026	-0.0269	-0.0305	-0.1491**	-0.0244	-0.1488***
[0.0310]	[0.0337]	[0.0147]	[0.0314]	[0.0363]	[0.0636]	[0.0315]	[0.0529]
0.3133***	0.2665***	0.3420**	0.3170***	0.2547***	0.1414	0.3346***	0.1723
[0.1029]	[0.0979]	[0.1421]	[0.1108]	[0.0848]	[0.1277]	[0.0989]	[0.1180]
0.0038***	0.0039***	0.0034***	0.0038***	0.0040***	0.0046***	0.0037***	0.0047***
[0.0005]	[0.0004]	[0.0005]	[0.0005]	[0.0005]	[0.0006]	[0.0006]	[0.0006]
-0.0278	-0.0279	0.0126	-0.0189	-0.0179	-0.0036	-0.0216	-0.0044
[0.0531]	[0.0609]	[0.0377]	[0.0569]	[0.0613]	[0.0473]	[0.0587]	[0.0520]
-0.4476**	-0.3101*	-0.2654	-0.3348*	-0.2973	-0.2538*	0	0
[0.1823]	[0.1831]	[0.2028]	[0.1954]	[0.2597]	[0.1443]	[.]	[.]
-0.1740***	-0.1598***	-0.1753***	-0.1594***	-0.1758***	-0.4031***	-0.1659***	-0.3777***
[0.0228]	[0.0215]	[0.0291]	[0.0225]	[0.0250]	[0.0382]	[0.0247]	[0.0367]
2,445	2,247	2,058	2,283	2,106	1,748	2,391	1,671
0.1048	0.1093	0.1112	0.1063	0.0991	0.1151	0.0949	0.1016

Note (continued): Specifications (13) - (14) show a time-specific analysis. Specification (13) excludes years of the dotcom bubble (1999/2000), specification (14) excludes years of the financial crisis (2008/2009) and specification (15) excludes all years of both crises. The dependent variable equals one for withdrawn offerings and zero for completed offerings. Descriptions of the variables and data sources are depicted in Table B.1. Marginal effects at the mean (MEM) are reported. Pseudo R-squared corresponds to McFadden's pseudo R-squared. Standard errors clustered at industry level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table B.4: Corporate governance and high volatility

	(1)	(2)	(3)	(4)	(5)	(6)
Corporate governance characteristics						
Board size	0.0201	-0.0040	-0.0016	-0.0022	-0.0014	-0.0083
	[0.0157]	[0.0134]	[0.0138]	[0.0137]	[0.0137]	[0.0152]
Board experience	-0.0008	0.0124**	-0.0005	-0.0005	-0.0009	0.0034
	[0.0030]	[0.0052]	[0.0029]	[0.0028]	[0.0028]	[0.0045]
Female board members	0.0247	0.0506	-0.3230	0.0438	0.0296	0.4674*
	[0.2585]	[0.2610]	[0.3488]	[0.2614]	[0.2694]	[0.2502]
CEO duality	-0.0495	-0.0476	-0.0470	-0.1418	-0.0528	-0.0468
	[0.0847]	[0.0839]	[0.0862]	[0.1255]	[0.0852]	[0.0864]
Lock-up	-0.5527***	-0.5158***	-0.5382***	-0.5552***	-1.1533***	-0.5070***
	[0.1784]	[0.1825]	[0.1880]	[0.1853]	[0.1974]	[0.1904]
Retained shares						0.1374
						[0.1532]
Intermediary characteristics						
VC backing	0.0450	0.0478	0.0448	0.0479	0.0455	0.0451
	[0.0420]	[0.0408]	[0.0421]	[0.0406]	[0.0417]	[0.0771]
Number of underwriters	-0.0535	-0.0487	-0.0517	-0.0491	-0.0521	0.0149
	[0.0360]	[0.0352]	[0.0351]	[0.0352]	[0.0357]	[0.0459]
Underwriter market share	-0.8881	-0.8943	-0.8847	-0.8992	-0.8123	-1.9413*
	[0.6171]	[0.6100]	[0.6157]	[0.6206]	[0.6085]	[1.0308]
CM Rank	-0.0197	-0.0158	-0.0176n	-0.0196	-0.0204	-0.0640**
	[0.0252]	[0.0242]	[0.0255]	[0.0254]	[0.0250]	[0.0259]
Issuer and issue characteristics						
Filing_size (bn USD)	-0.4187	-0.4494	-0.4993	-0.4705	-0.4747	-0.6054
	[0.3707]	[0.3578]	[0.3749]	[0.3709]	[0.3709]	[0.4288]
Firm size (bn USD)	0.0292	0.0532	0.0476	0.0479	0.0424	-0.1134
	[0.1191]	[0.1174]	[0.1206]	[0.1209]	[0.1209]	[0.1241]
Firm age	-0.0921***	-0.0919***	-0.0933***	-0.0927***	-0.0958***	-0.1095
	[0.0331]	[0.0321]	[0.0316]	[0.0330]	[0.0312]	[0.0694]
High-tech	-0.0591	-0.0601	-0.0521	-0.0574	-0.0551	-0.0867
	[0.0713]	[0.0726]	[0.0699]	[0.0702]	[0.0696]	[0.0945]
Debt ratio	0.1236**	0.1208**	0.1216**	0.1228**	0.1275**	0.1116*
	[0.0587]	[0.0584]	[0.0575]	[0.0603]	[0.0600]	[0.0592]
BM ratio	0.7034**	0.6662**	0.6923**	0.6810**	0.7144***	0.6259**
	[0.2851]	[0.2791]	[0.2851]	[0.2785]	[0.2730]	[0.2936]
Asset turnover	0.0333	0.0366*	0.0365*	0.0377*	0.0402**	0.0281
	[0.0204]	[0.0209]	[0.0190]	[0.0205]	[0.0181]	[0.0314]
Negative news	0.0463	0.0313	0.0461	0.0463	0.0473	0.0423
	[0.0392]	[0.0365]	[0.0382]	[0.0402]	[0.0385]	[0.0576]
Market characteristics at filing						
Nasdaq	0.0207***	0.0190***	0.0209***	0.0202***	0.0212***	0.0134***
	[0.0052]	[0.0052]	[0.0052]	[0.0053]	[0.0049]	[0.0039]
Implied volatility	0.2652***	0.2712**	0.2566**	0.2496**	0.2682***	-0.0499
	[0.0958]	[0.1065]	[0.1012]	[0.0969]	[0.1012]	[0.1632]
Yield spread	0.0185	0.0003	0.0377	0.0286	0.0059	-0.2674
	[0.1828]	[0.1843]	[0.1876]	[0.1825]	[0.1767]	[0.1749]
Number of filings	-2.2391***	-2.3021***	-2.2390***	-2.2416***	-2.2398***	-0.1959
	[0.4333]	[0.4160]	[0.4170]	[0.4125]	[0.4084]	[0.8738]
Market characteristics after filing						
Nasdaq return	-0.1974*	-0.2164**	-0.1978*	-0.2008*	-0.2061**	-0.0346
	[0.1084]	[0.1059]	[0.1078]	[0.1074]	[0.1046]	[0.0935]
high volatility	0.2402	1.5409***	-0.2440***	-0.3328***	-0.2413***	0.2942

	[0.1840]	[0.3890]	[0.0668]	[0.1014]	[0.0824]	[0.3577]
Yield spread change	1.1075***	1.1005***	1.0927***	1.1010***	1.0962***	1.0807***
	[0.3594]	[0.3785]	[0.3490]	[0.3607]	[0.3529]	[0.2991]
Lending standard change	0.0139***	0.0140***	0.0138***	0.0139***	0.0141***	0.0136***
	[0.0018]	[0.0017]	[0.0018]	[0.0018]	[0.0018]	[0.0013]
Regulation 1	-0.1030	-0.1010	-0.1051	-0.1021	-0.1028	-0.0384
	[0.1920]	[0.1926]	[0.1934]	[0.1925]	[0.1880]	[0.2121]
Regulation 2	-0.8705	-0.9057	-0.9643	-0.9421	-0.8759	-0.3243
	[0.6664]	[0.6488]	[0.6574]	[0.6293]	[0.6739]	[0.7539]
Regulation 3	-0.5978***	-0.5866***	-0.5931***	-0.5949***	-0.5994***	-0.6910***
	[0.0929]	[0.0968]	[0.0940]	[0.0960]	[0.0934]	[0.1964]
Board size	-0.0682***					
* high volatility	[0.0236]					
Board experience		-0.0348***				
* high volatility		[0.0081]				
Female board members			0.8501*			
* high volatility			[0.4580]			
CEO duality				0.2515**		
* high volatility				[0.1098]		
Lock-up					1.3653***	
* high volatility					[0.3627]	
Retained shares						-0.6674
* high volatility						[0.4357]
Obs.	2,468	2,468	2,468	2,468	2,468	2,145
Pseudo R-Squared	0.1061	0.1079	0.1048	0.1057	0.1072	0.0992

Note: Probit model with interaction terms of high volatility and corporate governance characteristics. The dependent variable equals one for withdrawn offerings and zero for completed offerings. High volatility is a binary variable that equals one if the Nasdaq return volatility in the post-filing period is in its highest tertile and zero otherwise. Descriptions of the other variables and data sources are depicted in Table B.1. Apart from the interaction effects, the model corresponds to the model estimated in Table 3.3 specification (5). Coefficients are reported. Pseudo R-squared corresponds to McFadden's pseudo R-squared. Standard errors clustered at industry level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table B.5: VC characteristics and high volatility

	(1)	(2)	(3)	(4)
Corporate governance characteristics				
Board size	-0.0019	-0.0048	-0.0067	-0.0063
	[0.0135]	[0.0139]	[0.0137]	[0.0147]
Board experience	-0.0004	-0.0006	-0.0007	0.0027
	[0.0029]	[0.0031]	[0.0032]	[0.0027]
Female board members	0.0093	-0.0005	0.0051	-0.0338
	[0.2698]	[0.2768]	[0.2733]	[0.2603]
CEO duality	-0.0508	-0.0488	-0.0448	-0.0716
	[0.0859]	[0.0812]	[0.0849]	[0.0834]
Lock-up	-0.5413***	-0.5272***	-0.5426***	-0.5547***
	[0.1778]	[0.1844]	[0.1896]	[0.1694]
Intermediary characteristics				
VC backing	0.1106			
	[0.0772]			
Non-VC		-0.1788	-0.2446*	-0.2168***
		[0.1537]	[0.1322]	[0.0712]
VC syndication		-0.0782		
		[0.1370]		
VC US			-0.1725	
			[0.1916]	
VC reputation				-0.0919*
				[0.0534]
Number of underwriters	-0.0509	-0.0499	-0.0502	0.0030
	[0.0350]	[0.0371]	[0.0369]	[0.0265]
Underwriter market share	-0.8655	-0.8378	-0.8545	-0.7848
	[0.6085]	[0.6192]	[0.6369]	[0.6241]
CM Rank	-0.0190	-0.0207	-0.0207	-0.0206
	[0.0255]	[0.0255]	[0.0256]	[0.0264]
Issuer and issue characteristics				
Filing size (bn USD)	-0.4737	-0.4935	-0.4725	-0.2962
	[0.3758]	[0.3883]	[0.3808]	[0.3644]
Firm size (bn USD)	0.0458	0.0603	0.0615	0.0090
	[0.1229]	[0.1254]	[0.1240]	[0.1171]
Firm age	-0.0933***	-0.0893***	-0.0900***	-0.0735**
	[0.0328]	[0.0331]	[0.0332]	[0.0346]
High-tech	-0.0629	-0.0557	-0.0634	-0.0065
	[0.0704]	[0.0692]	[0.0696]	[0.0652]
Debt ratio	0.1202**	0.1123*	0.1154*	0.1199*
	[0.0579]	[0.0660]	[0.0660]	[0.0640]
BM ratio	0.6916**	0.6743**	0.6995**	0.9263***
	[0.2839]	[0.2811]	[0.2846]	[0.2808]
Asset turnover	0.0344*	0.0354*	0.0394**	0.0341**
	[0.0195]	[0.0188]	[0.0189]	[0.0174]
Negative news	0.0427	0.0301	0.0242	0.0312
	[0.0398]	[0.0437]	[0.0408]	[0.0416]
Market characteristics at filing				
Nasdaq	0.0207***	0.0208***	0.0199***	
	[0.0053]	[0.0052]	[0.0052]	
Implied volatility	0.2526**	0.2637***	0.2499**	
	[0.1027]	[0.1018]	[0.1116]	
Yield spread	0.0180	-0.0075	-0.0131	
	[0.1883]	[0.1848]	[0.1865]	

Number of filings	-2.1927***	-2.2773***	-2.2325***	
	[0.4012]	[0.4032]	[0.3973]	
Market characteristics after filing				
Nasdaq return	-0.1896*	-0.2077*	-0.2028**	-0.1703*
	[0.1108]	[0.1081]	[0.1000]	[0.1034]
high volatility	-0.1250	-0.1132	-0.3195**	-0.2657**
	[0.0969]	[0.2331]	[0.1454]	[0.1324]
Yield spread change	1.1017***	1.0798***	1.1156***	1.2468***
	[0.3601]	[0.3460]	[0.3461]	[0.2955]
Lending standard change	0.0140***	0.0140***	0.0141***	0.0163***
	[0.0018]	[0.0019]	[0.0018]	[0.0016]
Regulation 1	-0.0987	-0.0895	-0.0913	-0.0558
	[0.1888]	[0.1850]	[0.1817]	[0.1665]
Regulation 2	-0.9789	-0.9540	-0.9631	-1.1251
	[0.6402]	[0.6484]	[0.6517]	[0.7777]
Regulation 3	-0.5985***	-0.5987***	-0.5817***	-0.2196***
	[0.0937]	[0.0937]	[0.0943]	[0.0747]
VC-backing	-0.1753			
* high volatility	[0.1391]			
Non-VC		-0.0107	0.2040	0.2718**
* high volatility		[0.2470]	[0.1593]	[0.1262]
VC syndication		-0.2154		
* high volatility		[0.1866]		
VC US			0.0498	
* high volatility			[0.2491]	
VC reputation				0.1398*
* high volatility				[0.0794]
Obs.	2,468	2,435	2,438	2,442
Pseudo R-Squared	0.1049	0.1050	0.1058	0.0951

Note: Probit model with interaction terms of high volatility and VC characteristics. The dependent variable equals one for withdrawn offerings and zero for completed offerings. High volatility is a binary variable that equals one if the Nasdaq return volatility in the post-filing period is in its highest tertile and zero otherwise. Descriptions of the other variables and data sources are depicted in Table B.1. Apart from the interaction effects, the model corresponds to the model estimated in Table 3.3 specification (5). Coefficients are reported. Pseudo R-squared corresponds to McFadden's pseudo R-squared. Standard errors clustered at industry level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

B.2 Figures

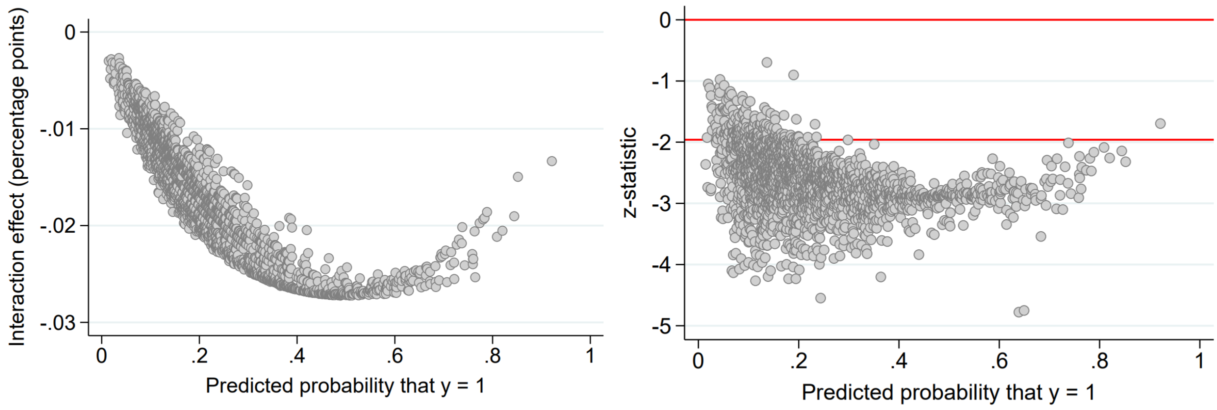


Figure B.1: Interaction effect of board size and high volatility

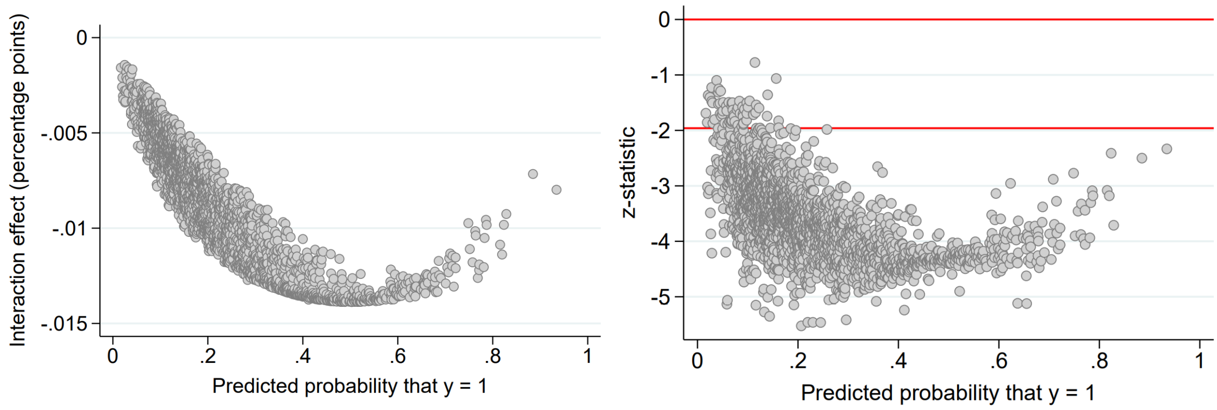


Figure B.2: Interaction effect of board experience and high volatility

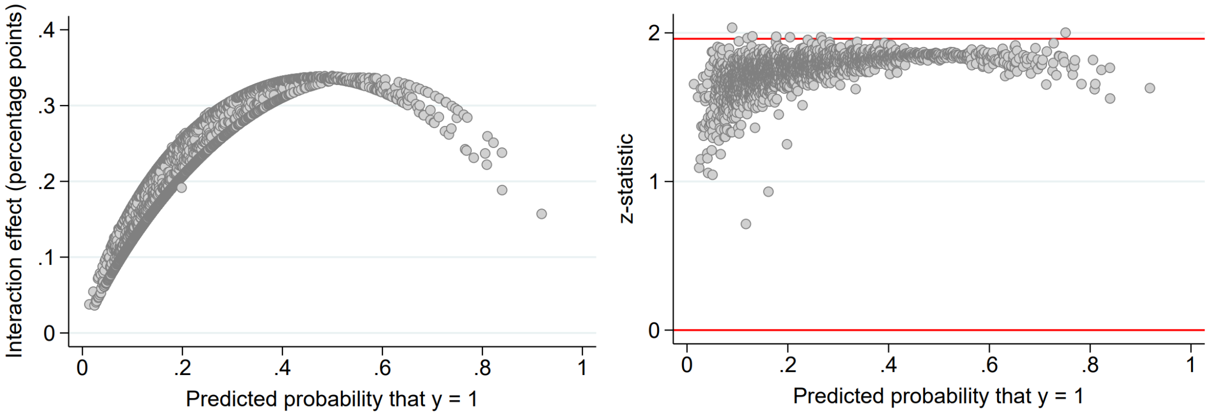


Figure B.3: Interaction effect of female board members and high volatility

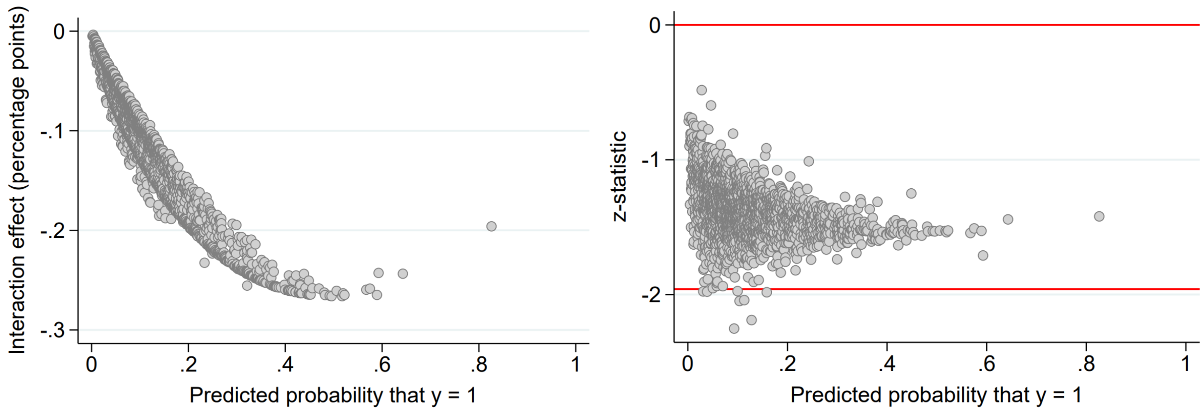


Figure B.4: Interaction effect of retained shares and high volatility

Chapter 4

Lucky coincidence or ominous threat? New evidence on the causal mechanisms behind the effect of IPOs on industry rivals[1]

4.1 Introduction

It is undisputable that an initial public offering (IPO) is a major event in the lifecycle of a firm. Therefore, a vast amount of literature deals with the question of how an IPO affects a firm's short-run (see Aggarwal et al., 2002; Bradley and Jordan, 2002; Ellul and Pagano, 2006; Ljungqvist, 2007; Loughran and Ritter, 2004; Ritter, 1984, among others) and long-run performance after going public (see Aggarwal and Rivoli, 1990; Jain and Kini, 1994; Loughran, 1993; Loughran and Ritter, 1995; Pástor et al., 2009; Ritter, 1991, among others). However, an IPO does not only have a life-changing influence on the firm that goes public but might also impact the performance of industry rivals. Pervious literature mainly discusses two possible effects of an IPO on industry rivals: a competition effect and an information effect. The idea behind the competition effect is that the firm that goes public gains some kind of competitive advantage over its rivals which in turn can lead to an increase in the competitive pressure for firms in the industry. Consequently, this would lead to negative valuation effects on rival firms (Akhigbe et al., 2003). The idea behind the information effect is that an IPO does not only release information about the firm that goes public but also about the industry in which it operates. The majority of studies take the view that an IPO signals good growth prospects for the whole industry and has

[1] This chapter is a single authored manuscript by the candidate. I thank Tereza Tykvová for her significant contribution in developing the idea of the paper.

thus positive valuation effects on rival firms (see Akhigbe et al., 2003; Cotei and Farhat, 2013; Lee et al., 2011). However, some studies argue that an IPO could also foreshadow future negative industry trends (Spiegel and Tookes, 2020) or reveal that the industry is overvalued (Slovin et al., 1995). In this case, an IPO would rather induce negative valuation effects on industry rivals. In sum, the theoretical reasoning concerning the effect of an IPO on rival firms suggests that there is a negative competition effect, whereas the information effect can be either positive or negative.

Previous empirical literature on the valuation effects of IPOs on rivals is rather scarce and reports mixed results. Slovin et al. (1995) find negative share price reactions of rival firms after IPOs and argue that this is due to a negative information effect. Hsu et al. (2010) and McGilvery et al. (2012) also report significant negative valuation effects of IPOs but argue that this finding is due to a negative competition effect. Akhigbe et al. (2003) report insignificant valuation effects of IPOs on rival firms and conclude that this is due to the fact that a negative competition and a positive information effect cancel each other out. In more fine-grained approaches, Lee et al. (2011) find positive valuation effects of IPOs in uncertain industries and negative valuation effects in highly concentrated industries, whereas Cotei and Farhat (2013) find positive valuation effects for venture-backed IPOs and insignificant effects for non-venture-backed IPOs. With regard to the methodological approach, previous research mostly conducts event study analyses to estimate short-term valuation effects of IPOs and focuses on different factors that explain cross-sectional differences in rival's reaction. However, they have not yet developed research designs that allow to explicitly test for the causal mechanisms discussed above. When looking at cumulative abnormal returns obtained in an event study, it should be considered that they reflect the sum of all possible valuation effects of a firm's IPO on industry rivals. For example, negative cumulative abnormal returns could reflect a negative competition effect (as e.g. argued by Hsu et al. (2010) and McGilvery et al. (2012)) but also a negative information effect (as argued by Slovin et al. (1995)). At the same time, cumulative abnormal returns would also be negative if a negative competition effect outweighs a positive information effect. Thus, looking at total valuation effects alone does not allow to make inferences about the underlying mechanisms which drive the effect an IPO might have on rival firms.

The aim of this paper is to contribute to the literature by taking a new methodological approach that allows to actually test for the existence of the competition and information effect. To this end, I employ an identification strategy consisting of two steps. In the first step, I follow previous literature and conduct an event study around IPO filings stemming from a sample of 3.438 US domestic firms that filed between 1997 and 2014. In the second step, I compare the development of the CARs obtained in the event study around two exogenous events in a Difference-in-Difference (DiD) setting. This framework is based on the assumption that the competition and information effect vary with the starting level of competition and information in the industry, respectively. The idea is that the higher the starting level of competition and information, the smaller the influence of an additional IPO in the industry and thus the smaller the competition and information effect. For example, if the competition level in an industry is already high (e.g. under perfect competition) an IPO does not lead to a substantial increase in the competitive pressure for rivals anymore. In contrast, if the competition level is low (e.g. in the monopoly case) an IPO probably leads to a stronger increase in the competitive pressure. The same reasoning applies for the information effect. The lower the information level in an industry, the more substantial the information brought by an IPO. Assuming that this condition holds, rival firms would show a significant change in their reaction to IPOs in their industry after an exogenous change in the competition and information level, respectively. Thus, estimating the effect of exogenous changes in the information and competition level on rival firms' reactions to IPOs can reveal insights about the existence of the two effects. I exploit two exogenous events that either only influence the competition level or the information level in the industry. In order to assess the competition effect, I exploit import tariff reductions (ITRs) as exogenous change in the competition level (see Aktas and Dupire-Declerck, 2015; Flammer, 2015; Frésard, 2010; Frésard and Valta, 2016; Valta, 2012, among others). It is likely to assume that ITRs lead to an exogenous increase in the level of competition in the domestic market by reducing trade barriers. In order to assess the information effect, I exploit broker closures and broker mergers as exogenous events (see Derrien and Kecskés, 2013; Derrien et al., 2016; Guo et al., 2019; He and Tian, 2013; Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012, among others). Broker closures and mergers are both associated with a loss of information and are thus considered to lead to an exogenous decrease in the information level. Both treatments affect the competition and information level only for a certain group of rival firms. This allows to compare the development of CARs by affected firms (treated firms) and un-affected firms (control firms) before and after the change in either the competition or information level.

Results suggest that the competition effect is likely to exist and thus IPO filings seem to harm industry rivals to a certain extent. One explanation could be that IPOs lead to a reassessment of the competitive situation in the industry in favor of the IPO firm. These results are robust over different specifications. In contrast, I find a zero effect in the information effect analysis and thus no explicit evidence for the existence of the information effect.

The remainder of the paper is organized as follows. Section 4.2 outlines the theoretical background of my research design, while section 4.3 provides and overview over previous literature on intra-industry effects of IPOs. The subsequent sections provide a description of data and methodology and present the results of the competition and information effect analysis. Finally, section 4.6 concludes with a discussion on the insights gained and points out directions for future IPO research.

4.2 Theoretical background and related literature

Previous literature mainly discusses two possible effects of an IPO on its industry rivals: a competition and an information effect. From a theoretical point of view, the competition effect predicts a negative valuation effect of an IPO on rivals in the industry. The idea is that an IPO can lead to a reassessment of the competitive situation in the industry in favor of the firm that goes public and thus might increase the competitive pressure of firms in the industry (Akhigbe et al., 2003). There are different approaches that try to explain why the IPO firm might gain a competitive advantage over their industry peers.[2] The probably most obvious argument states that IPO firms gain a competitive advantage as they raise new capital in their IPO which they can use to invest in profitable projects and expansion (Akhigbe et al., 2003) or which provides them the opportunity to lower their financing costs (Spiegel and Tookes, 2020). Chod and Lyandres (2011) add that firms that go public adopt risky and aggressive output market strategies because they are better able to diversify idiosyncratic risk in the capital market. This in turn brings them into a better competitive position. Another argument is based on Stoughton et al. (2001) who states that only better-quality firms go public. They reason that an IPO has a positive impact on the image of a firm and thus consumers perceive the products of listed firms as qualitatively better. In line with that, the high-quality firms have better market prospects than their rivals. Similarly, Akhigbe et al. (2003) note that the IPO increases the visibility of the firm and thus investors assess the competitive position of the IPO firm more positively. Hsu et al. (2010) state that IPO firms might have some attractive non-financial advantage like knowledge capital that can improve their competitive situation. They further argue that firms that go public are recently certified by (top) investment banks, which lead investors to prefer the issues of the IPO firms to issues of other firms in the industry. All in all, it is very likely that a firm gains some kind of competitive advantage over its industry rivals by going public. This in turn can induce a redistribution of market shares away from industry rivals to the IPO firm which coincides with negative price effects for rival firms (Hsu et al., 2010; Lee et al., 2011).

The second potential intra-industry effect that is considered in this study is the

[2]It should be noted that the literature deals with arguments for both a competitive advantage of the IPO firm over private firms and over their publicly traded rivals. As this study estimates intra-industry effects in an event study, its scope is limited to the impact on publicly traded rivals.

information effect. In the US, the Securities Act of 1933 ensures that each firm that goes public has to file a registration statement (prospectus) with the Securities and Exchange Commission (SEC) that contains detailed information about the firm's business operations, financial statements, risk factors and management and has to be published to all possible investors (SEC, 2017). The comprehensive information released during the registration process does not only deliver information about the issuing firm but probably also about the industry in which the firm operates and thus might also affect industry rivals. The argument about the information effect can be based on managerial signaling models (originally developed by Leland and Pyle (1977) and Titman and Trueman (1986)), in which an adverse selection problem arises in the going public process. The models assume that the managers and owners of the issuing firm are better informed than investors. The announcement of an IPO then reveals information about the firm's real asset values. If the IPO announcement also releases industry-specific information, IPOs can decrease information asymmetries in the whole industry and lead investors to update their valuations about industry rivals (see Benveniste et al., 2002; Lee et al., 2011; Slovin et al., 1995; Spence, 1974; Subrahmanyam and Titman, 1999, among others).

However, the direction of the information effect is not straightforward from a theoretical point of view, as it depends on the sort of information that is released. Most studies argue that the IPO serves as a positive signal for the whole industry and therefore causes a positive valuation effect on industry rivals (see Akhigbe et al., 2003; Cotei and Farhat, 2013; Lee et al., 2011, among others). This reasoning is based on Ritter (1991), who states that firms tend to go public if investors are especially optimistic about a certain industry and thus IPOs signal good growth prospects and possibly growing market demand (see Lee et al., 2011). If this holds true, IPOs would be a positive externality for industry rivals and investors would update their expectations about industry rivals upwards, which would be reflected in an increase in rival firms' valuation (see Lee et al., 2011).

Although having obtained less attention in the literature so far, it should be considered that the information effect could also be negative. The first argument is also related to industry factors. Instead of signaling good growth prospects, Spiegel and Tookes (2020) reason that it is possible that there are industry-wide changes (e.g. industry-wide shifts in consumer loyalty for products or industry-wide shifts in profitability) that lead to both a situation in which it is less attractive for firms to remain private and thus induce firms to go public and to a decline in industry rival's performance. In this case, IPOs can be considered to foreshadow future negative industry trends and can thus also be considered as signal for bad industry prospects. This would be in line with negative valuation effects of rival firms after an IPO. The second argument is based on the idea that firms tend to go public if valuations are especially high (see e.g. Benninga et al., 2005; Lerner, 1994; Pagano et al., 1998). Slovin et al. (1995) argue that an IPO in an industry can therefore also reveal that the industry is overvalued. In this situation, an IPO can also be interpreted as unfavorable signal. As a consequence, investors would correct their valuation of industry peers downwards (Akhigbe et al., 2003).

It is likely to assume that the information and competition effect are not equally strong in every context and rather depend on the actual situation in the industry. In particular, I assume that the start level of information and competition in the industry are crucial for the strength of both effects. For example, if the information level in an industry is high and there are almost no information asymmetries, an IPO would not bring about new substantial information that changes investor's valuations about industry peers. In contrast, if the information level in an industry is rather low, an IPO can bring about new information that changes investor's valuations substantially. Thus, I argue that the information effect should be weaker if the information level in an industry is already high before the IPO takes place. The same reasoning applies for the competition effect. In highly competitive industries, it is rather unlikely that the entry of an additional rival changes the actual market concentration substantially. In other words, the increase in the competitive pressure on rival firms is negligible after an IPO if the competition level in the industry is already high. The opposite is true when considering an IPO in an industry with a relatively low initial level of competition (like in a monopoly situation). Here, the market concentration decreases heavily with rivals' entry and the industry will probably be more competitive after an IPO. Following this reasoning, the competition effect should also be weaker if the competition level in the industry is already high before the IPO takes place.

In sum, from a theoretical point of view, there are reasonable arguments for a negative competition effect and either a positive or negative information effect. Moreover, both effects are likely to vary with the start level of information and competition in the industry, respectively.

Empirical evidence on the effect of IPOs on industry peers is still limited and dispersed. Most of them estimate valuation effects using the event study methodology. The first study to be mentioned is the one by Slovin et al. (1995). Using US data from 1980 to 1991, they find negative share price reactions for rival firms and argue that this is due to a negative information effect. Building on their study, Akhigbe et al. (2003) find insignificant average valuation effects of IPO filings on rival firms in a sample of 2,493 US IPO filings between 1989 and 2000. They explain their finding by a positive information and a negative competition effect that cancel each other out. Further, they find positive valuation effects for the first IPO in an industry and for IPOs in regulated industries. In contrast, they report negative valuation effects for large IPOs in highly competitive or risky industries and IPOs in better-performing or technology industries. Hsu et al. (2010) analyze the competition effect of IPOs using a sample of 4,188 US firms that went public between 1980 and 2001. While they find negative valuation effects and a decline in rival's operating performance after completed IPOs, they report positive valuation effects after IPO withdrawals. In addition, they report that rivals in a better competitive situation, which is especially reflected in less leverage, higher spending on research and development and certification by top investment banks, suffer less after an IPO in their industry. Lee et al. (2011) focus on intra-industry effects of IPOs in growing and uncertain industries using a sample of 48 US IPOs between 2000 and 2004. They find positive valuation effects of IPOs in uncertain industries and negative valuation effects for firms operating in highly concentrated industries. Cotei and Farhat (2013) highlight the role of venture capital backing in intra-industry effects of IPOs and argue that venture-backed IPOs deliver superior information and thus for those firms the information effect should be stronger. Using a sample of 3,810 US firms that went public between 1983 and 2001, they find positive valuation effects for venture-backed IPOs and no significant effect for non-venture-backed IPOs. McGilvery et al. (2012) analyze intra-industry valuation effects of 106 Australian IPOs in the period between 1999 and 2009, thereby focusing on the role of the IPO firm's governance profile and the intended use of their offer proceeds. They find negative valuation effects and conclude that this is due to the prevalence of a competition effect. Further, their results suggest that the negative valuation effects get stronger if firms declare to use their proceeds for debt reduction or investment. In contrast, a larger board size and CEO ownership are negatively associated to rival firms' abnormal returns. Lately, Spiegel and Tookes (2020) analyze the evolution of rival's post-IPO performance using a dynamic structural oligopoly model, in which they examine each IPO individually in a sample of US data including the period from 1983 to 2012. They argue that the observed performance decline of rivals after an IPO in their industry is due to the fact that IPOs are triggered by industry trends that also lead to changes in rivals' performance.

In a broader sense, this research is further related to literature on the effect of different firm-specific events on industry peers. One of the first of those studies is the one by Firth (1976), who finds positive effects of earning announcements on firms in the same industry, which he attributes to intra-industry information transfers. A large amount of literature analyzes intra-industry effects of bankruptcy announcements, which also discusses the implications of possible information transfers and competition effects (see Cheng and McDonald, 1996; Ferris et al., 1997; Iqbal, 2002; Jorion and Zhang, 2007; Lang and Stulz, 1992, among others). Further examples in the literature include Slovin et al. (1991) who report positive valuation effects of going-private transactions on industry peers, Hertzel (1991) who focus on intra-industry effects of stock repurchase announcements, or

Slovin et al. (1992) who report negative valuation effects of bank common stock issuance announcements on rival banks.

To sum up, there is a growing literature on the effect of firm-events on industry rivals during the last years. With regard to the effect of IPOs on industry rivals, previous literature (except for Spiegel and Tookes (2020)) mainly conducts event study analyses to examine whether IPOs have positive or negative valuation effects on rivals and identify different factors that determine rival firms' cumulative abnormal returns. Based on their findings, they argue in favor of the prevalence of an information effect, a competition effect or both. However, it should be considered that cumulative abnormal returns reflect all valuation effects a firm's IPO might have on industry rivals. In this regard, the direction of the valuation effects as such cannot be regarded as an explicit test for the causal mechanisms behind this effect. To the best of my knowledge, no study exists that tries to disentangle the underlying mechanisms an IPO has on industry rivals and actually tests for the existence of a competition and information effect. I therefore contribute to the literature by applying a new methodology that allows to test for the existence of both effects, thereby aiming to provide new insights on the underlying mechanism of how an IPO might affect rival's valuation.

4.3 Data and methodology

4.3.1 Data

The IPO sample comprises information on first-time filings of US domestic firms between 1997 and 2014, which stems from Thomson Reuters Securities Data Company (SDC) Platinum. Following IPO literature (see Busaba et al., 2001; Ritter and Welch, 2002, among others), American depositary receipts, convertible issues, unit offerings, closed-end funds, REITs, limited partnerships, small best effort offers, SPACs, issuers that are not seeking a listing on NYSE, NASDAQ or any other American exchange and financial firms are excluded from the sample. Filing dates, withdrawal dates and IPO dates are checked against EDGAR filings and observations are removed from the sample if these dates do not match. After these steps the sample comprises 3.438 filings. All variables and data sources are also summarized in Table C.1 in the appendix.

When analyzing intra-industry effects, the definition of rival firms is crucial. Usually, studies in this research area define rivals as firms with the same SIC code. The narrowest industry definition uses 4-digit SIC codes (used by e.g. Slovin et al. (1995)), while 2-digit SIC codes reflect a broader definition (used by e.g. Hsu et al. (2010)). I choose an intermediate solution and define rivals as firms in an industry with the same 3-digit SIC code. On the one hand, in order to identify direct rivals, the industry definition should be chosen as fine as possible (Lee et al., 2011). On the other hand, Kahle and Walkling (1996) claim that SIC codes tend to differ more across data sources the finer they are. Using 3-digit SIC codes partly accounts for both arguments. The sample of rival firms stems from CRSP and initially consists of 25.670 listed firms. SIC codes are primarily taken from Compustat and supplemented with and checked against data stemming from SDC Platinum and S&P Capital IQ. I limit the sample to rival firms that are listed on AMEX, NASDAQ or NYSE and only include common shares. After these steps, 14.386 potential rival firms remain in the sample.

4.3.2 Methodology: First step – Event study

The applied methodology is new in this strand of literature and allows to test for the existence of the competition and the information effect. Stated otherwise, it contributes to the literature by analyzing whether the reactions of industry rivals' returns to IPOs can be explained by a (positive or negative) information effect, the competition effect or a combination of both effects. In order to shed light on this question, the crucial part of the empirical analysis is to isolate both effects. For this purpose, I use an identification strategy comparable to the one applied by Aktas and Dupire-Declerck (2015) who analyze the effect of increased entry threat on industry merger activity. The identification strategy consists of two steps: an event study in the first step and a Difference-in-Difference (DiD) analysis in the second step. In order to increase the likelihood of the common trend assumption of the DiD approach to hold, control firms are matched to treatment firms via mahalanobis matching on different pre-treatment IPO and firm characteristics stemming from Compustat, S&P Capital IQ and EDGAR filings.

In the first step, I follow previous research in this field (see Akhigbe et al., 2003; Hsu et al., 2010; Lee et al., 2011; McGilvery et al., 2012, among others) and conduct an event study analysis (based onMcWilliams and Siegel (1997)) to estimate the short-term reaction of industry rivals on IPO announcements. As the prospectus is considered to be the main source of information, I consider the official filing date as the crucial date for information release on which the rivals react and thus use it as announcement date in the event study. In order to calculate abnormal returns, a market model is estimated by OLS over an estimation period (-218, -39) before the announcement.

Abnormal returns are calculated with the following equation:

$$AR_{it} = R_{it} - (\alpha_i + \beta_i R_{mt}), \tag{4.1}$$

where AR_{it} are the abnormal returns of rival firm i at time t, R_{it} is the actual return of rival firm i at time t and R_{mt} is the market return at time t, for which I use the value weighted S&P 500 index. Further, α_i and β_i are the estimates obtained from the market model (see McWilliams and Siegel, 1997).

Cumulative abnormal returns are then calculated by summing up the daily abnormal return over different event windows:

$$CAR_i = \sum_{t=1}^{t} AR_{i,t} \tag{4.2}$$

(see Brooks, 2014; McGilvery et al., 2012, among others).

The analysis by Hsu et al. (2010) suggests that there is already a price reaction about ten days before the actual event. Thus, the event window should contain a sufficient number of days before the event. Different event windows are considered in the analysis below. If a firm reacts to more than one IPO filing within the event window, these observations are excluded. This procedure ensures that the actual price reaction of this IPO filing is measured and that it is not confounded by another IPO filing. A further requirement is that there is no event (IPO filing) in the estimation period. In order to test whether CARs are statistically significant different from zero, the method suggested by Brooks (2014) is applied.

4.3.3 Methodology: Second step – Difference-in-Difference

In the second step, I compare the development of the CARs obtained in the first step around two exogenous experiments in a DiD setting. The treatments either influence only the competition or the information level for a sub-group of rival firms in the industry (treatment group), whereas another group of rival firms remains unaffected by the treatments (control group). In the DiD setting, the effect of the treatments can therefore be estimated by comparing the development of both groups from before to after the treatment. The main idea of this step is to test whether the strength of the competition or information effect changes due to an exogenous change in the competition or information level. As explained in the theoretical section above, it is likely to assume that the strength of the competition and information effect vary with the start level of competition and information in the industry respectively. If this crucial assumption holds and the respective effects do exist, an exogenous variation in the competition and information level should lead to a change in the strength of the competition and information effect, as reflected in a change in rival's reaction to IPOs (measured by CARs). Stated otherwise, a change in rival firm's reaction to IPOs due to the treatment (reflected in a significant treatment effect) can be considered as evidence for the existence of both effects. Using two different exogenous experiments that are assumed to either influence only the information or competition level allows to consider both effects in isolation.

In order to estimate the treatment effect, I employ the fixed effects implementation of the DiD estimator:

$$CAR_{ist} = \beta_1 treat_{ist} + \eta_{is} + \gamma_t + X'\beta, \tag{4.3}$$

where CAR_{ist} is the cumulative abnormal return of rival firm i at treatment s at time t. The estimated treatment effect is indicated by β_1, η_{is} are firm-treatment fixed effects for firm i at treatment s, γ_t are year fixed effects and X' is a row vector of a set of covariates. The next subsections describe the exogenous experiments that are used for the analysis.

Competition effect: Cuts in import tariffs

In order to test for the competition effect, I exploit import tariff reductions (ITRs) as natural experiments that are assumed to only influence the competition level in the market but leave the information level unchanged. Exploiting ITRs has been established in empirical finance in several contexts beyond IPO research as exogenous change in competition (see Aktas and Dupire-Declerck, 2015; Flammer, 2015; Frésard, 2010; Frésard and Valta, 2016; Valta, 2012, among others). They argue that cutting import tariffs reduces the costs of market entry for rival firms and thus creates exogenous variation in the likelihood of entry by rival firms. According to e.g. Bernard et al. (2006) a reduction in market entry barriers in form of ITRs can lead to a reallocation of market share away from domestic firms to foreign firms. Following this reasoning, ITRs can be considered as exogenous change in competition and should consequently lead to an increase in competitive pressure for firms in the market (see Aktas and Dupire-Declerck, 2015; Frésard, 2010; Frésard and Valta, 2016; Valta, 2012, among others). In line with that, I expect the level of competition in an industry to be higher after the cut in import tariffs than before the cut takes place. As explained above, the strength of the competition effect is likely to differ depending on the competition level in the industry. More precisely, I argue that the competition effect is weaker, the higher the initial competition level in the industry. In line with that, the strength of the potential competition effect should also be weaker after the ITR (which exogenously increase the competition level). As the competition effect itself is assumed to be negative, a weaker competition effect implies more positive CARs, which is in line with a positive treatment effect. It should be noted that the CARs reflect all reactions to IPO filings and not just the reaction that is due to the competition effect. Thus, CARs itself can be positive or negative before and after the ITR. However, if the competition effect does exist and varies with the starting level of competition, CARs should be more positive after the treatment (for treated firms). As the information level is likely to be unaffected by the ITRs, the information effect is assumed to be constant before and after the treatment and thus should not change the CARs.

Following Frésard (2010) and Frésard and Valta (2016) and using data provided by Peter Schott's Website, the import tariff reductions are calculated in the following way. First, the ad-valorem tariff rate is calculated as the duties collected by US Customs divided by the Free-on-Board value of imports for each industry year. Secondly, the change in tariff cuts per year is calculated. A tariff cut is then defined as a negative change in import tariffs that is two times larger than the industry's median change. The median is chosen instead of the mean as it is less sensitive to outliers. Robustness checks are reported in which a cut is defined to be three times larger than the industry median (see Aktas and Dupire-Declerck, 2015; Frésard, 2010; Frésard and Valta, 2016). Tariff cuts that are smaller than 1% and cuts that are followed by large tariff increases in the following three years are excluded (see Frésard, 2010). As tariff data is only available for manufacturing industries (SIC codes 200-399), the analyses focus on these industries. This also leads to a reduction in the sample of potential rivals to 6.057 firms.

It should be noted, that an ITR always affects all firms in one industry. Thus, there is no intra-industry variation between firms. The considered observation period is three years before and three years after the ITRs, thus cuts are required to lie at least seven years apart to be included in the sample. Table C.2 in the appendix provide an overview over the cuts that are included in the analysis for the different definitions of cuts, respectively.

Following this reasoning, the treatment group consists of all rival firms that operate in an industry in which an ITR takes place in the considered time period, while the control sample consists of firms in another manufacturing industry that are not subject to an ITR over the considered time period. Both treatment and control firms have to react to at least one IPO filing before and one IPO filing after the ITR.[3] Data is cumulated at the year level.

Information effect: Broker closures and mergers

In order to test for the information effect, I exploit broker closures and broker mergers as natural experiments. I assume that these events only influence the information level in the market but leave the competition level unchanged. Kelly and Ljungqvist (2012) were the first who exploited broker closures as natural experiments and Hong and Kacperczyk (2010) introduced broker mergers. Both events have been used in subsequent studies to analyze the effect of external shocks in analyst coverage on different outcomes (see Derrien and Kecskés, 2013; Derrien et al., 2016; Guo et al., 2019; He and Tian, 2013, among others). The underlying idea is that financial analysts are important players in the market that produce valuable information for investors and firms by e.g. evaluating the actual

[3]Note that these conditions lead to a further reduction in the number of rival firms and filings that can be used for the analysis. Details on the number of firms and filings used in each analysis are reported in the respective analyses below.

performance of firms or making forecasts about their future performance (see Asquith et al., 2005; Brennan and Subrahmanyam, 1995; He and Tian, 2013; Hong and Kubik, 2003; Jegadeesh et al., 2004; Kothari et al., 2016; Loh and Stulz, 2011; Womack, 1996, among others). Broker closures and broker mergers are both events that are accompanied with terminations of analysts and thus lead to a loss of analysts for the firms they previously covered (Derrien et al., 2016). While this is straightforward for broker closures, it is less clear for broker mergers because it is not necessarily the case that the merged firm employs less analysts than the two separate firms before the merger. However, Wu and Zang (2009) and Hong and Kacperczyk (2010) argue that broker firms usually dismiss analysts due to redundancy (e.g. if acquirer and target followed the same firm before) and lose analysts also due to merger turmoil and cultural clash after a merger. As a consequence, the merged broker firm employs less analysts than acquirer and target before the merger. This loss in analysts (through broker closure or broker merger) would lead to an increase in information asymmetries and consequently to a reduction of information that is publicly available. It is possible that a reduction in analyst coverage either leads to a total loss of information or that information becomes private. Both channels should be associated with an overall lower level of (publicly available) information in the market (Kelly and Ljungqvist, 2012). Kelly and Ljungqvist (2012) show that broker closures are indeed correlated with an increase in different measures for information asymmetries and thus broker closures should be a suitable measure for an increase in information asymmetries (see Derrien and Kecskés, 2013). Derrien and Kecskés (2013) add that research by analysts that disappear are not qualitatively worse than research produced by analysts that do not disappear after broker closures or mergers. This finding further strengthens the argument that there is actually less information available after the merger or closure and that the loss of analysts is not compensated by other analysts that produce qualitatively better information. It is important to note that broker closures and mergers are mainly due to the business strategy of the broker firm rather than to characteristics of the firms that are covered and can thus be considered as exogenous shocks for the firms they cover (see Derrien et al., 2016; Derrien and Kecskés, 2013; Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012, among others).

Similar to the analysis of the competition effect, the crucial assumption when using broker closures and mergers to analyze the information effect is that the strength of the information effect varies systematically with the start level of information in the industry. As explained in the theoretical section, I assume the information effect to be stronger, the lower the initial information level in the market. This means that after a decrease in information due to a broker closure or merger, an additional IPO in an industry gains in importance and thus the information effect should be stronger after the treatment. If the crucial assumption holds, a significant estimated treatment effect would strongly suggest that there is an information effect. However, it should be noted that the direction of the treatment effect is less clear as the information effect can either be positive or negative. Given that the competition level in the industry remains unaffected by the broker closure and merger, the CARs would become more positive after the treatment if the information effect is positive. This would be in line with a positive treatment effect. In contrast, if the IPO filing delivers negative information about the industry (negative information effect), CARs should become more negative after the treatment, which would be in line with a negative treatment effect. In sum, the direction of the treatment effect is less clear, whereas significant treatment effects would still imply the existence of the information effect. At the same time, it should be noted that a zero estimated treatment effect does not necessarily imply that there is no information effect. It is e.g. possible that some IPOs have a positive information effect while others have a negative one and that both effects cancel each other out.

Analyst information is gathered by I/B/E/S, which allows to identify the brokers that follow the firms in the sample. A broker is defined to follow a firm if there is at least one estimate by a broker about the firm before the event. After identifying the broker, I checked whether these brokers are involved in any broker merger or closure on S&P Capital IQ and on the website by the Financial Industry Regulatory Authority (FINRA). If brokers are involved in closures or mergers, I further checked whether the firms actually lost an analyst due to this event, that is, whether the broker actually stopped covering the respective firms. The advantage of this procedure is that it allows to identify affected firms on a firm-level basis.

In order to make the analysis as similar as possible to the analysis of the competition effect, data is also cumulated at year level and I also choose an observation period of three years before and after the treatment. However, there are also some differences. Firstly, this analysis is not limited to the manufacturing industry. However, following Spiegel and Tookes (2020) I exclude the finance and utility industries as firms operating in these industries probably react differently. Secondly, information on analyst coverage is available on the firm-level and thus it is possible to identify exactly which firm lost an analyst due to a broker closure or merger, which leads to the fact that there is intra-industry variation.

Following this reasoning, treatment firms are defined as firms that lose an analyst due to a broker closure or merger. In addition, they have to react to at least one IPO in the three years before the loss and to an IPO in the three years after the loss. Control firms are defined as firms that operate in the same industry (based on 3-digit SIC codes) and react to exactly the same IPOs in the industry like the treatment firm. Control firms are not allowed to be treated in the before and after period but are allowed to be treated at another point in time.

4.3.4 Methodology: Pre-matching

The crucial assumption for DiD analyses to be valid is the common trend assumption. This means that in the absence of the treatment, treatment and control group would have evolved in the same way. As diverging trends can be caused by differences in covariates between treatment and control group (e.g. if some unobserved determinants interact with observable covariates), the common trend assumption is more likely to hold when covariates are similar between treatment and control group. In the case at hand, this means that both characteristics of the rival firms and the IPOs to which the rivals react should be quite similar. In order to ensure that treatment and control firms are as similar as possible, control firms are matched to treatment firms via mahalanobis matching before conducting the DiD analysis (see Derrien et al., 2016).

As matching variables, I consider variables that can be expected to have a major influence on the CARs. They are measured pre-treatment and on a yearly basis. On the one hand, rival firms' reaction to IPOs in their industry depend on the competitive strength of the rival firm. Therefore, I choose variables that are assumed to somehow influence the competitive situation of the rival firm. Variables that are assumed to foster the competitive situation and are thus associated with a less strong reaction of rivals are firm size (see e.g. McGilvery et al., 2012), a firm's performance (measured as asset turnover) (see e.g. Hsu et al., 2010), a firm's growth prospects (measured as the market-to-book ratio) (see e.g. Cotei and Farhat, 2013; McGilvery et al., 2012), and a firm's R&D intensity as these firms might be more innovative (see e.g. Hsu et al., 2010). In addition, I use a rival's debt ratio, which presumably weakens the competitive situation of the rival and would thus be associated with a stronger reaction to IPOs.

On the other hand, rival firm's reaction on IPOs might also depend on characteristics of the IPO firm and the IPO itself.[4] IPO and IPO firm characteristics that are associated with a stronger reaction of rival firms, as they are assumed to increase the IPO firm's power to compete are firm size, firm age (see e.g. Lee et al., 2011; McGilvery et al., 2012), and certification by a high reputable underwriter (measured as the Carter-Manaster (CM) rank). In addition, larger IPOs have probably a stronger effect on rival firms than smaller IPOs, why I add filing size (see e.g. Hsu et al., 2010) to the set of matching variables. Finally, I also add the debt ratio of the IPO firm, which is associated with a less strong reaction of rivals. Applying mahalanobis matching requires a sample of non-missing matching variables, which further reduces the number of observations.

[4] As in the analysis of the information effect it is already required that treatment and control firms react to the same IPO, these matching variables are only used in the analysis of the competition effect.

I decide to include the matching variables also as parametric controls in some specifications in the DiD analysis, as the pre-matching reduces differences in covariates between groups but allows certain differences to remain. Therefore, I check the robustness of the results by running the analysis with and without control variables. To further assess the validity of the research design and especially the common trend assumption, I conduct different placebo tests in order to test for diverging trends before the treatment in the following way: I reduce the sample to the pre-treatment period when neither treatment nor control group have experienced a treatment, but define certain time spans as placebo-treatment period. If the placebo-DiD estimation leads to significant treatment effects, this would suggest diverging trends already before the treatment. In this case, the common trend assumption would be unlikely to hold.

4.4 Results

4.4.1 Competition effect

Table 4.1 reports event study results, obtained in the first step in the competition effect analysis. It shows cumulative abnormal returns in percent for different event windows and the respective p-values indicating whether the CARs are statistically significant different from zero. It should be noted that in the pre-matching stage, control firms are matched to the treatment firms by matching with replacement and therefore control firms are allowed to show up more than once in the sample. As each firm is allowed to have a maximum of five control firms, the number of control firms reported in the table is much higher than the number of treated firms. Control firms that are chosen more than once are weighted according to their frequency of occurrence in the table. One observation reflects a rival firm's reaction to one IPO. Remember, that each firm has to react to at least two IPOs (one before and one after the treatment) to be included in the sample.

Table 4.1: Competition effect: Event study results

Event window	CAR (mean)	p-value	CAR (median)	p-value	Obs
[-1, 1]	-0.39	0.00	-0.52	0.00	26,777
[-2, 2]	-0.50	0.00	-0.61	0.00	26,777
[-5, 5]	-1.09	0.00	-1.21	0.00	26,777
[-10, 1]	-1.46	0.00	-1.44	0.00	26,777
[-10, 5]	-1.67	0.00	-1.66	0.00	26,777
[-10, 10]	-2.80	0.00	-2.58	0.00	26,777

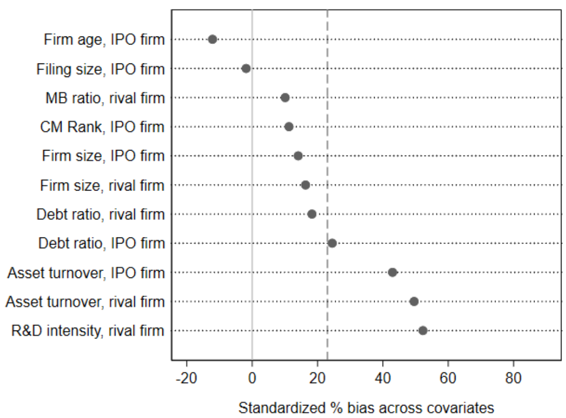
Note: Event study results of the rival firms included in the sample to analyse the competition effect. Cumulative abnormal returns (CARs) are reported in percent for different event windows. P-values of t-tests are reported that indicate whether the CARs are statistically different from zero.

While the CARs are relatively small for narrow event windows around the announcement dates (e.g. -0.13% for the narrowest event window of [-1;1]), the price reaction gets stronger when taking the pre-event price reaction into account (mean CAR jumps to -1.46% when also including the 10 days before the announcement). This suggests that there is also some information leakage somedays before the announcement date which should be taken into account. This finding is in line with Hsu et al. (2010) who also states that it is important to account for the pre-event price reaction. In addition, it should be noted that rival's reactions are secondary effects which generally take longer to become observable. Thus, the event window should also contain some days after the event. Figure C.1 in the appendix illustrates the development of the mean cumulative abnormal returns from 10 days before to 10 days after the event date. It supports the view that the reaction already starts about 10 days before the announcement date and lasts until several days after the event. Therefore, I use the results of the event window of [-10;10] as preferred one in the following DiD analysis. Robustness tests are reported for the [-5;5] event window. The mean cumulative abnormal return of the [-10;10] event window is -2.80% and statistically significant, indicating an overall negative reaction of rival firms to IPOs in the same industry. Overall, the results of the event study analysis are comparable to the results of the majority of studies in this research area (e.g Hsu et al. (2010) report statistically significant mean CARs of -0.82% and -1.00% for the [-10;10] event window using US data between 1980 and 2001 and McGilvery et al. (2012) report mean CARs of -2.95% for a [-6;6] event window using data on Australian IPOs between 1999 and 2009).

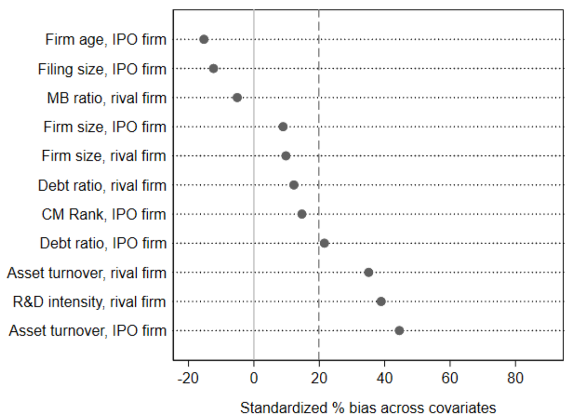
The second step of the analysis turns to the DiD analysis on the effect of ITRs on CARs. In order to increase the likeliness of the common trend assumption of the DiD analysis to hold, control firms are matched to treatment firms in a pre-stage to the DiD analysis. Table C.3 shows descriptive statistics of the variables that are used to match control firms to treatment firms. It should be noted that firms are matched via mahalanobis matching and not via exact matching. Therefore it is possible that some differences between the variables remain. Figures 4.1a to 4.1c show the calculated standardized bias for all covariates in percent for samples that are matched in different ways. The standardized bias measures the differences between covariates of the treatment and control sample and should thus be minimized. The dashed vertical line indicates the mean standardized bias. Figure 4.1a shows the bias of the basic sample in which treatment and control firms are matched via mahalanobis matching and in which each treatment firm is allowed to have a maximum of five control firms. The reported mean bias for this sample is 23.03%. In line with the descriptive statistics in Table C.3, the variables asset turnover of rival firm, asset turnover of the IPO firm and the rival firm's R&D intensity are the variables with the biggest differences. Figure 4.1b shows the standardized mean bias of a sample in which only the nearest matched rival of each treatment firm is included as control firm. As expected, the mean standardized bias decreases slightly to 19.85% in this sample. Figure 4.1c shows the standardized bias for a sample in which five control firms are chosen randomly out of a set of possible control firms. In this sample, the mean standardized bias increases substantially to 35.44%. The decrease in the standardized bias from the randomly matched sample to the sample created by mahalanobis matching implies that the applied matching procedure is indeed useful to make treatment and control group more similar to each other and make parametric controls less prone to error due to functional form misspecification (see e.g. Imbens and Rubin, 2015).

Table 4.2 shows the main results of the DiD analysis around ITRs, where a cut in import tariff is defined as a negative tariff change that is two times larger than the industry's median change. For each specification, the number of treated and control firms is reported, while control firms are allowed to be matched to different treatment firms and are thus allowed to show up more than once in the sample. The mean number of treatment per control firms is also reported in the bottom part of the table. The dependent variable is

(a) Mahalanobis matching (5 matches)



(b) Mahalanobis matching (1 match)



(c) Randomly chosen control sample

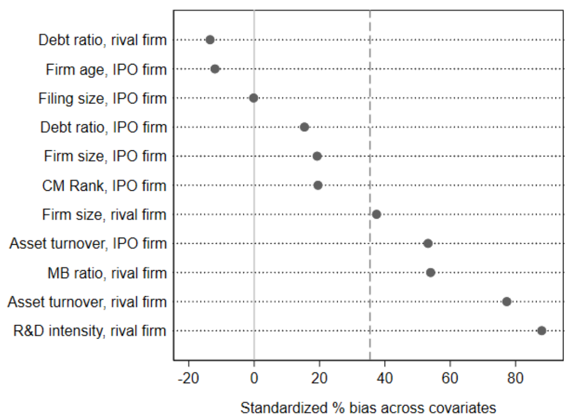


Figure 4.1: Competition effect samples: standarized biases

the rival firm's CAR over the [-10; 10] event window, estimated in the first stage. The sample includes all reactions of a rival firm to all IPO filings in its industry in the period from three years before to three years after the treatment. The number of IPO filings to which the firms react are also reported in the table for each specification.

The first specification does not include control variables and delivers a highly significant and positive treatment effect of 0.0774. This result holds when adding control variables in specification (2). A positive treatment effect means that the CARs become more positive after the treatment. As outlined in the theory section above, the ITRs are assumed to increase the competition level. Moreover, the competition effect of an IPO filing is assumed to be negative and weaker if the competition level is higher. Taking these considerations into account, the CARs are expected to be more positive after the ITRs. In this regard, the positive treatment effect confirms the theoretical expectation. Stated otherwise, the competition effect gets weaker once competition increases through an exogenous event. Given that the information level is likely to be unchanged in this setting, the more positive reaction to IPO filings of rivals after the ITRs compared to rivals' reaction in industries without an ITR can be considered as evidence for the existence of the competition effect. Thus, IPO filings seem to harm industry rivals to a certain extend. As the significance tests indicate, the treatment effect is highly significant from a statistical point of view. To further assess the economic significance of the results, I calculate a standardized measure of the treatment effect. For this purpose, I divide the treatment effect by the pre-treatment standard deviation of the dependent variable (standardized treatment effect). The treatment effect estimated in specification (1) corresponds to a treatment effect of 0.4733 standard deviations (and 0.5505 standard deviations in specification (2)), indicating that the treatment effect is also rather strong in substantive terms.

As outlined in the theory section, the framework is based on the assumption that the competition effect varies systematically with the start level of competition in the industry. More precisely, the competition effect is assumed to be stronger the lower the initial competition level in the industry. In order to substantiate this assumption, the sample is split into two sub-samples according to the pre-treatment level of competition in the respective industry. If the assumption holds, the treatment effect should be weaker for highly competitive industries, as the additional increase of competition induced by the treatment is of less importance compared to industries with low starting level of competition.

In order to measure the pre-treatment competition level in the industry, I use the Herfindahl-Hirschman Index (HHI) which is a measure for industry concentration. In this connection, a high industry concentration is line with a low level of competition. Industries are defined to have a high level of competition (low industry concentration) if the HHI is smaller than 0.15 and they are defined to have a low level of competition (high industry concentration) if the HHI is larger or equal to 0.15 (see e.g. Brezina et al., 2016). Treatment firms need to have at least one control firm with the same pre-treatment competition level to be included in the analysis. Note that this condition leads to a further reduction in the number of observations for this sub-analysis.

Specification (3) includes firms operating in industries with a low pre-treatment competition level and specification (4) includes firms in industries with a high pre-treatment competition level. Indeed, the treatment effect is positive (0.0873) and both statistically significant and substantive (standardized treatment effect of 0.5674) for rivals in industries with a low pre-treatment competition level (specification (3)). In contrast, the treatment effect becomes insignificant and nearly zero (standardized treatment effect of -0.0332) in industries with an already high level of competition (specification 4). Thus, the results indeed seem to be driven by firms operating in industries with a rather low initial competition level. This suggests that an additional increase of the competition level is more important when the starting level of competition is low. In addition, these findings imply that the crucial assumption of the identification strategy – that the competition effect varies systematically with the start level of competition in the industry – seems to hold.

Specification (5) and (6) test the sensitivity of the results to the exclusion of different firms. Specification (5) excludes rival firms that have an own stock event in the time around the IPO. It is possible that firms that have an own stock event react differently to IPO filings in their industry. The stock events that are accounted for are Mergers and Acquisitions (rival firm is either acquirer or target), liquidations, reorganizations, and issuances, for which information is gathered by Thomson Reuters SDC Platinum and CRSP. The exclusion of these firms leads to a reduction in the number of treated firms from 436 to 344. The estimated treatment effect remains positive and highly significant or, stated otherwise, stays robust to the exclusion of rival firms that experience an own stock event around the IPO filing. Therefore, results do not seem to be driven by firms that have an own stock event in the considered time period. Specification (6) excludes IPO filings that are withdrawn later on, which reduces the number of treated firms to 398 and the number of IPO filings included in the analysis to 120. Hsu et al. (2010) show that withdrawn IPOs have an opposite effect on rivals than successful IPOs, that is, rival firm's reaction to an IPO withdrawal is positive. If the information about the outcome of the IPO filing (issue or withdrawal) becomes already available during the event window, the effect is likely to be weaker compared to successful IPOs. More precisely, the estimated treatment effect would be underestimated by including these IPO filings, as the effect of these filings cancel part of the treatment effect out. The estimated treatment effect is again positive, of high economic (standardized treatment effect of 0.5885) and statistical significance, suggesting that including filings that are withdrawn later on do not lead to an underestimation of the treatment effect.

Further robustness tests related to the estimation technique are reported in Table 4.3. Firstly, I test the robustness of the results when varying the event window over which the dependent variable is calculated in the first step of the analysis. Instead of using the rival firm's CAR over the [-10; 10] event window, specification (1) uses CARs calculated over the [-5;5] event window as dependent variable. The estimated treatment effect stays positive and statistically significant but becomes slightly weaker in substantive terms (standardized treatment effect decreases to 0.3233). However, the treatment effect can still be considered to be robust to a change in the event window used in the event study.

Secondly, specification (2) and (3) test the robustness of the results to a different choice of the control firms, as the choice of rival firms is crucial for the analysis. Instead of using the five nearest matches of each treatment firm as control firms, specification (2) performs the analysis with only using the nearest match as control firm. Compared to using five matches, this usually leads to a lower number of observations (i.e. a higher variance) but to a better matching quality (i.e. lower bias). Specification (3) choses five control firms randomly out of the set of possible controls without using matching to compare the results from a pre-matched sample with a purely parametric approach. As already discussed above, the mean standardized bias is higher for the randomly matched control sample than for the samples that apply mahalanobis matching. For both specifications ((2) and (3)), the estimated treatment effect remains positive and statistically significant (although the standardized treatment effect decreases slightly (to 0.3698 in specification (2) and to 0.5264 in specification (3)), implying that the choice of the estimation technique does not influence the results to a large extent.

Specifications (4) to (6) report the results of different placebo tests in order to evaluate the validity of the common trend assumption. In order to test for diverging trends before the treatment, the sample is reduced to the pre-treatment period when neither treatment nor control group have experienced an ITR. It should be noted that the placebo tests can only be performed for firms that react to at least two IPOs in the pre-treatment period. In specification (4), the last observation in the pre-treatment period is pretended to be treated and all others are pretended to be untreated. In specification (5), the first half of the observations are pretended to be treated and the other half of the observations are pretended to be untreated. Finally, in specification (6), the first observation is pretended to be untreated and all other observations are pretended to be treated. All three placebo tests are insignificant and close to zero (ranging between -0.0235 and 0.0245 or in terms of standardized treatment effects ranging from -0.1602 to 0.1671). This finding suggests

Table 4.2: Competition effect: Difference-in-Difference results

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.0774***	0.0865***	0.0873**	-0.0051	0.0729***	0.1016***
	[0.0173]	[0.0171]	[0.0342]	[0.0134]	[0.0223]	[0.0208]
IPO and IPO firm characteristics						
Firm size		0.0436	0.0046	0.1248***	0.0524	0.2661***
		[0.0313]	[0.0711]	[0.0413]	[0.0386]	[0.0835]
Asset turnover		0.0157*	0.0333**	0.0068	0.0278***	0.0060
		[0.0092]	[0.0129]	[0.0090]	[0.0095]	[0.0080]
Debt ratio		0.0123	-0.0027	0.0001	0.0084	0.0243**
		[0.0089]	[0.0178]	[0.0155]	[0.0110]	[0.0112]
Firm age		-0.0010	-0.0003	-0.0012**	-0.0008	-0.0015**
		[0.0006]	[0.0009]	[0.0005]	[0.0007]	[0.0008]
Filing size		-0.0560	-0.0067	-0.0825	-0.1162	-0.4620***
		[0.0888]	[0.1697]	[0.0727]	[0.1195]	[0.1672]
CM Rank		0.0029	0.0020	0.0004	0.0044	0.0105***
		[0.0027]	[0.0044]	[0.0028]	[0.0040]	[0.0035]
Rival firm characteristics						
Firm size		0.0007	0.0023	0.0014	0.0011	-0.0003
		[0.0011]	[0.0016]	[0.0017]	[0.0040]	[0.0014]
Asset turnover		-0.0039	-0.0182**	-0.0014	-0.0029	-0.0011
		[0.0025]	[0.0092]	[0.0051]	[0.0026]	[0.0019]
Debt ratio		-0.0718**	-0.0752	-0.0596**	-0.0390	-0.0999
		[0.0340]	[0.1271]	[0.0270]	[0.0304]	[0.0641]
MB ratio		0.0016	0.0018	0.0065***	0.0017*	0.0002
		[0.0011]	[0.0030]	[0.0021]	[0.0010]	[0.0018]
R&D intensity		0.0138**	0.0113	0.0157	0.0111*	0.0053
		[0.0066]	[0.0131]	[0.0147]	[0.0067]	[0.0063]
Total observations	26,777	22,246	1,659	4,985	13,117	13,260
Number treated firms	449	436	45	117	344	398
Number control firms	2,245	2,042	176	306	1,295	1,798
Control firms per treated firms (mean)	5.00	4.68	3.91	2.62	3.76	4.52
Number IPO filings	141	122	101	108	120	82

Note: Results of the DiD analysis around ITRs. ITRs are defined as negative tariff changes that are two times larger than the industry's median change. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. In specification (3), the sample is restricted to firms operating in industries with a low pre-treatment competition level (HHI>=0.15). In specification (4), the sample is restricted to firms operating in industries with a high pre-treatment competition level (HHI<0.15). Specification (5) excludes rival firms that have an own stock event in the time around the IPO filing. Specification (6) excludes IPO filings that are withdrawn later on. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table 4.3: Competition effect: Robustness tests

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.0508***	0.0569***	0.0772***			
	[0.0121]	[0.0185]	[0.0210]			
Placebo effect				-0.0235	0.0245	-0.0165
				[0.0176]	[0.0152]	[0.0157]
IPO and IPO firm characteristics						
Firm size	0.0242	0.0288	0.0493	0.1638**	0.1824**	0.1803**
	[0.0202]	[0.0242]	[0.0902]	[0.0773]	[0.0755]	[0.0762]
Asset turnover	-0.0091	0.0087	0.0715***	0.0461***	0.0469***	0.0456***
	[0.0066]	[0.0069]	[0.0205]	[0.0150]	[0.0150]	[0.0150]
Debt ratio	0.0178***	0.0197**	0.0262	-0.0296	-0.0269	-0.0285
	[0.0062]	[0.0084]	[0.0214]	[0.0203]	[0.0200]	[0.0200]
Firm age	-0.0012***	-0.0005	-0.0015	-0.0013*	-0.0015**	-0.0014*
	[0.0004]	[0.0006]	[0.0009]	[0.0007]	[0.0007]	[0.0007]
Filing size	0.0067	-0.0156	-0.1750	-0.1423	-0.1452	-0.1554
	[0.0682]	[0.0657]	[0.1875]	[0.1142]	[0.1076]	[0.1106]
CM Rank	-0.0013	0.0022	0.0092	0.0036	0.0030	0.0034
	[0.0014]	[0.0025]	[0.0095]	[0.0037]	[0.0037]	[0.0037]
Rival firm characteristics						
Firm size	0.0005	0.0001	-0.0085	0.0071	0.0069	0.0073
	[0.0007]	[0.0019]	[0.0073]	[0.0070]	[0.0071]	[0.0070]
Asset turnover	-0.0036**	-0.0032	0.0099***	-0.0037*	-0.0039**	-0.0039**
	[0.0017]	[0.0024]	[0.0038]	[0.0020]	[0.0020]	[0.0020]
Debt ratio	-0.0274	-0.0366	0.0582	0.0633	0.0525	0.0665
	[0.0239]	[0.0247]	[0.0477]	[0.0561]	[0.0571]	[0.0561]
MB ratio	0.0017*	0.0008	0.0022*	0.0041	0.0039	0.0041
	[0.0009]	[0.0010]	[0.0013]	[0.0029]	[0.0029]	[0.0029]
R&D intensity	0.0117**	0.0128**	-0.0429*	0.0118	0.0120	0.0123
	[0.0048]	[0.0063]	[0.0254]	[0.0080]	[0.0079]	[0.0080]
Total observations	22,246	7,466	32,339	7,932	7,932	7,932
Number treated firms	436	421	473	290	290	290
Number control firms	2,042	421	2,365	1,284	1,284	1,284
Control firms per treated firms (mean)	4.68	1.00	5.00	4.43	4.43	4.43
Number IPO filings	122	122	132	69	69	69

Note: Robustness tests of results of the DiD analysis around ITRs. ITRs are defined as negative tariff changes that are two times larger than the industry's median change. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. In specification (1) the dependent variable is the rival firm's CAR over the [-5;5] event window. Specification (2) uses only the nearest match as control firm. Specification (3) choses five control firms randomly out of a set of possible control firms. Specification (4) to (6) report results of placebo tests in which the treatment is shifted to different places in the pre-treatment period. In specification (4), the last observation in the pre-treatment period is pretended to be treated. In specification (5), half of the observations in the pre-treatment period are pretended to be treated. In specification (6), the first observation in the pre-treatment period is pretended to be treated. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

that there are no diverging trends before the ITRs, making the common trend assumption more credible in this context.

As additional robustness test, analyses in Table 4.2 and 4.3 are repeated by using a different definition of the ITRs, that is, an ITRs is defined as a negative tariff change that is three times larger than the industry's median change. Defining ITRs in this way is more restrictive and thus leads to a slightly change in the sample. Results are reported in Table C.4 and Table C.5 in the appendix. All specifications are estimated analogously to the specifications in Table 4.2 and Table 4.3. Overall, results are quite similar to the main results and thus they seem to be also robust to a change in the definition of the ITRs.

In sum, the estimated treatment effect is positive and statistically significant. Following the assumption that the competition effect is weaker if the competition level in the market is higher, this suggests that a competition effect is likely to exist and an IPO filing in an industry is likely to harm industry rivals to a certain extent. The estimated treatment effects are robust to different changes in the sample and for different definitions of the control group and the treatment.

4.4.2 Information effect

This section reports the results of the analysis of the information effect. Due to the sample construction described above, the sample composition varies between the competition effect and the information effect. Thus, descriptive statistics differ compared to the descriptive statistics shown in the competition effect analysis and are therefore reported separately. Table 4.4 reports event study results performed in the first step and Figure C.2 shows the development of CARs over the [-10;10] event window. Overall, all CARs are negative and statistically significant. The pattern of CARs is quite similar to the ones in the samples used for the competition effect, but CARs are slightly smaller. Again, the measurable price reaction is stronger when choosing larger event windows (e.g. CAR of -0.46% for an

Table 4.4: Information effect: Event study results

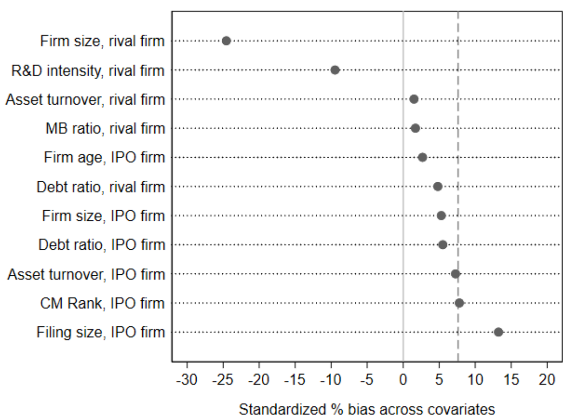
Event window	CAR (mean)	p-value	CAR (median)	p-value	Obs
[-1, 1]	-0.46	0.00	-0.50	0.00	4,319
[-2, 2]	-0.43	0.00	-0.45	0.00	4,319
[-5, 5]	-0.87	0.00	-1.00	0.00	4,319
[-10, 1]	-1.25	0.00	-1.06	0.00	4,319
[-10, 5]	-1.54	0.00	-1.43	0.00	4,319
[-10, 10]	-2.04	0.00	-1.46	0.00	4,319

Note: Event study results of the rival firms included in the sample to analyse the information effect. Cumulative abnormal returns (CARs) are reported in percent for different event windows. P-values of t-tests are reported that indicate whether the CARs are statistically different from zero.

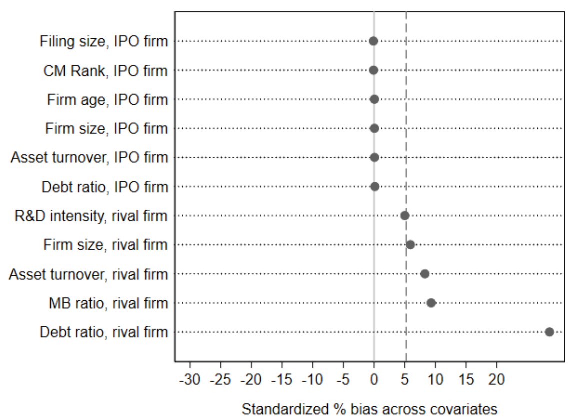
event window of [-1;1] compared to CAR of -2.04% for the [-10;10] event window). In this setting, the event window of [-10;10] is also considered to be the preferred one and used for the DiD analysis. Robustness tests are reported for the [-5;5] event window.

Analogously to Table C.3, Table C.6 reports descriptive statistics of the matching variables. Figures 4.2a to 4.2c plot the respective standardized biases of the covariates for different samples. Figure 4.2a shows the standardized bias of the sample in which control firms are matched to treatment firms via mahalanobis matching which allows a maximum number of five control firms per treatment firms. Figure 4.2b shows the standardized bias of the respective sample that uses only the nearest match as control firm. Figure 4.2c plots the standardized biases for the sample that matches five control firms randomly to the treatment firms. The dashed vertical line corresponds to the mean standardized bias in all figures, respectively. Standardized biases overall are substantially smaller than in the competition effect sample, reflecting the fact that treatment and control firms react to the same IPO in this setting. The standardized mean bias is quite similar for all samples here

(a) Mahalanobis matching (5 matches)



(b) Mahalanobis matching (1 match)



(c) Randomly chosen control sample

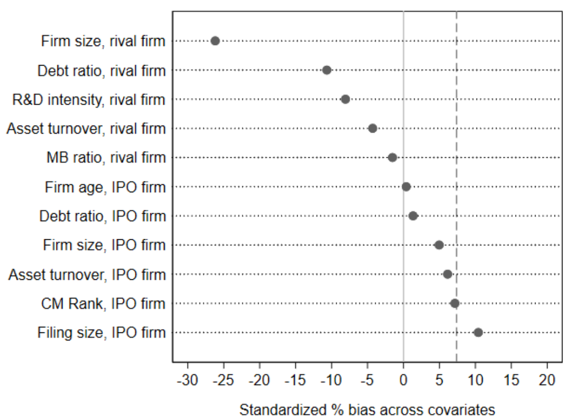


Figure 4.2: Information effect samples: standarized biases

(varying between 5.21 and 7.62), as the variables are already similar before the matching. At the same time, there are less possible matches for each treatment firm due to this condition. Therefore, many treatment firms do have less than five possible control firms. In this case, it is irrelevant whether these firms are matched randomly to the treatment firms or based on covariates, as all possible control firms end up in the sample. For this reason, the mean standardized bias does not vary much between the different information effect samples. It should also be noted that for the same reason, the number of observations differs to the number of observations in the competition effect analysis.

The main results for the information effect analysis are reported in Table 4.5. The estimated treatment effect is insignificant and very small in specification (1) without control variables (standardized treatment effect measure of 0.0459) and specification (2) with control variables (standardized treatment effect measure of -0.0171). This implies that the existence of the information effect cannot be confirmed by the analysis at the first place. Specification (3) excludes rival firms that have an own stock event around the IPO filing as these firms might react differently to an IPO. However, the estimated treatment effect remains nearly zero and insignificant, suggesting that the zero effect is not due to other events biasing the results. Specification (4) excludes IPO filings that are withdrawn later on, as these filings might bias the treatment effect downwards. The treatment effect remains still nearly zero and insignificant and thus withdrawn filings seem not to be the reason why results are insignificant. Specification (5) only includes firms in the manufacturing industries in order to make the sample better comparable to the competition effect sample and to test whether differences in the results are driven by industry specific characteristics. Again, the estimated treatment effect remains insignificant and nearly zero, suggesting that the differences in the results are not driven by special industry characteristics.

Further robustness tests are reported in Table 4.6. The first specification uses rival firm's CAR over the [-5;5] event window as dependent variable to test the robustness of the results with regard to a slight change in the dependent variable. The estimated treatment effect stays statistically insignificant to this variation suggesting that the finding is not specific to the chosen event window in the event study. The following two specifications test the robustness of the results with regard to variations in the control sample. Instead of using the five nearest matches as control firms, specification (2) uses only the nearest matched rival as control firm and specification (3) choses up to five control firms randomly out of the possible control firms without applying pre-matching. The estimated treatment effect does not change in either of both specifications.

Specifications (4) to (6) report different placebo tests in order to implicitly test the common trend assumption. Analogously to the placebo tests reported in Table 4.3, the

Table 4.5: Information effect: Difference-in-Difference results

	(1)	(2)	(3)	(4)	(5)
Treatment effect	0.0072	-0.0026	-0.0268	-0.0074	-0.0031
	[0.0144]	[0.0161]	[0.0249]	[0.0182]	[0.0230]
IPO and IPO firm characteristics					
Firm size		-0.0178*	0.0036	-0.0135	-0.0264
		[0.0093]	[0.0185]	[0.0108]	[0.0167]
Asset turnover		0.0035	-0.0009	0.0047	-0.0073
		[0.0070]	[0.0130]	[0.0080]	[0.0100]
Debt ratio		0.0003	-0.0021	-0.0069	0.0052
		[0.0064]	[0.0106]	[0.0097]	[0.0100]
Firm age		0.0007*	0.0009	0.0005	0.0010
		[0.0004]	[0.0010]	[0.0004]	[0.0006]
Filing size		0.0371	-0.0732	0.0024	-0.0105
		[0.0472]	[0.0853]	[0.0619]	[0.0724]
CM Rank		-0.0046	-0.0038	-0.0050	0.0006
		[0.0036]	[0.0057]	[0.0037]	[0.0059]
Rival firm characteristics					
Firm size		-0.0017	-0.0056**	-0.0027	-0.0035
		[0.0021]	[0.0024]	[0.0036]	[0.0101]
Asset turnover		0.0004	0.0038	0.0014	0.0013
		[0.0033]	[0.0033]	[0.0036]	[0.0032]
Debt ratio		0.0319	0.1121	0.0681	0.0022
		[0.0990]	[0.1194]	[0.0986]	[0.1672]
MB ratio		-0.0003	0.0010	-0.0002	-0.0003
		[0.0021]	[0.0025]	[0.0023]	[0.0021]
R&D intensity		0.0006	-0.0069	0.0017	-0.0005
		[0.0078]	[0.0072]	[0.0097]	[0.0093]
Total observations	4,319	3,278	1,496	2,508	1,646
Number treated firms	107	91	52	80	47
Number control firms	357	296	137	275	137
Control firms per treated firms (mean)	3.22	3.12	2.54	3.27	2.85
Number IPOs	379	304	226	224	168

Note: Results of the DiD analysis around broker closures and broker mergers. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. Specification (3) excludes rival firms that have an own stock event in the time around the IPO filing. Specification (4) excludes IPO filings that are withdrawn later on. Specification (5) only includes firms in manufacturing industries. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table 4.6: Information effect: Robustness tests

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.0089	0.0106	0.0094			
	[0.0116]	[0.0221]	[0.0158]			
Placebo effect				-0.0346	-0.0129	0.0154
				[0.0253]	[0.0233]	[0.0249]
IPO and IPO firm characteristics						
Firm size	-0.0117	-0.0242*	-0.0096	0.0109	0.0111	0.0096
	[0.0071]	[0.0136]	[0.0058]	[0.0256]	[0.0255]	[0.0254]
Asset turnover	-0.0107**	-0.0058	0.0036	0.0193	0.0192	0.0186
	[0.0050]	[0.0088]	[0.0070]	[0.0118]	[0.0118]	[0.0118]
Debt ratio	0.0024	-0.0029	-0.0047	-0.0117	-0.0114	-0.0110
	[0.0042]	[0.0080]	[0.0081]	[0.0114]	[0.0113]	[0.0114]
Firm age	0.0004	0.0010*	0.0006*	0.0014*	0.0014*	0.0014*
	[0.0003]	[0.0006]	[0.0003]	[0.0008]	[0.0008]	[0.0008]
Filing size	0.0288	0.0374	0.0111	-0.2700*	-0.2735*	-0.2690*
	[0.0369]	[0.0681]	[0.0507]	[0.1478]	[0.1472]	[0.1475]
CM Rank	-0.0034	0.0004	-0.0017	-0.0007	-0.0003	-0.0006
	[0.0023]	[0.0073]	[0.0039]	[0.0075]	[0.0075]	[0.0075]
Rival firm characteristics						
Firm size	0.0005	-0.0026	-0.0003	-0.0020	-0.0016	-0.0026
	[0.0017]	[0.0085]	[0.0005]	[0.0235]	[0.0232]	[0.0239]
Asset turnover	-0.0041**	-0.0004	-0.0003	-0.0020	-0.0017	-0.0017
	[0.0021]	[0.0038]	[0.0014]	[0.0092]	[0.0093]	[0.0094]
Debt ratio	0.0011	-0.0456	0.0371	-0.2387	-0.2414	-0.2536
	[0.0748]	[0.1214]	[0.0607]	[0.1508]	[0.1497]	[0.1543]
MB ratio	0.0005	0.0005	-0.0017	-0.0014	-0.0012	-0.0009
	[0.0014]	[0.0031]	[0.0024]	[0.0062]	[0.0063]	[0.0063]
R&D intensity	0.0114*	0.0010	-0.0013	-0.0105	-0.0116	-0.0119
	[0.0066]	[0.0132]	[0.0038]	[0.0238]	[0.0242]	[0.0240]
Total observations	3,278	1,278	3,967	1,090	1,090	1,090
Number treated firms	91	73	115	61	61	61
Number control firms	296	75	401	178	178	178
Control firms per treated firms (mean)	3.12	1.00	3.34	2.83	2.83	2.83
Number IPOs	304	270	324	141	141	141

Note: Robustness tests of results of the DiD analysis around broker closures and broker mergers. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. In specification (1) the dependent variable is the rival firm's CAR over the [-5;5] event window. Specification (2) uses only the nearest match as control firm. Specification (3) choses five control firms randomly out of a set of possible control firms. Specification (4) to (6) report results of placebo tests in which the treatment is shifted to different places in the pre-treatment period. In specification (4), the last observation in the pre-treatment period is pretended to be treated. In specification (5), half of the observations in the pre-treatment period are pretended to be treated. In specification (6), the first observation in the pre-treatment period is pretended to be treated. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

treatment is shifted to different places in the pre-treatment period. In specification (4), the last observation in the pre-treatment period is pretended to be treated and all others are pretended to be untreated. In specification (5), the first half of the observations are pretended to be treated and the other half of the observations are pretended to be untreated. Finally, in specification (6), the first observation is pretended to be untreated and all other observations are pretended to be treated. All estimated placebo tests are not statistically significant suggesting that there are no diverging trends before the treatment which makes it very likely that the common trend assumption does hold. The estimated treatment effects are therefore likely to be valid.

To sum up, regardless of the specification, there is no significant treatment effect and thus no direct evidence of the existence of an information effect can be found. However, this does not certify that there is no information effect. From a theoretical point of view, the information effect can either be positive or negative. Thus, it would be possible that the estimated zero mean effect is due to the fact that the information effect is positive for some IPO filings while it is negative for other IPO filings and that both effects cancel each other out. In addition, it would also be possible that the crucial assumption, namely that the information effect varies systematically with the information level in the industry does not hold. Unfortunately, this assumption is hard to test in this set-up.

4.5 Conclusion

This paper contributes to the IPO literature by shedding new light on intra-industry effects of IPO filings on rival firms. To this end, it applies a new methodology that allows to test for the existence of the competition and information effect. In the first step, an event study is performed in order to estimate short-term valuation effects of IPO filings. In the second step, it tests whether these valuation effects change in a DiD setting by exploiting exogenous events that either influence the competition or the information level in the industry. If rival firms react differently to IPOs if the information or competition level in the industry changes, this can be considered as evidence for the existence of the respective effects.

Results provide sufficient evidence for the existence of the competition effect. This implies that IPO filings tend to harm rivals operating in the same industry due to a deterioration of the competitive situation for rival firms. In contrast, no direct evidence could be found for the existence of the information effect. However, this does not necessarily imply that there is no information effect. One reason for the results could also be that the crucial assumption in the identification strategy - that the information effect varies systematically with the start level of information in the industry - does not hold. Another reason could be that a positive and a negative information effect cancel each other out.As the mechanisms behind the information effect are more complex and assumptions are harder to test, the employed research design provides a more thorough test for the existence of the competition rather than for the information effect.

Beyond that, there are further limitations to the analysis. First of all, the sample is restricted in time and space. Relatedly, the firms need to experience at least one IPO in the treatment and control period, which excludes firms in industries with a fewer number of IPOs. In addition, due to the differences in sample construction between the competition and information effect, the comparison of the results of both analyses is limited and should be drawn with caution. It is possible that results might differ in other contexts and can only be generalized to a certain extent. It remains subject to future research to assess whether the results can be replicated in other contexts.

In addition, the treatments considered need to be interpreted with some caution. The calculation of cuts of ITRs is based on import data, which does not allow to check whether the used cuts are unilateral or bilateral. The effect of a bilateral cut might have a less strong effect on the competition level, as it might not only increase the competitive pressure for domestic firms, but would also expand the market on which they can sell their products. However, this implies that the actual increase in the competition level is lower, suggesting that estimated effects rather present lower bounds. The true effects are therefore (if at all) likely to be even stronger. Relatedly, the use of different treatments can further limit the scope for comparison of the competition and information effect. For example, it is possible that the ITRs might constitute a stronger treatment compared to the broker closures or mergers. Consequently, the clearer evidence for the existence of a competition effect might be due to the intensity of the related treatment rather than the importance of the effect itself.

Another point is that intra-industry analyses depend heavily on the definition of rival firms. However, it is far from straightforward to identify firms that actually compete in the same market. There are many aspects that have to be considered, for example whether firms actually produce competing products that are substitutes or whether they produce strategic complements which coincides with other competitive interdependencies (see Cooper, 1989; Lee et al., 2011, among others). In addition, sometimes markets are geographically constrained (see Shelegia, 2012, among others). Although highly established and often used in literature, SIC codes might not be sufficient to identify rivals as they do not explicitly account for all these aspects. In fact, SIC codes classify firms into industries based on a firm's primary business line and do not focus on whether these firms produce competing products. In addition, they do not account for innovations or changes over time (Hoberg and Phillips, 2016). One alternative would be to use the industry classification introduced by Hoberg and Phillips (2016), who classify firms according to firm product descriptions using text-based analysis. However, the time varying characteristic of their industry classification brings about new problems for the analysis at hand, as the DiD analysis requires constant classifications at least for the before-and-after period. Thus, there is still no perfect measure to identify rival firms and further research is required to check whether results change when changing the definition of rival firms.

Relatedly, due to the event study approach, the scope of this paper is limited to rival firms that are public. Further research is required to analyze the effects of IPOs on private firms, as these firms might be affected differently by IPOs.

All in all, the results deepen the understanding for the causal mechanisms behind the valuation effects of IPOs on rivals. At the same time, there are some limits to the generalizability of the results, and conclusions on the existence of the information effect need to be drawn with caution. Overall, a lot of work to unpack the black box of the underlying mechanisms behind the effect of IPOs on rival firms remains to be done.

C Appendix

C.1 Tables

Table C.1: Variable descriptions and data sources

Variable	Description	Source
Dependent Variable		
CARs	Cumulative abnormal returns winsorized at the 1% level at each tail	Estimated in an event study analysis (1st step)
Matching variables (winsorized at 1% level at each tail)		
IPO and IPO firm characteristics		
Firm size	Total assets in bn USD in most recent financial period	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.
Asset turnover	Ratio of total revenues to total assets in most recent financial period	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.
Debt ratio	Ratio of total liabilities to total assets in most recent financial period	Data primarily taken from Compustat. Supplemented with and checked against data stemming from S&P Capital IQ and EDGAR filings.

Firm age	Firm age of IPO firm since founding date	Founding and outcome date primarily taken from Thomson Reuters Securities Data Company (SDC) Platinum. Supplemented with and checked against data stemming from Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`) and S&P Capital IQ.
Filing size	Filing size in bn USD. Calculated as the product of original global shares filed and original mid filing price.	Thomson Reuters Securities Data Company (SDC) Platinum
CM Rank	Carter - Manaster Rank of lead underwriter in filing year. It ranges from 0 to 9.	Jay Ritter's website (`https://site.warrington.ufl.edu/ritter/ipo-data/`).Originally developed by Carter and Manaster (1990), updated by Carter, Dark, and Singh (1998) and by Loughran and Ritter (2004).
Rival firm characteristics		
Firm size	Total assets in bn USD in most recent financial period	Compustat
Asset turnover	Ratio of total revenues to total assets in most recent financial period	Compustat
Debt ratio	Ratio of total liabilities to total assets in most recent financial period	Compustat

MB ratio	Ratio of market value to book value in most recent financial period	Compustat
R&D intensity	Ratio of R&D expenditure to sales in most recent financial period	Compustat
Treatments		
Import tariffs reductions (ITRs)	The ad-valorem tariff rate is calculated as the duties collected by US Customs divided by the Free-on-Board value of imports for each industry year. Then, the change in tariff cuts per year is calculated. A tariff cut is defined as a negative change in import tariffs that is two or three times larger than the industry's median change.	Peter Schott's Website (`http://faculty.som.yale.edu/peterschott/sub/international.htm`)
Information on analyst coverage		I/B/E/S
Information on broker closures & broker mergers		S&P Capital IQ, Financial Industry Regulatory Authority (FINRA) (`https://www.finra.org`)
Others		
SIC codes		Data primarily taken from Compustat. Supplemented with and checked against data stemming from Thomson Reuters Securities Data Company (SDC) Platinum and S&P Capital IQ

SIC description		`https://siccode.com/sic-code-lookup-directory`
Market data (stock prices)		CRSP
Info about other stock events	Other stock events that are considered are: M&As, liquidations, reorganizations, offer/issuances	Thomson Reuters Securities Data Company (SDC) Platinum, CRSP
Herfindahl Hirschman Index (HHI)	Measure for the level of market concentration. It is defined as the sum of the squares of the market shares s_i (i = 1, 2, . . . , n) of all entities in the industry: $H = \sum_{i=1}^{n}(s_i)^2$ (Brezina et al., 2016, p.53). It ranges from $1/n$ (lowest market concentration) to 1 (highest market concentration) (Brezina et al., 2016, p.53). In the case at hand, an industry is defined to be unconcentrated if HHI <0.15 and concentrated, if HHI >=0.15. (Due to data limitations, the categories "moderately concentrated" and "highly concentrated" reported in Brezina et al. (2016) are summarized to "concentrated" for the purpose of this study).	Own calculations based on market share data stemming from Compustat. Supplemented with data stemming from Refinitiv Eikon. Based on 3-digit SIC codes.

Table C.2: Overview over import tariff reductions

SIC code	Industry description	Year	Change in industry per year in percent (relative)	Industry median
Panel A: An ITRs is defined as a negative change in import tariffs that is two times larger than the industry's median change				
Chemicals and Allied Products				
283	Drugs	2004	-0.0577	-0.0109
286	Industrial Organic Chemicals	2007	-0.1752	-0.0356
289	Miscellaneous Chemical Products	2000	-0.0384	-0.0113
Rubber and Miscellaneous Plastics Products				
308	Miscellaneous Plastic Products	2002	-0.0118	-0.0046
Industrial and Commercial Machinery and Computer Equipment				
351	Engines and Turbines	2003	-0.0568	-0.0078
353	Construction, Mining, and Material Handling	2000	-0.2052	-0.0153
356	General Industrial Machinery and Equipment	2006	-0.0218	-0.0059
357	Computer and Office equipment	2006	-0.0469	-0.0052
Electronic and other Electrical Equipment and Components, except Computer Equipment				
366	Communications Equipment	2009	-0.0479	-0.0033
367	Electronic Components and Accessories	2003	-0.0582	-0.0138
369	Miscellaneous Electrical Machinery, Equipment, and Supplies	2001	-0.0540	-0.0051
Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks				
382	Laboratory Apparatus and Analytical, Optical, Measuring, and Controlling Instruments	2001	-0.0509	-0.0066

Panel B: An ITRs is defined as a negative change in import tariffs that is three times larger than the industry's median change

Chemicals and Allied Products				
283	Chemicals and Allied Products	2004	-0.0577	-0.0109
286	Industrial Organic Chemicals	2007	-0.1752	-0.0356
287	Agricultural Chemicals	2007	-0.1073	-0.0144
289	Miscellaneous Chemical Products	2000	-0.0384	-0.0113
Industrial and Commercial Machinery and Computer Equipment				
351	Engines and Turbines	2003	-0.0568	-0.0078
353	Construction, Mining, and Material Handling	2000	-0.2052	-0.0153
356	General Industrial Machinery and Equipment	2006	-0.0218	-0.0059
357	Computer and Office equipment	2006	-0.0469	-0.0052
Electronic and other Electrical Equipment and Components, except Computer Equipment				
366	Communications Equipment	2009	-0.0479	-0.0033
367	Electronic Components and Accessories	2003	-0.0582	-0.0138
369	Miscellaneous Electrical Machinery, Equipment, and Supplies	2001	-0.0540	-0.0051
Measuring, Analyzing, and Controlling Instruments; Photographic, Medical and Optical Goods; Watches and Clocks				
382	Laboratory Apparatus and Analytical, Optical, Measuring, and Controlling Instruments	2001	-0.0509	-0.0066

Note: Overview over import tariff reductions (ITRs). In Panel A, an ITRs is defined as a negative change in import tariffs that is two times larger than the industry's median change. In Panel B, an ITRs is defined as a negative change in import tariffs that is three times larger than the industry's median change.

Table C.3: Competition effect: Descriptives covariates

	Treated firms			**Control firms**		
Variable	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
IPO and IPO firm characteristics						
Firm size	0.1059	0.1713	4,655	0.0830	0.1547	21,755
Asset turnover	0.8490	0.8042	4,602	0.5545	0.5414	20,688
Debt ratio	0.6702	0.6449	4,655	0.5221	0.5610	21,712
Firm age	9.9795	9.6256	4,438	11.0395	7.7409	21,755
Filing size	0.1007	0.0651	4,529	0.1022	0.1002	20,839
CM Rank	8.0264	1.4631	4,655	7.8637	1.4255	21,755
Rival firm characteristics						
Firm size	2.2273	5.8598	4,714	1.4358	3.5517	21,355
Asset turnover	5.4575	10.5593	4,656	1.6025	3.0871	21,348
Debt ratio	0.1530	0.1968	4,706	0.1214	0.1447	21,355
MB ratio	4.1093	4.2278	4,714	3.7262	3.3022	21,343
R&D intensity	1.3924	3.2472	4,646	0.1644	0.7070	21,293

Note: Descriptive statistics of matching variables for rival firms included in the sample to analyse the competition effect by treatment status. Variable descriptions and data sources are summarized in Table C.1.

Table C.4: Competition effect: DiD robustness - different definition of ITRs I

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.0751***	0.0796***	0.0880**	-0.0070	0.0627***	0.0970***
	[0.0198]	[0.0184]	[0.0385]	[0.0188]	[0.0233]	[0.0249]
IPO and IPO firm characteristics						
Firm size		0.0238	0.0100	0.0790**	0.0317	0.2229***
		[0.0260]	[0.0570]	[0.0347]	[0.0319]	[0.0721]
Asset turnover		0.0153	0.0430***	-0.0137	0.0282**	-0.0044
		[0.0106]	[0.0158]	[0.0127]	[0.0111]	[0.0099]
Debt ratio		0.0160*	-0.0065	0.0272**	0.0113	0.0233
		[0.0096]	[0.0212]	[0.0128]	[0.0117]	[0.0142]
Firm age		-0.0008	-0.0005	-0.0005	-0.0007	-0.0014*
		[0.0007]	[0.0010]	[0.0005]	[0.0008]	[0.0008]
Filing size		-0.0356	0.0643	-0.1977*	-0.0816	-0.4026**
		[0.0908]	[0.1707]	[0.1095]	[0.1213]	[0.1695]
CM Rank		0.0018	-0.0022	0.0068	0.0030	0.0095**
		[0.0033]	[0.0051]	[0.0054]	[0.0045]	[0.0042]
Rival firm characteristics						
Firm size		-0.0007	0.0071	-0.0010	0.0016	-0.0028
		[0.0025]	[0.0250]	[0.0035]	[0.0044]	[0.0031]
Asset turnover		-0.0035	-0.0266***	0.0007	-0.0024	-0.0008
		[0.0024]	[0.0089]	[0.0054]	[0.0025]	[0.0017]
Debt ratio		-0.0539	-0.0910	-0.0740**	-0.0118	-0.0703
		[0.0353]	[0.1499]	[0.0344]	[0.0313]	[0.0675]
MB ratio		0.0013	0.0009	0.0020	0.0015	0.0004
		[0.0011]	[0.0032]	[0.0027]	[0.0010]	[0.0019]
R&D intensity		0.0104**	0.0144	0.0041	0.0077	0.0035
		[0.0052]	[0.0197]	[0.0148]	[0.0054]	[0.0043]
Total observations	23,135	19,095	1,415	2,948	11,688	11,257
Number treated firms	452	439	44	121	345	359
Number control firms	2,260	2,022	166	334	1,252	1,617
Control firms per treated firms (mean)	5.00	4.61	3.77	2.76	3.63	4.50
Number IPOs	162	133	102	115	131	90

Note: Results of the DiD analysis around ITRs. ITRs are defined as negative tariff changes that are three times larger than the industry's median change. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. In specification (3), the sample is restricted to firms operating in industries with a low pre-treatment competition level (HHI>=0.15). In specification (4), the sample is restricted to firms operating in industries with a high pre-treatment competition level (HHI<0.15). Specification (5) excludes rival firms that have an own stock event in the time around the IPO. Specification (6) excludes IPO filings that are withdrawn later on. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table C.5: Competition effect: DiD robustness - different definition of ITRs II

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	0.0475***	0.0504**	0.0802***			
	[0.0122]	[0.0212]	[0.0221]			
Placebo effect				-0.0253	0.0233	-0.0196
				[0.0186]	[0.0175]	[0.0173]
IPO and IPO firm characteristics						
Firm size	0.0219	0.0256	0.0558	0.1040	0.1261**	0.1182*
	[0.0162]	[0.0228]	[0.0918]	[0.0631]	[0.0589]	[0.0611]
Asset turnover	-0.0072	0.0109	0.0681***	0.0379*	0.0401*	0.0376*
	[0.0071]	[0.0090]	[0.0202]	[0.0218]	[0.0223]	[0.0220]
Debt ratio	0.0219***	0.0156	0.0290	-0.0092	-0.0061	-0.0076
	[0.0068]	[0.0098]	[0.0208]	[0.0211]	[0.0207]	[0.0208]
Firm age	-0.0011***	-0.0008	-0.0015	-0.0008	-0.0011	-0.0009
	[0.0004]	[0.0007]	[0.0010]	[0.0008]	[0.0008]	[0.0008]
Filing size	-0.0075	0.0077	-0.1802	-0.1701	-0.1533	-0.1760
	[0.0664]	[0.0712]	[0.1808]	[0.1257]	[0.1269]	[0.1241]
CM Rank	-0.0006	0.0016	0.0085	0.0028	0.0019	0.0025
	[0.0017]	[0.0029]	[0.0095]	[0.0045]	[0.0044]	[0.0045]
Rival firm characteristics						
Firm size	0.0002	-0.0012	-0.0087	0.0015	0.0009	0.0019
	[0.0014]	[0.0030]	[0.0065]	[0.0065]	[0.0065]	[0.0065]
Asset turnover	-0.0031*	-0.0027	0.0105***	-0.0035*	-0.0035*	-0.0036*
	[0.0017]	[0.0024]	[0.0038]	[0.0020]	[0.0020]	[0.0020]
Debt ratio	-0.0119	-0.0296	0.0511	0.0737	0.0660	0.0780
	[0.0287]	[0.0270]	[0.0467]	[0.0631]	[0.0648]	[0.0630]
MB ratio	0.0014	0.0007	0.0023*	0.0027	0.0025	0.0027
	[0.0010]	[0.0011]	[0.0013]	[0.0029]	[0.0028]	[0.0029]
R&D intensity	0.0082**	0.0088	-0.0595*	0.0086	0.0086	0.0091
	[0.0037]	[0.0053]	[0.0344]	[0.0064]	[0.0064]	[0.0063]
Total observations	19,095	6,607	32,043	6,146	6,146	6,146
Number treated firms	439	411	461	286	286	286
Number control firms	2,022	411	2,305	1,249	1,249	1,249
Control firms per treated firms (mean)	4.61	1.00	5.00	4.37	4.37	4.37
Number IPOs	133	128	124	72	72	72

Note: Robustness tests of results of the DiD analysis around ITRs. ITRs are defined as negative tariff changes that are two times larger than the industry's median change. The dependent variable is the rival firm's CAR over the [-10; 10] event window. Variable descriptions and data sources are summarized in Table C.1. Control firms are matched to treatment firms in a pre-stage via mahalanobis matching. In specification (1) the dependent variable is the rival firm's CAR over the [-5;5] event window. Specification (2) uses only the nearest match as control firm. Specification (3) choses five control firms randomly out of a set of possible control firms. Specification (4) to (6) report results of placebo tests in which the treatment is shifted to different places in the pre-treatment period. In specification (4), the last observation in the pre-treatment period is pretended to be treated. In specification (5), half of the observations in the pre-treatment period are pretended to be treated. In specification (6), the first observation in the pre-treatment period is pretended to be treated. Firm-cut fixed effects and year fixed effects are included. Standard errors clustered at firm level are in parentheses. Statistics with significance at the 10%, 5%, and 1% level are denoted with *, **, and ***.

Table C.6: Information effect: Descriptives covariates

	Treated firms			**Control firms**		
Variable	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
IPO and IPO firm characteristics						
Firm size	0.1847	0.4811	1,010	0.1599	0.4544	3,097
Asset turnover	0.9602	0.8117	995	0.9014	0.8068	3,033
Debt ratio	0.8202	0.7850	1,001	0.7784	0.7330	3,072
Firm age	12.3022	12.5838	973	11.9622	12.6944	2,991
Filing size	0.0976	0.1036	987	0.0848	0.0893	3,052
CM Rank	7.6708	2.0278	1,010	7.5109	2.0727	3,107
Rival firm characteristics						
Firm size	1.4463	3.9416	1,024	3.2486	9.6005	3,233
Asset turnover	2.2987	3.8601	1,022	2.2433	3.5253	3,222
Debt ratio	0.1257	0.1518	1,019	0.1185	0.1473	3,204
MB ratio	4.5380	4.5939	1,024	4.4661	3.8193	3,233
R&D intensity	0.4353	1.3656	1,001	0.5759	1.5955	3,218

Note: Descriptive statistics of matching variables for rival firms included in the sample to analyse the information effect by treatment status. Variable descriptions and data sources are summarized in Table C.1.

C.2 Figures

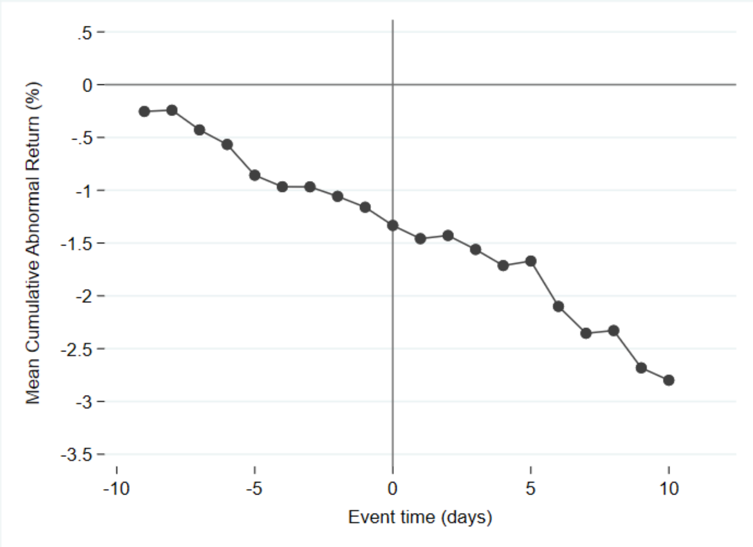


Figure C.1: Competition effect: Development of CARs

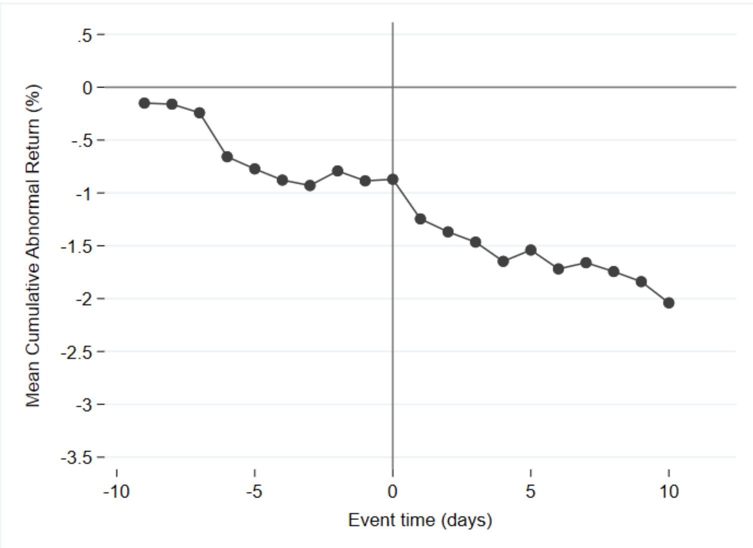


Figure C.2: Information effect: Development of CARs

Chapter 5

Conclusion

The overarching goal of this thesis is to contribute to the IPO literature by analyzing unresolved questions with regard to the phenomenon of IPO withdrawal and the effect of IPOs on industry rivals. With regard to IPO withdrawal, it particularly addresses the role of quality signals and monitoring mechanisms, examines the importance of different variables and adds a more forward-looking perspective to explain and predict IPO withdrawal. With regard to intra-industry effects of IPOs, it adds to previous discussions on the underlying mechanisms of the possibly occurring effects more closely. These questions are addressed in three separate but theoretically linked empirical studies.

The first study contributes to the IPO withdrawal literature mainly in a methodological way. As one of the first, it applies two machine learning methods, namely lasso and random forest, to predict IPO withdrawal. The performance of the machine learning methods is compared to a logistic regression model, which has been commonly applied in this research field and therefore provides a good standard for comparison. Results show that random forest performs quite well in predicting IPO withdrawal and clearly outperforms logit when considering in-sample and cross-sectional out-of-sample prediction performance. In contrast, lasso and logit show a comparable and if at all moderate performance with regard to in-sample and cross-sectional out-of-sample prediction. When trying to predict future IPO withdrawal outcomes based on historical data, the predictive power of all models falls below a purely random prediction. One reason for this quite puzzling finding is the presence of concept drift, which, at the case at hand, means that the relationship between the predictors and the outcome variable changes over time. All methods identify market characteristics at filing to be the most important variables for prediction, while corporate governance and intermediary characteristics seem to be less important.

The second study moves away from a data-driven approach towards a theory-based approach to shed new light on IPO withdrawals. As the first study identifies market characteristics to be important determinants for IPO withdrawal, the focus of the second study is to test the hypothesis whether certain variables can serve as quality signal for investors in volatile markets thereby reducing the withdrawal probability even under bad market conditions. The idea is that investors tend to be especially careful in these situations and thus quality signals or monitoring are expected to be especially meaningful. Results suggest that corporate governance characteristics, like large and experienced boards, serve as signals in volatile markets, thereby decreasing the withdrawal probability under these circumstances. Further, the study contributes to the literature by deepening our understanding of the role of VC backing by considering different VC characteristics rather than an average effect of VC backing. US VCs tend to reduce the withdrawal probability. The same holds true for syndicated VCs, which tend to be associated with a lower withdrawal probability than stand-alone VCs, especially in highly volatile markets. Both results are in line with the theory that local VCs and syndicated VCs may serve as signals and decrease investors uncertainties. In contrast, reputable VCs are associated with lower withdrawal rates only in less volatile markets, which is not in line with the signaling explanation. One explanation for this finding could be that firms backed by reputable VCs rather postpone their IPO or follow the dual-track strategy in volatile markets.

Overall, the first two studies contribute to the field of IPO withdrawal in two different ways. While the first study aims at improving the statistical methods applied, the second study assesses the application of the signaling theory in the IPO context under different conditions. Both studies together demonstrate that data-driven and theory-based approaches are two sides of the same coin and both sides should be considered to understand the whole picture of IPO withdrawal. Machine learning methods - especially random forest - can provide a promising toolbox to improve the prediction performance in the context of IPO withdrawals. This adds to previous research in empirical finance which confirms that machine learning methods outperform regression-based approaches when prediction is the main purpose (see e.g. Barboza et al. (2017) for bankruptcy prediction or Reber et al. (2005) for the prediction of IPO underpricing). In this regard, further developing and applying machine learning methods can be a promising approach in empirical IPO research. In addition, the first study deepens our understanding of which variables are most important for prediction. Although logit is able to identify the most important variables, too, random forest and lasso provide a more explicit ranking of the variables and their importance for predicting IPO withdrawal. While this seems to be a somewhat incremental improvement at the first glance, it should be considered that identifying the most important variables may be more complex when the importance is less clear cut (e.g. in case of a higher number of highly predictive variables) and may therefore be helpful for future research.

However, machine learning methods can suffer from the same problems as regression-based methods when it comes to the prediction of future IPO withdrawals based on historical data. Understanding the mechanisms that lead to this poor prediction is a central question for future research. This includes the question of why the influence of certain variables seems to change over time. In this regard, it would be essential to consider theoretical advancements in addition to methodological ones. The importance of considering theoretical approaches is also demonstrated by the second study in this thesis, which shows that signaling theory can help to explain certain withdrawal patterns. Taking further theories into account to explain IPO withdrawals should therefore also be part of future research. For example, Helbing (2019) suggests to account for behavioral and non-rational approaches when analyzing the decision to withdraw an IPO. In addition, theory can possibly add further suggestions about new influential predictors.

While the first two studies contribute to the IPO withdrawal literature, the third study contributes to another rather less explored strand within the IPO literature, namely the analysis of intra-industry effects of IPOs. In particular, a new two-step methodology delivers new insights about the mechanisms behind the effects of IPOs on industry rivals by explicitly testing for the existence of the competition and information effect. In this regard, the crucial idea of the methodology is to compare rival firm's reactions to exogeneous events that either only influence the competition or information level in the industry for a subgroup of firms. Conditional on the assumption that the influence of an additional IPO in an industry gets weaker the higher the starting level of competition and information in the industry are, this analysis can provide insights about the existence of both effects. Results suggest that there is a competition effect that seems to harm industry rivals to a certain extent. One reason for this finding could be that firms gain some kind of competitive advantage over their industry peers by going public, which increases the competitive pressure for firms in the industry. In contrast, the analysis does not provide direct evidence for the existence of the information effect. However, the lack of evidence for an aggregate information effect does not necessarily imply that the information effect does not exist. Unlike to the competition effect, the direction of the information effect is less clear from a theoretical point of view. On the one hand, IPOs could signal good growth prospects for the whole industry which would be in line with positive valuation effects on rivals. On the other hand, IPOs could foreshadow future negative industry trends or reveal that the industry is overvalued, which would be in line with negative valuation effects. The insignificant estimated treatment effect could therefore also result from the fact that a positive and a negative information effect cancel each other out.

This study demonstrates how methodological advancements can help to learn more about existing theories. However, the study also reveals a lot of un-resolved puzzles in the analysis of the effects and underlying mechanisms of IPOs on industry rivals. The results show that especially disentangling the information effect deserves closer consideration. For this purpose, methodologies are needed that go beyond the scope of event study approaches, which only allow to consider average valuation effects. Although extending the scope of event studies, the methodology applied in this study is still not sufficient to go beyond average effects and demonstrate the need for more fine-grained analyses. The recently published study by Spiegel and Tookes (2020) provides a start for the application of such methods by performing an event-by-event analysis and demonstrates that rival firm's reactions seem to be IPO specific. However, studies are not only limited with regard to the applied methodologies but also with regard to the definition of rival firms. Most studies rely on SIC codes to identify rival firms. However, in order to get a more detailed picture of the effects of IPOs on rival firms, different definitions of rival firms should be considered, including an extension to private firms. Thus, there is a lot of scope for future research to deepen our understanding of the effects of IPOs on rivals, which is equally important for rivals and investors.

From a more general perspective, there are at least two directions for future research that should be considered. First, this thesis as well as the majority of IPO research focusses on the US market. However, properties of IPO markets in different countries and jurisdictions might differ and thus it is also likely that characteristics of IPO withdrawal or intra-industry effects might differ across countries (for IPO withdrawals see Helbing, 2019). As IPO withdrawals seem to be a global phenomenon (see Helbing, 2019), and there is also no reason why intra-industry effects of IPOs would be limited to the US market, it is rather surprising that research on other countries is very scarce. As findings from US data cannot be generalized to other countries, future research should extend the analysis to different countries and also replicate studies based on US data using data from different other countries. A first start of extending the IPO withdrawal research in space is provided by Helbing et al. (2019) for Europe and by Fan and Yamada (2020) for Asia, while McGilvery et al. (2012) analyze intra-industry effects of Australian IPOs. Second, another under-explored issue in the IPO literature is the simultaneous consideration of both sub-research-areas, that is, analyzing the effect of IPO withdrawals on industry rivals more closely. The findings by Hsu et al. (2010) suggest that the effect of IPO withdrawals on rival firms is opposed to the effect of IPO filings and successful IPOs. Exploring these effects in more detail would also be a task for future research.

In conclusion, this thesis contributes to the IPO literature at multiple levels. At the methodological level, this includes new research designs to disentangle causal mechanisms underlying intra-industry effects of IPOs as well as machine learning methods that focus on improving the prediction performance in the context of IPO withdrawals. At a theoretical level, it goes beyond previous approaches by highlighting the context dependence of different factors which influence IPO withdrawal. The results suggest that these new approaches are promising to further develop IPO research and highlight that theoretical and methodological issues need to be considered simultaneously. At the same time, this thesis points to new, unanswered research questions and reveals that there remains a long way to go to fully understand all decisions and mechanisms connected to IPOs. Considering the high IPO volume in different regions of the world (see Helbing, 2019) and the great importance of IPOs for the economy, addressing these unresolved issues is an important task for future research.

Bibliography

Aggarwal, R., L. Krigman and K. L. Womack (2002): "Strategic IPO underpricing, information momentum, and lockup expiration selling", *Journal of Financial Economics* 66(1), 105–137.

Aggarwal, R. and P. Rivoli (1990): "Fads in the initial public offering market?", *Financial Management*, 45–57.

Ai, C. and E. C. Norton (2003): "Interaction terms in logit and probit models", *Economics Letters* 80(1), 123–129.

Akerlof, G. A. (1970): "The market for "lemons": Quality uncertainty and the market mechanism", *Quarterly Journal of Economics* 84, 488–500.

Akhigbe, A., S. F. Borde and A. M. Whyte (2003): "Does an industry effect exist for initial public offerings?", *Financial Review* 38(4), 531–551.

Aktas, N., C. Andres and A. Ozdakak (2018): "The Interplay of IPO and M&A Markets", *The Oxford Handbook of IPOs*, Oxford University Press, 201.

Aktas, N. and M. Dupire-Declerck (2015): "Increased entry threat and merger activity", *Finance* 36(1), 75–115.

Allen, F. and G. R. Faulhaber (1989): "Signaling by underpricing in the IPO market", *Journal of Financial Economics* 23(2), 303–323.

Almeida, P. R. L., L. S Oliveira, A.S. Britto Jr and R. Sabourin (2018): "Adapting dynamic classifier selection for concept drift", *Expert Systems with Applications* 104, 67–85.

Altı, A. (2005): "IPO market timing", *The Review of Financial Studies* 18(3), 1105–1138.

Altı, A. and J. Sulaeman (2012): "When do high stock returns trigger equity issues?", *Journal of Financial Economics* 103(1), 61–87.

Archer, K. J. and R. V. Kimes (2008): “Empirical characterization of random forest variable importance measures”, *Computational Statistics & Data Analysis* 52(4), 2249–2260.

Asquith, P., M. B. Mikhail and A. S. Au (2005): “Information content of equity analyst reports”, *Journal of Financial Economics* 75(2), 245–282.

Aue, A. and L. Horváth (2013): “Structural breaks in time series”, *Journal of Time Series Analysis* 34(1), 1–16.

Bancel, F. and U. R. Mittoo (2009): “Why do European firms go public?”, *European Financial Management* 15(4), 844–884.

Barboza, F., H. Kimura and E. Altman (2017): “Machine learning models and bankruptcy prediction”, *Expert Systems with Applications* 83, 405–417.

Barua, S., M. M. Islam, X. Yao and K. Murase (2012): “MWMOTE – majority weighted minority oversampling technique for imbalanced data set learning”, *IEEE Transactions on Knowledge and Data Engineering* 26(2), 405–425.

Bastı, E., C. Kuzey and D. Delen (2015): “Analyzing initial public offerings’ short-term performance using decision trees and SVMs”, *Decision Support Systems* 73, 15–27.

Batista, G. E., R. C. Prati and M. C. Monard (2004): “A study of the behavior of several methods for balancing machine learning training data”, *ACM SIGKDD Explorations Newsletter* 6(1), 20–29.

Bekkar, M., H. K. Djemaa and T. A. Alitouche (2013): “Evaluation measures for models assessment over imbalanced data sets”, *Journal of Information Engineering and Application* 3(10).

Benninga, S., M. Helmantel and O. Sarig (2005): “The timing of initial public offerings”, *Journal of Financial Economics* 75(1), 115–132.

Benveniste, L. M., W. Y. Busaba and W. J. Wilhelm Jr (2002): “Information externalities and the role of underwriters in primary equity markets”, *Journal of Financial Intermediation* 11(1), 61–86.

Benveniste, L. M. and P. A. Spindt (1989): “How investment bankers determine the offer price and allocation of new issues”, *Journal of Financial Economics* 24(2), 343–361.

Bergbrant, M. C., D. Bradley and D. M. Hunter (2017): “Does bank loan supply affect the supply of equity capital? Evidence from new share issuance and withdrawal”, *Journal of Financial Intermediation* 29, 32–45.

Bernard, A. B., J. B. Jensen and P. K. Schott (2006): "Trade costs, firms and productivity", *Journal of Monetary Economics* 53(5), 917–937.

Bodnaruk, A., E. Kandel, M. Massa and A. Simonov (2007): "Shareholder diversification and the decision to go public", *The Review of Financial Studies* 21(6), 2779–2824.

Boeh, K. and C. Dunbar (2013): "Post IPO Withdrawal Outcomes", Available at: `https://ssrn.com/abstract=2135772`.

— (2014): "IPO waves and the issuance process", *Journal of Corporate Finance* 25, 455–473.

— (2016): "Raising Capital after IPO Withdrawal", Available at: `https://ssrn.com/abstract=2863266`.

Boeh, K. and C. Southam (2011): "Impact of initial public offering coalition on deal completion", *Venture Capital* 13(4), 313–336.

Bouis, R. (2009): "The short-term timing of initial public offerings", *Journal of Corporate Finance* 15(5), 587–601.

Bradley, A. P. (1997): "The use of the area under the ROC curve in the evaluation of machine learning algorithms", *Pattern Recognition* 30(7), 1145–1159.

Bradley, D. J. and B. D. Jordan (2002): "Partial adjustment to public information and IPO underpricing", *Journal of Financial and Quantitative Analysis* 37(4), 595–616.

Brander, J. A., R. Amit and W. Antweiler (2002): "Venture-capital syndication: Improved venture selection vs. the value-added hypothesis", *Journal of Economics & Management Strategy* 11(3), 423–452.

Brau, J. C. and S. E. Fawcett (2006): "Initial public offerings: An analysis of theory and practice", *The Journal of Finance* 61(1), 399–436.

Brau, J. C., N. K. Sutton and N. W. Hatch (2010): "Dual-track versus single-track sell-outs: An empirical analysis of competing harvest strategies", *Journal of Business Venturing* 25(4), 389–402.

Brav, A. and P. A. Gompers (2003): "The role of lockups in initial public offerings", *The Review of Financial Studies* 16(1), 1–29.

Breiman, L. (2001): "Random forests", *Machine Learning* 45(1), 5–32.

Breiman, L. and A. Cutler (2003): *Manual - Setting Up, Using, And Understanding Random Forests V4.0: Statistics Department University of California Berkeley*, Available at: `https://www.stat.berkeley.edu/~breiman/Using_random_forests_v4.0.pdf`.

Brennan, M. J. and A. Subrahmanyam (1995): "Investment analysis and price formation in securities markets", *Journal of Financial Economics* 38(3), 361–381.

Brezina, I., J. Pekár, Z. Čičková and M. Reiff (2016): "Herfindahl–Hirschman index level of concentration values modification and analysis of their change", *Central European Journal of Operations Research* 24(1), 49–72.

Brooks, C. (2014): *Introductory Econometrics for Finance*, 3rd ed., Cambridge University Press.

Busaba, W. Y. (2006): "Bookbuilding, the option to withdraw, and the timing of IPOs", *Journal of Corporate Finance* 12(2), 159–186.

Busaba, W. Y., L. M. Benveniste and R. J. Guo (2001): "The option to withdraw IPOs during the premarket: empirical analysis", *Journal of Financial Economics* 60(1), 73–102.

Busaba, W. Y., Y. Li and G. Yang (2015): "Market Volatility and IPO Filing Activity", *Quarterly Journal of Finance* 5(04), 1550017.

Butler, A. W., M. O. C. Keefe and R. Kieschnick (2014): "Robust determinants of IPO underpricing and their implications for IPO research", *Journal of Corporate Finance* 27, 367–383.

Carter, R. B., F. H. Dark and A. K. Singh (1998): "Underwriter reputation, initial returns, and the long-run performance of IPO stocks", *The Journal of Finance* 53(1), 285–311.

Carter, R. B. and S. Manaster (1990): "Initial public offerings and underwriter reputation", *The Journal of Finance* 45(4), 1045–1067.

Celikyurt, U., M. Sevilir and A. Shivdasani (2010): "Going public to acquire? The acquisition motive in IPOs", *Journal of Financial Economics* 96(3), 345–363.

Chawla, N. V. (2009): "Data mining for imbalanced datasets: An overview", *Data mining and knowledge discovery handbook*, Springer, 875–886.

Chawla, N. V., K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer (2002): "SMOTE: synthetic minority over-sampling technique", *Journal of Artificial Intelligence Research* 16, 321–357.

Chemmanur, T. J. and P. Fulghieri (1999): “A theory of the going-public decision”, *The Review of Financial Studies* 12(2), 249–279.

Chemmanur, T. J. and J. He (2011): “IPO waves, product market competition, and the going public decision: Theory and evidence”, *Journal of Financial Economics* 101(2), 382–412.

Chemmanur, T. J., S. He and D. K. Nandy (2009): “The going-public decision and the product market”, *The Review of Financial Studies* 23(5), 1855–1908.

Chen, G., N. Sutton and J. Qi (2017): “From setback to comeback: Motivations for withdrawn IPO firms to return”, *The Quarterly Review of Economics and Finance* 66, 259–264.

Chen, N., B. Ribeiro and A. Chen (2016): “Financial credit risk assessment: a recent review”, *Artificial Intelligence Review* 45(1), 1–23.

Cheng, L. T.W. and J. E. McDonald (1996): “Industry structure and ripple effects of bankruptcy announcements”, *Financial Review* 31(4), 783–807.

Chod, J. and E. Lyandres (2011): “Strategic IPOs and product market competition”, *Journal of Financial Economics* 100(1), 45–67.

Cohen, J. (1960): “A coefficient of agreement for nominal scales”, *Educational and Psychological Measurement* 20(1), 37–46.

Cooney, J. W., T. Moeller and M. Stegemoller (2009): “The underpricing of private targets”, *Journal of Financial Economics* 93(1), 51–66.

Cooper, T. E. (1989): “Indirect competition with spatial product differentiation”, *The Journal of Industrial Economics*, 241–257.

Corwin, S. A. and P. Schultz (2005): “The role of IPO underwriting syndicates: Pricing, information production, and underwriter competition”, *The Journal of Finance* 60(1), 443–486.

Cotei, C. and J. Farhat (2013): “Informational externalities of initial public offerings: Does venture capital backing matter?”, *Journal of Economics and Finance* 37(1), 80–99.

Cumming, D. (2008): “Contracts and exits in venture capital finance”, *The Review of Financial Studies* 21(5), 1947–1982.

Cumming, D., A. Knill and K. Syvrud (2016): “Do international investors enhance private firm value? Evidence from venture capital”, *Journal of International Business Studies* 47(3), 347–373.

Derrien, F. and A. Kecskés (2013): “The real effects of financial shocks: Evidence from exogenous changes in analyst coverage”, *The Journal of Finance* 68(4), 1407–1440.

Derrien, F., A. Kecskés and S. A. Mansi (2016): “Information asymmetry, the cost of debt, and credit events: Evidence from quasi-random analyst disappearances”, *Journal of Corporate Finance* 39, 295–311.

Dunbar, C. (1998): “The choice between firm-commitment and best-efforts offering methods in IPOs: The effect of unsuccessful offers”, *Journal of Financial Intermediation* 7(1), 60–90.

Dunbar, C. and S. Foerster (2008): “Second time lucky? Withdrawn IPOs that return to the market”, *Journal of Financial Economics* 87(3), 610–635.

Edelen, R. M. and G. B. Kadlec (2005): “Issuer surplus and the partial adjustment of IPO prices to public information”, *Journal of Financial Economics* 77(2), 347–373.

Ellul, A. and M. Pagano (2006): “IPO underpricing and after-market liquidity”, *The Review of Financial Studies* 19(2), 381–421.

Esfahanipour, A., M. Goodarzi and R. Jahanbin (2016): “Analysis and forecasting of IPO underpricing”, *Neural Computing and Applications* 27(3), 651–658.

Fan, P. and K. Yamada (2020): “Same Bed Different Dream: The Analysis of Composition of IPO Shares and Withdrawal Decision”, *Small Business Economics* 55, 955–974.

Fawcett, T. (2006): “An introduction to ROC analysis”, *Pattern Recognition Letters* 27(8), 861–874.

Ferri, C., J. Hernández-Orallo and R. Modroiu (2009): “An experimental comparison of performance measures for classification”, *Pattern Recognition Letters* 30(1), 27–38.

Ferris, S. P., N. Jayaraman and A. K. Makhija (1997): “The response of competitors to announcements of bankruptcy: An empirical examination of contagion and competitive effects”, *Journal of Corporate Finance* 3(4), 367–395.

Firth, M. (1976): “The impact of earnings announcements on the share price behaviour of similar type firms”, *The Economic Journal* 86(342), 296–306.

Flammer, C. (2015): “Does product market competition foster corporate social responsibility? Evidence from trade liberalization”, *Strategic Management Journal* 36(10), 1469–1485.

Frésard, L. (2010): “Financial strength and product market behavior: The real effects of corporate cash holdings”, *The Journal of Finance* 65(3), 1097–1122.

Frésard, L. and P. Valta (2016): “How does corporate investment respond to increased entry threat?”, *The Review of Corporate Finance Studies* 5(1), 1–35.

Friedman, J., T. Hastie and R. Tibshirani (2009): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer series in statistics.

— (2010): “Regularization paths for generalized linear models via coordinate descent”, *Journal of Statistical Software* 33(1), 1.

Gama, J., I. Žliobaitė, A. Bifet, M. Pechenizkiy and A. Bouchachia (2014): “A survey on concept drift adaptation”, *ACM Computing Surveys (CSUR)* 46(4), 1–37.

Gao, J., W. Fan, J. Han and P. S. Yu (2007): “A general framework for mining concept-drifting data streams with skewed distributions”, *Proceedings of the 2007 siam international conference on data mining*, Society for Industrial and Applied Mathematics, 3–14.

Guo, B., D. Pérez-Castrillo and A. Toldrà-Simats (2019): “Firms’ innovation strategy under the shadow of analyst coverage”, *Journal of Financial Economics* 131(2), 456–483.

Hanley, J. A. and B. J. McNeil (1982): “The meaning and use of the area under a receiver operating characteristic (ROC) curve.”, *Radiology* 143(1), 29–36.

Hao, Q. (2011): “Securities litigation, withdrawal risk and initial public offerings”, *Journal of Corporate Finance* 17(3), 438–456.

He, J. Jack and X. Tian (2013): “The dark side of analyst coverage: The case of innovation”, *Journal of Financial Economics* 109(3), 856–878.

Helbing, P. (2019): “A review on IPO withdrawal”, *International Review of Financial Analysis* 62, 200–208.

Helbing, P., B. M. Lucey and S. A. Vigne (2019): “The determinants of IPO withdrawal – Evidence from Europe”, *Journal of Corporate Finance* 56, 415–436.

Henrique, B. M., V. A. Sobreiro and H. Kimura (2019): "Literature review: Machine learning techniques applied to financial market prediction", *Expert Systems with Applications* 124, 226–251.

Hertzel, M. G. (1991): "The effects of stock repurchases on rival firms", *The Journal of Finance* 46(2), 707–716.

Ho, T. K. (1995): "Random decision forests", *Proceedings of 3rd international conference on document analysis and recognition*, vol. 1, IEEE, 278–282.

Hoberg, G. and G. Phillips (2016): "Text-based network industries and endogenous product differentiation", *Journal of Political Economy* 124(5), 1423–1465.

Holmström, B. and J. Tirole (1993): "Market liquidity and performance monitoring", *Journal of Political Economy* 101(4), 678–709.

Hong, H. and M. Kacperczyk (2010): "Competition and bias", *The Quarterly Journal of Economics* 125(4), 1683–1725.

Hong, H. and J. D. Kubik (2003): "Analyzing the analysts: Career concerns and biased earnings forecasts", *The Journal of Finance* 58(1), 313–351.

Hsieh, J., E. Lyandres and A. Zhdanov (2011): "A theory of merger-driven IPOs", *Journal of Financial and Quantitative Analysis* 46(5), 1367–1405.

Hsu, H.C., A. V. Reed and J. Rocholl (2010): "The new game in town: Competitive effects of IPOs", *The Journal of Finance* 65(2), 495–528.

Huang, J. and C. X. Ling (2005): "Using AUC and accuracy in evaluating learning algorithms", *IEEE Transactions on Knowledge and Data Engineering* 17(3), 299–310.

Ibbotson, R. G. and J. F. Jaffe (1975): ""Hot issue" markets", *The Journal of Finance* 30(4), 1027–1042.

Imbens, G. W. and D. B. Rubin (2015): *Causal Inference in Statistics, Social, and Biomedical Sciences*, Cambridge University Press.

Iqbal, Z. (2002): "The effects of bankruptcy filings on the competitors' earnings", *International Review of Economics & Finance* 11(1), 85–99.

Jain, B. A. and O. Kini (1994): "The post-issue operating performance of IPO firms", *The Journal of Finance* 49(5), 1699–1726.

Jain, B. A. and B. N. Nag (1995): "Artificial neural network models for pricing initial public offerings", *Decision Sciences* 26(3), 283–302.

— (1998): "A neural network model to predict long-run operating performance of new ventures", *Annals of Operations Research* 78, 83–110.

James, G., D. Witten, T. Hastie and R. Tibshirani (2013): *An Introduction to Statistical Learning*, Springer.

Janitza, S., C. Strobl and A. L. Boulesteix (2013): "An AUC-based permutation variable importance measure for random forests", *BMC Bioinformatics* 14(1), 1–11.

Japkowicz, N. (2001): "Concept-learning in the presence of between-class and within-class imbalances", *Conference of the Canadian society for computational studies of intelligence*, Springer, 67–77.

Jegadeesh, N., J. Kim, S. D. Krische and C. M. C. Lee (2004): "Analyzing the analysts: When do recommendations add value?", *The Journal of Finance* 59(3), 1083–1124.

Jensen, M. C. (1986): "Agency costs of free cash flow, corporate finance, and takeovers", *The American Economic Review* 76(2), 323–329.

Jensen, M. C. and W. H. Meckling (1976): "Theory of the firm: Managerial behavior, agency costs and ownership structure", *Journal of Financial Economics* 3(4), 305–360.

Jorion, P. and G. Zhang (2007): "Good and bad credit contagion: Evidence from credit default swaps", *Journal of Financial Economics* 84(3), 860–883.

Kahle, K. M. and R. A. Walkling (1996): "The impact of industry classifications on financial research", *Journal of Financial and Quantitative Analysis* 31(3), 309–335.

Kelly, B. and A. Ljungqvist (2012): "Testing asymmetric-information asset pricing models", *The Review of Financial Studies* 25(5), 1366–1413.

Kim, W. and M. S. Weisbach (2008): "Motivations for public equity offers: An international perspective", *Journal of Financial Economics* 87(2), 281–307.

Kothari, S. P., E. So and R. Verdi (2016): "Analysts' forecasts and asset pricing: A survey", *Annual Review of Financial Economics* 8, 197–219.

Krigman, L., W. H. Shaw and K. L. Womack (2001): "Why do firms switch underwriters?", *Journal of Financial Economics* 60(2-3), 245–284.

Krishnan, C.N.V., V. I. Ivanov, R. W. Masulis and A. K. Singh (2011): “Venture capital reputation, post-IPO performance, and corporate governance”, *Journal of Financial and Quantitative Analysis*, 1295–1333.

Landis, J. R. and G. G. Koch (1977): “The measurement of observer agreement for categorical data”, *Biometrics*, 159–174.

Lang, L. H. P. and R. M. Stulz (1992): “Contagion and competitive intra-industry effects of bankruptcy announcements: An empirical analysis”, *Journal of Financial Economics* 32(1), 45–60.

Latham, S. and M. R. Braun (2010): “To IPO or not to IPO: Risks, uncertainty and the decision to go public”, *British Journal of Management* 21(3), 666–683.

Lee, P. M., T. G. Pollock and K. Jin (2011): “The contingent value of venture capitalist reputation”, *Strategic Organization* 9(1), 33–69.

Leland, H. E. and D. H. Pyle (1977): “Informational asymmetries, financial structure, and financial intermediation”, *The Journal of Finance* 32(2), 371–387.

Lemmon, M. L. and J. F. Zender (2010): “Debt capacity and tests of capital structure theories”, *Journal of Financial and Quantitative Analysis* 45(5), 1161–1187.

Lerner, J. (1994): “Venture capitalists and the decision to go public”, *Journal of Financial Economics* 35(3), 293–316.

Lian, Q. and Q. Wang (2009): “Does the Market Incorporate Previous IPO Withdrawals When Pricing Second-Time IPOs?”, *Financial Management* 38(2), 357–380.

— (2012): “Acquisition valuations of withdrawn IPOs: When IPO plans turn into mergers”, *Journal of Banking & Finance* 36(5), 1424–1436.

Liaw A.and Wiener, M. (2002): “Classification and regression by randomForest”, *R News* 2(3), 18–22.

Lin, W. C., C.F. Tsai, Y. H. Hu and J. S. Jhang (2017): “Clustering-based undersampling in class-imbalanced data”, *Information Sciences* 409, 17–26.

Lin, W. Y., Y. H. Hu and C. F. Tsai (2011): “Machine learning in financial crisis prediction: a survey”, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42(4), 421–436.

Ljungqvist, A. (2007): “IPO underpricing”, *Handbook of Empirical Corporate Finance*, 375–422.

Lobo, J. L., J. Del Ser, M. N. Bilbao, C. Perfecto and S. Salcedo-Sanz (2018): “DRED: An evolutionary diversity generation method for concept drift adaptation in online learning environments”, *Applied Soft Computing* 68, 693–709.

Loh, R. K. and R. M. Stulz (2011): “When are analyst recommendation changes influential?”, *The Review of Financial Studies* 24(2), 593–627.

Loughran, T. (1993): “NYSE vs NASDAQ returns: Market microstructure or the poor performance of initial public offerings?”, *Journal of Financial Economics* 33(2), 241–260.

Loughran, T. and J. R. Ritter (1995): “The new issues puzzle”, *The Journal of Finance* 50(1), 23–51.

— (2004): “Why has IPO underpricing changed over time?”, *Financial Management*, 5–37.

Lowry, M. (2003): “Why does IPO volume fluctuate so much?”, *Journal of Financial Economics* 67(1), 3–40.

Lowry, M. and G. W. Schwert (2002): “IPO market cycles: Bubbles or sequential learning?”, *The Journal of Finance* 57(3), 1171–1200.

Lucas, D. J. and R. L. McDonald (1990): “Equity issues and stock price dynamics”, *The Journal of Finance* 45(4), 1019–1043.

Maksimovic, V. and P. Pichler (2001): “Technological innovation and initial public offerings”, *The Review of Financial Studies* 14(2), 459–494.

McGilvery, A., R. Faff and S. Pathan (2012): “Competitive valuation effects of Australian IPOs”, *International Review of Financial Analysis* 24, 74–83.

McWilliams, A. and D. Siegel (1997): “Event studies in management research: Theoretical and empirical issues”, *Academy of Management Journal* 40(3), 626–657.

Moreno-Torres, J. G., T. Raeder, R. Alaiz-RodríGuez, N. V. Chawla and F. Herrera (2012): “A unifying view on dataset shift in classification”, *Pattern Recognition* 45(1), 521–530.

Muthuramu, P. and T. U. Maheswari (2019): “Tests for structural breaks in time series analysis: A review of recent development”, *International Journal of Economics* 7(4), 66–79.

Norton, E. C., H. Wang and C. Ai (2004): "Computing interaction effects and standard errors in logit and probit models", *The Stata Journal* 4(2), 154–167.

Owen-Smith, J., N. C. Cotton-Nessler and H. Buhr (2015): "Network effects on organizational decision-making: Blended social mechanisms and IPO withdrawal", *Social Networks* 41, 1–17.

Pagano, M., F. Panetta and L. Zingales (1998): "Why do companies go public? An empirical analysis", *The Journal of Finance* 53(1), 27–64.

Page, E. S. (1954): "Continuous inspection schemes", *Biometrika* 41(1/2), 100–115.

Pástor, L., L. A. Taylor and P. Veronesi (2009): "Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability", *The Review of Financial Studies* 22(8), 3005–3046.

Pástor, L. and P. Veronesi (2005): "Rational IPO waves", *The Journal of Finance* 60(4), 1713–1757.

Perron, P. (2006): "Dealing with structural breaks", *Palgrave Handbook of Econometrics* 1(2), 278–352.

Quintana, D., Y. Sáez and P. Isasi (2017): "Random forest prediction of IPO underpricing", *Applied Sciences* 7(6), 636.

Reber, B., B. Berry and S. Toms (2005): "Predicting mispricing of initial public offerings", *Intelligent Systems in Accounting, Finance & Management: International Journal* 13(1), 41–59.

Ritter, J. R. (1984): "The "hot issue" market of 1980", *Journal of Business*, 215–240.

— (1987): "The costs of going public", *Journal of Financial Economics* 19(2), 269–281.

— (1991): "The long-run performance of initial public offerings", *The Journal of Finance* 46(1), 3–27.

Ritter, J. R. and I. Welch (2002): "A review of IPO activity, pricing, and allocations", *The Journal of Finance* 57(4), 1795–1828.

Rock, K. (1986): "Why new issues are underpriced", *Journal of Financial Economics* 15(1-2), 187–212.

Schill, M. J. (2004): "Sailing in rough water: market volatility and corporate finance", *Journal of Corporate Finance* 10(5), 659–681.

SEC (2017): *What is a registration statement?*, Available at: `https://www.sec.gov/smallbusiness/goingpublic/registrationstatement` (visited on 09/09/2021).

— (2019): *Going public*, Available at: `https://www.sec.gov/smallbusiness/goingpublic` (visited on 09/09/2021).

Shelegia, S. (2012): "Is the Competitor of my Competitor also my Competitor?", *Journal of Economics & Management Strategy* 21(4), 927–963.

Shi, Y., W. M. Liu and K. Y. Ho (2016): "Public news arrival and the idiosyncratic volatility puzzle", *Journal of Empirical Finance* 37, 159–172.

Signori, A. (2018): "Zero-revenue IPOs", *International Review of Financial Analysis* 57, 106–121.

Slovin, M. B., M. E. Sushka and Y. M. Bendeck (1991): "The intra-industry effects of going-private transactions", *The Journal of Finance* 46(4), 1537–1550.

Slovin, M. B., M. E. Sushka and S. R. Ferraro (1995): "A comparison of the information conveyed by equity carve-outs, spin-offs, and asset sell-offs", *Journal of Financial Economics* 37(1), 89–104.

Slovin, M. B., M. E. Sushka and J. A. Polonchek (1992): "Informational externalities of seasoned equity issues: Differences between banks and industrial firms", *Journal of Financial Economics* 32(1), 87–101.

Spence, M. (1974): *Market signaling: Informational transfer in hiring and related screening processes*, 143, Cambridge: harvard university press.

Spiegel, M. and H. Tookes (2020): "Why does an IPO affect rival firms?", *The Review of Financial Studies* 33(7), 3205–3249.

Statista (2021): *Value of initial public offerings (IPOs) in the United States from 2000 to 2020*, Available at: `https://www.statista.com/statistics/264607/ipo-volume-in-the-us/` (visited on 11/09/2021).

Stoughton, N. M., K. P. Wong and J. Zechner (2001): "IPOs and product quality", *The Journal of Business* 74(3), 375–408.

Strobl, C., A. L. Boulesteix, A. Zeileis and T. Hothorn (2007): "Bias in random forest variable importance measures: Illustrations, sources and a solution", *BMC Bioinformatics* 8(1), 25.

Strobl, C., A. L. Boulesteix, T. Kneib, T. Augustin and A. Zeileis (2008): “Conditional variable importance for random forests”, *BMC Bioinformatics* 9(1), 1–11.

Subrahmanyam, A. and S. Titman (1999): “The going-public decision and the development of financial markets”, *The Journal of Finance* 54(3), 1045–1082.

Tian, S., Y. Yu and H. Guo (2015): “Variable selection and corporate bankruptcy forecasts”, *Journal of Banking & Finance* 52, 89–100.

Tian, X. (2012): “The role of venture capital syndication in value creation for entrepreneurial firms”, *Review of Finance* 16(1), 245–283.

Tibshirani, R. (1996): “Regression shrinkage and selection via the lasso”, *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1), 267–288.

Titman, S. and B. Trueman (1986): “Information quality and the valuation of new issues”, *Journal of Accounting and Economics* 8(2), 159–172.

Valta, P. (2012): “Competition and the cost of debt”, *Journal of Financial Economics* 105(3), 661–682.

Wald, A. (1945): “Sequential tests of statistical hypotheses”, *The Annals of Mathematical Statistics* 16(2), 117–186.

Weiss, G. M. and F. Provost (2003): “Learning when training data are costly: The effect of class distribution on tree induction”, *Journal of Artificial Intelligence Research* 19, 315–354.

Womack, K. L. (1996): “Do brokerage analysts’ recommendations have investment value?”, *The Journal of Finance* 51(1), 137–167.

Wooldridge, J. M. (2009): *Introductory Econometrics: A Modern Approach*, 4th ed., South-Western Cengage Learning.

Wu, J. S. and A. Y. Zang (2009): “What determine financial analysts’ career outcomes during mergers?”, *Journal of Accounting and Economics* 47(1-2), 59–86.

Yan, M. M. W. (2020): “Accurate detecting concept drift in evolving data streams”, *ICT Express*.

Yen, S. J. and Y. S. Lee (2009): “Cluster-based under-sampling approaches for imbalanced data distributions”, *Expert Systems with Applications* 36(3), 5718–5727.

Zingales, L. (1995): “Insider ownership and the decision to go public”, *The Review of Economic Studies* 62(3), 425–448.

Zou, Hui (2006): “The adaptive lasso and its oracle properties”, *Journal of the American statistical association* 101(476), 1418–1429.

Declaration of co-authorship

UNIVERSITÄT HOHENHEIM
FAKULTÄT WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN

KO-AUTORENERKLÄRUNG *DECLARATION OF CO-AUTHORSHIP*
(Für kumulative Dissertationen)

Name des Kandidaten: (*Name of the candidate*)	Annika Reiff

Titel des Artikels (*Title of the article*):

IPO withdrawals: Are corporate governance and VC characteristics the guiding light in the rough sea of volatile markets?

☐ nicht eingereicht (*not submitted*)

☐ eingereicht bei (*submitted to*):

☒ Zur Veröffentlichung angenommen oder veröffentlicht in (*accepted for publication or published in*):

Journal of Corporate Finance, 67, 101908. Available at: https://doi.org/10.1016/j.jcorpfin.2021.101908

Arbeitsanteil des Kandidaten an vorgenanntem Artikel *Quantification of candidates contribution to the article (overall)***:**

☐ hat zur Arbeit beigetragen/*has contributed to the work* (<1/3)

☒ hat wesentlich zur Arbeit beigetragen/*has made a substantial contribution* (1/3 to 2/3)

☐ hat einen Großteil der Arbeit allein erledigt/*did the majority of the work independently* (>2/3)

Ko-Autoren *Co-authors* (Name und Kontaktdaten/*full name; contact*):

1.	Tereza Tykvová,	Swiss Institute of Banking and Finance (s/bf-HSG), University of St.Gallen (HSG), Unterer Graben 21, CH-9000 St.Gallen, tereza.tykvova@unisg.ch
2.		
3.		
4.		

Hiermit bestätige ich die Richtigkeit des oben beschriebenen Arbeitsanteils des Kandidaten.
I hereby confirm the candidate's contribution as quantified above.

St.Gallen, 29 March 2022	
Ort, Datum *Place, Date*	Unterschrift Ko-Autor *Signature (Co-Author)*

Stand 11/2016. Alle personenbezogenen Angaben gelten stets für Frauen und Männer gleichermaßen. 1 I 1