Cost of IPOs, IPO dynamics in the short and long-run and value of textual tone of IPO prospectus

Mazen Alzamel

Doctor of Philosophy

University of York

Management

June 2016

Abstract

This thesis is an attempt to understand few aspects of the initial public offerings (IPO) market using the data from the London Stock Exchange (LSE) for both the Main Market (MM) and the Alternative Investment Market (AIM) segments. I examined the cost of raising money for the IPOs. I found that IPO spread is in the range of 4% - 6.43% and a median of 5%, which contrasts with previous US studies that report a spread of 7%. More interestingly, there is no clustering of fees at any level in the LSE. Economies of scale explain the spread charged in the AIM better than the MM, implying the spread decreases when the proceed increases.

As the final product of the IPO preparation stage is the IPO prospectus, I conducted a textual analysis to measure the tones' relationship with IPO short-run and long-run dynamics. I used a novel methodology of analysis the textual tone from the IPO prospectus utilising Loughrun and McDonald (2011) modified word list. I found that IPO tone is related to underpricing, spread and lockup length. However, I did not find any relationship with the volatility or idiosyncratic risk. That may be due to the fact that these are not useful measures of ex-ante uncertainty.

I also examined some of the long-run anomalies in the IPO literature. The tone measures show a significant relationship with the survival of the firms and time till dead. I showed that the information disclosed at the time of IPO is still valid for predicting the survival of IPO firms. However, I do not find much correlation with the IPO tone measures and lockup expiration returns and long-run IPO performance. That might be a result of methodological controversies of measuring the long-run performance of the IPO.

Keywords: AIM, Main market, London Stock Exchange, Underwriters spread, IPOs, lockup, underpricing, risk, Underpricing, spread, lockup length, volatility, idiosyncratic risk, Textual analysis, Lockup expiration returns, long-run IPO performance, survival analysis.

JEL Code: G14, G15, G24, G30, G32

List of Contents

List of Tables

List of Figures

List of Equations

Acknowledgements

I would like to thank several people. My supervisor, Dr Hafiz Hoque, for his support and insightful guidance throughout my PhD program. Dr Viktor Manahov for his suggestions. My Thesis Advisory Panel member, Dr Keith Anderson for his constructive comments and feedback. And most of all, my caring mother, loving wife and children for being supportive and patient. My father, who passed before I start my program, for encouraging me to start.

Author's Declaration

Except where stated otherwise, all material presented in this thesis represents the independent and original work of the author, which has neither been previously published nor is currently under consideration for publication in any form.

Chapter 1

1. Introduction

In this chapter, I give a brief introduction about the thesis and shade some light on the IPO literature. In the next section, I give a brief about the main findings and contribution of the thesis.

Initial Public Offerings (IPOs) are one of the most important events during the life cycle of a typical firm. There has been a sizable literature on numerous issues of IPOs including the first day returns, the cost of raising equity finance, underwriter price support, long-run IPO performance, survival of the newly listed firms, designs of the lockups, lockup expiration returns. Despite enormous research was done during the last 40 years there are few anomalies remain in the IPO literature: i) underpricing ii) long-run underperformance and iii) lockup expiration returns. The purpose of this thesis is to shed light on some of the unanswered questions in the IPO literature.

Since IPO is a costly event for the growth firms, researchers examine the IPO gross spreads in detail and the determinants of IPO gross spread. One of the pioneering studies in IPO gross spread is Chen and Ritter (2000). They report an IPO gross spread of 7% for most of the US IPOs. They conjecture that charging exactly 7% may be the result of some collusive practice by the investment banks. A recent study by Abrahamson, Jenkinson and Jones (2011) show that this is still the norm to charge 7%. Hansen (2001) reject that 7% is a result of collusive behaviour on the part of underwriters and it is the efficient contracting mechanism. Though Abrahamson et al. (2011) looked into the cost of raising equity in the Europe and UK (Main Market), there exists no study which examines the cost of raising money in the Alternative

Investment Market, which is the most popular second market in the world. I focus on the IPO gross spreads in the Alternative Investment Market in the United Kingdom.

First-day return or IPO underpricing is one of the most researched areas in IPO literature. The empirical studies report the various level of underpricing for different countries. For example, the US studies report an average underpricing of 7.6% (Ritter, 2016) and the UK studies report an underpricing of 16.7% (see Chambers and Dimson, 2009). The reasons for underpricing includes winners curse (Benvenist and Spindt, 1989), asymmetric information (Rock, 1986), IPOs as a marketing tool (Demers and Lewellen, 2003), information cascades (Welch, 1992), lawsuit avoidance (Hughes and Thakor, 1992; Tinic, 1988), behavioural biases (Welch, 1992, Loughran and Ritter, 2002). Still, it is not clear the reasons for underpricing. This thesis is another effort to examine the underpricing by using the tone measures from the IPO prospectus.

Lockups are voluntary contracts between the IPO underwriters and insiders of a company not to sell the shares before a stipulated period. It is to provide confidence for the investors in the IPOs where they are putting their money. Even lockups are voluntary agreements; they are present in almost all the IPOs. Why do lockups exist? Academic studies provide three explanations: i) asymmetric information ii) moral hazard and iii) agency problem (for a detailed discussion, please see Hoque (2011)). It is not clear why lockups exist and what determines their length. It is more puzzling than in the US where the lockup length is exactly 180 days for most of the IPOs. UK market provides a more interesting case where the lockup lengths are longer and diverse. I extend the information asymmetry explanation of lockups by extracting information from IPO prospectuses.

Hoque and Lasfer (2016) examine changes in director ownership during lockup periods. They find that these transactions are not additional signalling devices. Instead, they are contractual arrangements between directors and underwriters, as directors increase their holdings after a significant price decline, in line with the price support hypothesis, but they decrease their ownership after price run-up, consistent with the early lockup releases hypothesis. They also examine the rent-seeking hypotheses—whether underwriters gain from other transactions like seasoned equity offering, insider buys and sells. They find some evidence in line with this.

Though lockup expiration is a known event and clearly mentioned in the IPO prospectus and some databases report the lockup expiration date (at least for the US companies), there should not be any market reaction around the lockup expiration day. However, there is -1.50% (-2.00%) drop in share price in the US (UK). It is quite puzzling. Field and Hanka (2001) examine a number of hypothesis including the downward sloping demand curve, worse than expected insider selling, an increase in number of trade at bid prices, price pressure, trading costs, change in ownership due to distribution of shares by venture capitalists to their partners around unlock day and 180-day effect. Though they find partial support for several hypotheses that they test, their evidence is far from conclusive.

Another remaining puzzle in the IPO literature is IPOs underperform as compared to some benchmarks over 3-5 year period. The underperformance persists after the seminal work by Ibbotson (1978), and it is persistent across countries. A number of explanations have been put forward to explain the long-run underperformance because short selling is not possible for small public floats (Miller (1977) and Morris (1996), heterogeneous expectations among the investors (Ritter

(1991) and Loughran and Ritter (1995)), behavioural timing (Loughran and Ritter (2000)), and methodological problems what has been termed as "pseudo market timing" (Schultz (2003)). However, it remains a puzzle why IPOs underperform. This thesis sheds some light on underperformance using the tome measure from IPO prospectus.

One of the important issue in the IPO literature is the survival of the IPOs**.** Hensler, Rutherford and Springer (1997) examine the effects of various firm and IPO characteristics on the survival of IPO firms by applying an accelerated failure time (AFT) model. They find that survival time decreases with the increases in some risk factors mentioned in the prospectus. Demers and Joos (2007) develop an IPO failure risk model by incorporating accounting and other deal specific characteristics. Espenlaub, Khurshed and Mohamed (2012) examine whether the IPOs that join the Alternative Investment Market with more prestigious underwriters survives longer and they find evidence in support of that. Joining the market with prestigious underwriter could mitigate the information asymmetry, or it might be possible that prestigious underwriters bring good quality IPOs. Ahmad and Jelic (2014) examine whether the lockup agreements have any impact on the survival of UK Initial Public Offerings (IPOs) between the time of 1990–2011. They find a positive association between the lockup length and IPO survival in the UK market. To the best of our knowledge, no previous study considers the information from IPO prospectus to relate to the survival of the IPOs.

Recently, there is a growing number of contribution in accounting and finance literature which uses the tone and sentiment of various documents and announcements made by the company such as annual reports, news announcements, management

discussions and analysis (MD&A), articles in newspapers, and message boards for the investors. Examples include Antweiler and Frank (2004), Tetlock (2007), Engelberg (2008), Li (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008). The findings of these studies may be summarised as negative words show a significant relationship with different company level financial measures. Kothari, Li, and Short (2008) relate the tone of daily paper articles with the cost of capital, the variability of return, and forecasts by analysts. Henry (2008), Engelberg (2008), and Demers and Vega (2008) relates news releases with company earnings, drift in earnings, or share price returns. Some papers relate the information content of IPO prospectus with the share returns, price variability and trading volume (e.g., Li (2008, 2009), Feldman, Govindaraj, Livnat and Segal (2008), Hanley and Hoberg (2010)).

Previous studies are subject to numerous criticisms. Most of the previous studies use Harvard dictionary words to measure the tone of the IPO. However, Loughrun and Mcdonald (2010) criticise when a liability is not a liability as such. For example, the word “liability” is not a negative word but Harvard dictionary defines this as a negative word. This thesis use Loughrun and Mcdonald (2010) refined dictionary. Secondly, there is not much variation in the IPO gross spread and lockup length in the US, where most studies on IPOs are done. I used data from the UK where there is much variation in the IPO gross spread and lockup length.

While I fully acknowledge the previous work on IPOs, I considered the existing gaps and unsolved puzzles in the IPO literature to take advantage of the existing gaps to further explore the information content in the IPO prospectus and relate it to the different IPO phenomenon in the short and long-run. The IPO tone measures using the prospectus in the UK can provide additional insights and help us

understand the massive growth of the Alternative Investment Market and thus may help answer some previously unanswered questions in the IPO literature.

This thesis is such an initiative to examine the information content through textual tones and use this information to address some puzzles remain in IPO literature. For this purpose, I use a crucial corporate finance event in a firm's life cycle, called initial public offerings to shed light on some of the puzzles in the IPO literature; this thesis uses the information provided by the "tone" of IPO prospectus. In particular, the study uses the tone measures using the Loughrun and Mcdonald (2010) dictionary measures. Specifically, this study uses *positive, negative, superfluous* tone measures of IPO prospectus to address the IPO gross spread, underpricing, lockup length, lockup expiration returns, the long-run performance of IPOs and survival of IPO.

However, the objective of the theses is not to provide a closed form solution for the puzzles that remain in the IPO literature. The purpose of this thesis is to examine further and to shed additional light on these puzzles, which might help us in closing the gap between where I stand regarding the understanding of IPOs and the empirical puzzles that remain in the IPO literature. Specifically, this research empirically investigates what information is provided by the IPO prospectuses' tone? What is the role of information provided by different types of words for newly listed companies?

The issue of the information content of IPO prospectus is still debatable, and this study contributes to the literature by paying attention to the empirical debate of whether "tone" of IPO prospectus are informative by using more recent UK data. The first objective of the thesis is to examine whether IPO prospectuses contain significant information using an updated word lists from Loughrun and McDonald (2010).

Second, do the tone measures of IPO prospectus is related to short-run IPO dynamics like underpricing, gross spread and lockup length? Third, do the tone measures of IPO prospectus is linked to the long-run IPO dynamics such as lockup expiry returns, the long-run performance of IPOs and survival of IPOs? The empirical research questions in this study are motivated by the gaps and controversies in the literature that remains, the different nature of the UK market compared to the US where the IPO market is different in terms of IPO method, lockup length, quiet period, price support by the underwriters.[1]

In Saudi Arabia, there has been a large number of family-owned businesses that have been operating for a range of 80 – 20 years. Some of those companies grew up and became of a size that is comparable to some of the public companies. On the other hand, just until about the year 2000, most of the listed companies in Saudi Arabia were from the banking and utility sectors. There have been some other companies in some other sectors such as retail and industry. Recently, many of the family-owned business and some of the newly incorporated companies started to seek the opportunity of going public. With the increasing number of companies going public in Saudi Arabia, I found it to be a good time to contribute to the literature of IPO. This is accompanied with new underwriters and investment banking coming to the market of IPO. Although almost all the IPOs that took place in Saudi Arabia were over-subscribed, there were a few IPOs that failed to raise the required proceeds. In addition, there were a few companies that have bankrupted in just a few years after the

1 In the UK the preferred method of IPO is open offer and placing, whereas in the US the usual method is book building. Lockup length in the US is almost standardised at 180 days whereas in the UK this is much longer and the average is 391 days. There is no quiet period in the UK, whereas there is quiet period of 40 days in the US. There is no price support in the UK, but there are price supports in the US.

IPO. This motivated me to look into the underwriters main public contribution to the IPO which is the prospectuses. I decided to study it in a more mature market such as London Stock Exchange where the transparency is high, and the data is available.

Many studies that covered the IPO literature focus on results of the IPO. However, the process of IPO preparation and undertaking is less covered. I am trying to cover this part by examining the main public output from the underwriters and that is the IPO prospectuses. I look into the prospectuses from 2 main angles. First, I consider the cost associated with the process and cover the direct and indirect cost and their effect on the post-IPO performance. We also look into the tone of the prospectuses that is used by the underwriters in the IPO prospectuses. We measure it and then examine its effect on both the short and long run of the post-IPO performance.

1.1 Main Findings and Contribution

The results of the empirical tests conducted in this thesis show that IPO prospectuses convey information in the case of IPOs. The findings of this thesis are as follows. The first empirical paper (chapter four) finds that, in general, underwriters charge less in the London Stock Exchange as compared to the US market. I have found that the underwriters charge ranges between 4% - 6.43% with a median of 5% over the entire sample. That goes in line with Abrahamson et al. (2011) that the European underwriters charge less than their US counterparts. I also examined the cost of underpricing because underpricing is an indirect cost of raising finance. Our results show that for the AIM market, there is a tendency to underprice more. In comparison to the Main Market, this has been noticed more among smaller and new companies. However, this is more noticed in the first trading day as the discrepancy tends to ease where I can see a convergence towards the IPO issue price on the fifth trading day. It

is shown in this chapter that diseconomies of scale apply to the fees charged by the underwriters as it drops down when the size of the IPO increases.

The second empirical paper (chapter five) focuses on the short-run IPO dynamics and value of textual analysis. Our results show a significant relationship with the dictionary words and IPO underpricing, lockup length and spread. I found that superfluous words and underpricing are significantly related. I also believe that positive words are significantly related to underpricing. My results show a positive relationship between most of the tone measures and IPO spread. Higher information disclosure in the UK increases the spread, which shows the underwriters can gauge the riskiness of the company and charge an appropriate fee for the services they provide. I also found a significant negative relationship between negative, litigious and uncertainty words and lockup length. The results show higher the risk related words the lower the lockup length. To the best of my knowledge, there is no study which relates IPO tone with lockup length. I find that certain words significantly related to the lockup length.

Third empirical paper (Chapter six) focuses on long-run IPO dynamics and value of textual analysis. The finding that IPOs underperform in the long-run is analysed by using the information extracted from the prospectuses. I find a significant relationship with the IPO prospectus tone measure and survival of the UK IPOs. I use time till dead and dead dummy to relate it to the IPO prospectus tones. Our Tobit regression on time till dead and Logit and Cox hazard model on dead dummy show that IPO prospectus tone measures are significantly related to the survival of the firms. I also find that time till dead is negatively related to spread and multi-bookrunners. It shows that risky companies would have paid higher fees and hence they survive less

time and also risky companies would have come to the market with multiple bookrunners. The higher the lockup period, the higher the time companies survive. That is consistent with the previous paper by Ahmad and Jelic (2014).

In summary, this research contributes to the current literature in eight ways. First, this thesis is one of the first efforts to understand the growing popularity of the Alternative Investment Market (AIM) in the UK. I examined the cost of raising money in such a market that has not been addressed in the literature. The cost of raising money in the AIM is high as compared to the London Main Market. However, it is still growing and considered as one of the popular markets around the world. The higher cost is partly attributed to the size of the IPOs and riskiness of the companies that raise equity in the AIM.

Second, the results help understands the role of textual tones of IPO prospectus in the context of going public process. All previous research uses the US data where the market is much different in comparison to the UK. The US market uses book building method whereas the UK uses open offer and placing. The underpricing is much higher in the UK as compared to the US. There is a quiet period after the IPO in the US, but not in the UK. The lockup lengths in the US is standardised and shorter, whereas the lockup lengths are much longer and diverse in the UK. I teased out the previous US findings using IPO tone measures and extend the evidence in the context of a different market setting in the UK. This research extends the previous findings by utilising a unique UK data set.

Third, I used the tone measure to relate it to the higher gross spread in the UK. Our research shows a positive relationship between most of the tone measures and IPO spread. That is somehow consistent with Hanley and Hoberg (2010) who find a

positive relationship between standard content and informative content with the Spread in the US market. During the book building phase of IPO, much information is produced in the US. Higher information disclosure in the UK increases the spread, which shows the underwriters can measure the riskiness of the company and charge an appropriate fee for the services they provide. Thus bookrunners charge according to the risk of the companies as reflected by the IPO tones in the prospectuses.

Fourth, I extended the previous evidence on underpricing that try to relate underpricing with the information asymmetry or risk factors. For instance, Beatty and Ritter (1986) provides evidence that more detail data in the prospectuses regarding the use of proceeds increases underpricing. On the other hand, Leone, Rock and Willenborg, (2007) and Ljungqvist and Wilhelm (2003) find that organisations that are more (less) particular in their disclosure of the utilizations of the funds have lower (higher) underpricing. I extended Beatty and Welch (1996) and Arnold, Fishe, and North (2010) evidence who analyse the risk factors segment of the prospectus and find that more prominent exposure in this area is connected with higher IPO first day returns. I provided an out of sample robustness by using the data from the UK.

Fifth, while previous research shows that lockup contracts originate from information asymmetry (Brav and Gompers, 2003; Brau, Lambson and McQueen, 2005) I used the textual tone measures to relate it to the length of lockup. No previous research has done this, and I filled this gap in the literature. My research extends the previous evidence that the IPO tone information mitigates information asymmetry to some extent and it is related to the lockup lengths in the UK, where more information is produced and disclosed during the lockup period.

Six, this thesis sheds further light on the long-run underperformance of IPOs that remain as a puzzle in the literature. Most previous studies show that IPOs underperform in the long-run using a number of different methodologies. I followed the literature in terms of methodologies. However, I do not find that most IPOs underperform. This study contributes to the literature by relating tone information from IPO prospectuses and long-run IPO performance.

Seventh, survival of the IPOs is an important issue, and I provideed fresh evidence the IPO tone measures could be used to differentiate between the companies which are more likely to survive. While previous studies provide evidence using the accounting and Deal specific characteristics and number of risk factors mentioned in the IPO prospectuses to link it to the survival of IPOs (e.g., Deemers and Joos, 2007; Hensler et al., 1997), no previous studies use the tone measures to relate it to the survival of IPOs. I filled this gap in the literature.

Finally, by constructing unique hand-collected dataset on the IPOs such as gross spread and the "tone" information extracted from the prospectus, this research revisits the empirical evidence on the puzzles reported in the IPO literature. Regarding the UK market, IPO gross spread and lockup lengths are not readily reported in any databases and hence need to be hand-collected. I constructed a database of the IPO gross spread and lockup lengths in the UK. Also, the textual data used in this study is unique as no other studies have done any analysis in the UK, to the best of our knowledge. For this theses, I collected data on all IPOs issued in the UK over the period 1999 to 2011. The findings from this research are therefore up to date and more generalizable as compared to previous studies. As far as I know, no previous studies considered the IPO puzzles in a comprehensive way using the "tone" information from

IPO prospectus. As such, my results provide new insights and thought in terms of choosing the "tone" of IPOs. While some previous literature addresses some of the puzzles in the IPO literature, most are based on US data. Empirical results based on UK data using tone information from IPO prospectuses sheds further light on the empirical puzzles that have been documented in the IPO literature.

The remainder of the thesis is organised as follows. Chapter 2 reviews the related literature for this research. Chapter 3 describes the data and empirical methodology used throughout this thesis. Chapter 4 examines the gross spread in the United Kingdom. Chapter 5 focuses on short-run IPO dynamics and value of textual analysis. In Chapter 6, this thesis examines the long-run IPO dynamics and value of textual analysis. Chapter 7 summarises the findings, concludes the study, and discusses the limitations of this study and suggests some future research directions.

In addition to that, I have noticed that the information presented in the prospectuses take different forms. The regulator in the UK requires the presentation of certain information in the prospectuses. However, they do not require them in a certain format. In addition to that, some of the information presented in the prospectuses are explicit while some are implicit and one has to go through the details to come up with basic information. It would be advisable if the regulator comes up with a regulation same to the rule by the US Securities and Exchange Commission (SEC) to use plain English in the prospectuses. In addition to that, primary information that would affect the investors' decisions should be presented in an implicit and clear way. Till then, investment community can come up with the same. Investment analysts can issue a press where they clearly state the primary information.

Chapter 2

2. Literature review

In this chapter I go through the literature of the IPO dynamics over the short and long run. In the first section, I focused on the short-run and go through some literature regarding dynamics of the gross spread, underpricing, lockup length and ex-ante measure of risk. Then in the second section, I look into the long-run dynamics and discuss the literature of the event study and the return on the lockup expiration date, the long-run underperformance and the survival of the IPOs. Finally, in the third section, we discuss the literature of the IPO tone measures and discuss the textual anlysis.

2.1 Short-run IPO dynamics

2.2 IPO gross spread

(Ritter, 1991) has identified in his paper three main anomalies in the IPO pricing. These are the short-run underpricing when the trading price at the first trading day is higher than the issuing prince, the ‘‘hot issue’’ market phenomenon when there’s a high demand for IPO and the third anomaly is the long-run Underperformance of IPOs. These three are the main puzzles of the IPO.

Various studies have attempted to look at why bookrunners charge a fixed fee of 7% for the US IPOs. While Chen and Ritter (2000) asserted the 7% spread to be a sort of arrangement among the bookrunners, Hansen (2001) did not discover any confirmation of that; rather it is an efficient contracting of IPOs. Though that Hansen (2001) acknowledged that investment banks charge 7% for IPOs, yet, they rather compete on the basis of reputation, placement service, and underpricing.

A leading paper by (Carter and Manaster, 1990) examined 501 IPOs that took place in the US market between 1979 and 1983. In their paper, they looked into the underwriter reputation and the associated level of underpricing and hence the first-day return. They have constructed the underwriters ranking based on the tombstone announcements. Basically, in the tombstone announcements, the participating underwriters are divided into sections based on their reputability and prestigiousness starting from section B, then C, etc. where B is higher than C and so on. They went through all the tombstones of all the IPOs in their sample and kept shifting the underwriters up and down between the ranks according to their relative position in each tombstone announcement. They reported that underwriters underprice more to compensate for the level of information, and hence the risk associated. They showed that for the more prestigious underwriters, the return on the first-day trade after the IPO is less. That is a reflection of the level of risk associated with it.

After establishing that more prestigious underwriters take less risky IPOs compared to their less prestigious counterpart underwriters, (Carter, 1992) examined if underwriters seek to get the less risky IPOs to increase their potential of getting more IPOs. The empirical results reported in this paper support this hypothesis. In addition, he tested if the subsequent offering is positively related to the reputation of the underwriter, if the subsequent offering is negatively related to the gross spread and if an issuing firm is likely to switch firms with less prestigious underwriter rather than higher prestigious underwriters. He reported that all the four hypotheses proved to be significant.

(Carter, Dark, and Singh, 1998) used the same underwriters' ranking method that was used by (Carter and Manaster, 1990). They have examined the effect of a

more prestigious underwriter over the long-run post-IPO performance. They tested a three-year holding period and reported that the underperformance of the IPOs with more prestigious underwriters is less severe than the others. They have tested their ranking method using the tombstone announcement replacement method to other methods such as (Logue, 1973), (Beaty and Ritter, 1986), (Johnson and Miller, 1988) and (Megginson and Weiss, 1991). They reported that (Carter and Manaster, 1990) ranking method showed a better significance compared to other ranking proxies.

Torstilla (2003) inspect a global sample of IPOs and discover confirmation of grouping in IPO gross spreads extends beyond the US however at a lower level of a spread than the 7%. Moreover, the results do not show collusive practices by the bookrunners. Also, an examination of abnormal gross spreads following Hansen (2001) indicates that few groups contain unusual positive surpluses.

Chen and Mohan (2002) examine bookrunner reputation, spread and IPO first-day return. They conjecture that bookrunner spread is a direct cost for IPOs while underpricing is an implicit cost for IPOs. Henceforth, they examine spread and underpricing as simultaneous equation frameworks. Armitage (2000) analyse the direct expense of UK rights and open offer and find that the average is 5.78% and the median is 4.28% for the time of 1985-1996.

Torstilla (2001) examine the European IPOs and observe that IPO gross spread is lower for the European IPOs when contrasted with the US IPOs. In a more recent study, Abrahamson et al. (2011) demonstrate that the spread charged by the bookrunners in the US business sector is approaching on settled at 7%. It is turning into the standard for IPO ascending to $250 million. They have contrasted US spread

with the European IPO sector where they have reported bookrunners charge around 3% less than the US market.

2.3 IPO underpricing

Initial public offerings are described by high information asymmetry. Asymmetric information is one of the determinants of IPO underpricing. One such instrument to mitigate for information asymmetry is underpricing the issue. If the IPO sector is described by both educated and naive investors, then offer costs must be reduced to adjust for the naive investors (Rock, 1986) or to honestly revealing information (Benveniste and Spindt, 1989). Rock believed that a few investors, similar to the issuing firm, or its bookrunners are better informed over other investors. Informed investors offer just for good shares, while naive investors offer indiscriminately for every one of the IPOs. That results in a "winners curse" issue, as naive investors get all the demanded shares in a bad IPO, while informed investors get allotments in good IPOs.

There exist other explanations for IPO underpricing in the literature. For example, prospect theory (Loughran and Ritter, 2002), corruption (Loughran and Ritter, 2003), informational cascades (Welch, 1992), lawsuit avoidance (Hughes and Thakor, 1992), signalling (Allen and Faulhaber, 1988) and IPO as a marketing event (Chemmanur, 1993) has been proposed in the literature as an explanation of IPO underpricing.

To the best of my knowledge, no study has considered whether moral hazard problem can lead to underpricing. The survey evidence in Jenkinson and Ljungqvist, (2001) and Ibbotson, Sindelar and Ritter, (1994) is in line with the idea that moral hazard does not lead to underpricing. Since underpricing is determined promptly after

the issue, it is hard to envision how underpricing can mitigate moral hazard. Nonetheless, Brennan and Franks (1997) contend that underpricing is a component through which managers can ensure their private advantages by allotting shares deliberately when raising money through IPOs.

2.4 Lockup length

Past studies contend that lockup contracts lessen the asymmetric information and alleviate agency issues between the managers who are insiders and the shareholders who are outsiders (Brau, Carter, Christophe and Key, 2004). Ibbotson and Ritter (1995) contend that investors are prepared to pay more for a firm with a lockup contract for two reasons: (i) Insider managers cannot take advantage of the any private information because they need to hold on to the IPO allocations until the lockup expiration day reducing the value of any negative private information, and (ii) if the managers as insiders retain large holding their incentives will be aligned with the outside shareowners objectives (Ibbotson and Ritter, 1995). A substantial number of studies give backing to these contentions as insiders abstain from selling shares before the lockup expiration date inspired by the fear of passing on negative signals to the investors (Brau and Fawcett (2006)).

Brav and Gompers (2003) propose three extra contending hypothesis to clarify the presence and length of the lockup time (i) Lockups signal firms' quality, (ii) lockups are commitment device, and (iii) they are rent-seeking mechanisms by the underwriters. They find that lockups are driven by the commitment theory, however, dismiss the signalling and the rent-seeking theories. In any case, Brau et al. (2005) revisit these outcomes and give backing that lockups are commitment devices and they provide signals about firm quality. They demonstrate that Brav and Gompers (2003)

findings of an inverse relationship amongst transparency and lockup length support that the lockup is a signal as much as lockup is a commitment device. They additionally report that the length of the lockups is positively connected with high information asymmetries, low firm-specific risk and high potential for moral hazard.

The lockup agreements are more homogenous in terms of length and types in the US as compared to the UK. Hoque (2011) analyses heterogeneous lockup agreements from the London Stock Market. As compared to the US studies (Field and Hanka, 2001), Hoque (2011) reports four types of lockup agreements in the UK which are unique: absolute-date lockups, relative- date lockups, single lockups and staggered lockups. In line with the previous papers on lockups using US data (e.g., Brava and Gompers, 2003; Brau and Fawcett, 2005), Hoque (2011) tests several potential explanations for the choice of lockup contracts: (1) lockups mitigate information asymmetry, (2) they are signalling devices, (3) lockups reduce agency problem, and (4) they are certification tools. The study finds strong evidence that lockups mitigate information asymmetry and are certification tools and partial support for agency explanation for the choice of lockups. The study further provides support through insider selling activity around lockup expiration and drop in share price around lockup expiration returns are also consistent with asymmetric information, certification and agency hypothesis.

2.5 Ex-ante measures of risk of IPOs

Ritter (1984) and Beatty and Ritter (1986) assert that underpricing of the IPO should increase as ex-ante uncertainty increases. Since then, various studies have used different proxies of measures of ex-ante uncertainty. Studies use company characteristics, offer characteristics, disclosure, certification and aftermarket variables

to measure the ex-ante risk of IPO. For example, age has been employed by Ritter (1984, 1991) and Megginson and Weiss (1991). The size of the offerings is measured by sales has been employed by Ritter (1984). Offer characteristics like inverse gross proceeds (e.g., Beatty and Ritter, 1986; Prabhala and Puri, 1998), percentage width of filling price range (e.g., Hanley, 1993; Prabhala and Puri, 1998), offer price (Tinic, 1988; Prabhala and Puri, 1998; Brennan and Hughes, 1991; and Beatty and Welch, 1996) has been used in the literature. Disclosure variables like the number of uses of proceeds (Beatty and Ritter, 1986) and the number of risk factors (Beatty and Welch, 1996) has been used as well. Certification measures are utilised by a number of studies. Venture backing is used in Megginson and Weiss (1991), Barry, Muscarella, Peavy and Vetsuypens (1990), Lin and Smith (1998), Gompers and Lerner (1997) and Hamao, Packer and Ritter (2000). The reputation of underwriter has been used in Megginson and Weiss (1991), Carter and Manaster (1990), Habib and Ljungqvist (2001) and Beatty and Welch (1996). There has been a number of after-market variables that has been used as a proxy of ex-ante uncertainty. For example, the standard deviation of daily share price returns by Ritter (1984, 1987), Clarkson and Merkley (1994), Finn and Higham (1988), Prabhala and Puri (1998). Others use daily trade volume in early after-market (e.g., Miller and Reilly, 1987; Prabhala and Puri, 1998).

2.6 Long-run IPO dynamics

2.7 Lockup expiration returns

On the lockup expiry dates share price drops to a significant extent (e.g., Hoque (2011), Hoque (2014), Hoque and Lasfer (2016), Brau et al. (2004), Brav and Gompers (2000, 2003), Bradley, Jordan, Yi and Roten (2001), Ofek and Richardson (2000), Field and Hanka (2001)). A number of explanations have been put forward to explain

the share price drop of a known event. However, the empirical evidence is not anonymous (e.g., Ofek and Richardson (2000)). Brau et al. (2004) report a positive relationship between the insider ownership in the IPO, proxy for agency costs, and the cumulative abnormal returns over five days. Field and Hanka (2001) test a number of hypothesis including bid-ask spread, liquidity impacts and supply shocks, downward sloping demand curves, worse than expected insider selling. Field and Hanka (2001) find that higher trading volume is associated with the higher abnormal return, which lends support to the downward sloping demand curves. They additionally conclude that the insiders selling is related to the larger drop in share prices, yet fail to support that the decrease is exclusively determined by more terrible than anticipated insider selling. However, Hoque and Lasfer (2016) did not find any evidence that insider selling leads to a higher drop in share price around the lockup expiration. These outcomes are not clear, and it is ambiguous whether the abnormal drop in share price around the lockup expiry dates is consistent with the signalling, downward sloping demand curves, agency problems or an artefact of bid-ask bounce.

2.8 Long-run IPO underperformance

Almost 40 years back, Ibbotson (1975) was first to report the underperformance of IPOs. However, the issue of underperformance gets momentum after Ritter's (1991) work. Utilising a sample of 1,526 IPOs in the US over the ten- year period from 1975 to 1984, Ritter (1991) find that in 3 years post-IPO, these organisations show poor performance when compared with matched firms by size and industry. The 3-year buy and hold abnormal return for the IPOs in comparison to the matching firms is -29.13%. Ritter (1991) also notice that the underperformance varies to a significant extent across years and is different across industries. The IPOs that join the market in hot issue market are the worst performers. Ritter (1991) asserted that the results are

consistent with the notion that suboptimal companies might issue equity in the hot markets where investors are sporadically overoptimistic about the young growth firms, suggesting that companies take advantage of these 'windows of opportunity'.

Loughran and Ritter (1995) later increase the sample size and coverage in comparison to Ritter (1991) and examine the long-run performance of 4,753 firms that went public in the US between the twenty-year period from 1970 to 1990. Loughran and Ritter (1995) find that the IPOs over that period underperform in comparison to the non-issuing firms subsequent to five years after IPO. In particular, the non-issuing firms average buy-and-hold-return is 12% per annum whereas the IPO firms buy and hold return is 7%, so they underperform by 5%.

The US evidence of the long-run underperformance is examined in international settings, and it seems that IPOs underperform in other countries as well. For instance, an earlier study by Levis (1993) utilising data from the London Stock Exchange spanning a period of 1980-88. Contingent upon the benchmark utilised he reports a 3-year post-IPO underperformance of - 8.31% to - 22.96%. Aggarwal, Leal and Hernandez (1993) report that IPOs underperform in the long-run by - 47.0% for Brazil, Chile, Mexico. Keloharju (1993) examines 79 Finnish IPOs and find that IPOs underperform by - 21.1% after 3-years post-IPO period. Similarly, Cai and Wei (1997) in the context of Tokyo Stock Exchange provides evidence of that IPO underperform over the long-run by -26.0%. However, they only look a small number of IPOs issued on over the time of 1971-1992.

The long-run underperformance stems from a mixture of heterogeneous expectations among the investors, and short selling is expensive if not impossible for the small issue size on many IPOs (Ritter and Welch (2002) for a review). The reasons

as explained in Hoque (2010): “Miller (1977) and Morris (1996) explore asset pricing by relaxing the standard assumption of homogeneous expectations. By accommodating the divergence in opinions about the future cash flows and growth potential of a company presents an element of reality that can explain long-run underperformance. As the disagreement in belief gets smaller and smaller, the marginal investor’s belief of the IPOs and thus the trading price are lowered. Obviously, these arguments are based on the proposition that the heterogeneity is at peak at floatation but gradually drops through time with the arrival of new information. Consequently, the number of once positive investors alter their beliefs about the value of the company, and drive the price down, even though the average investor’s belief might never have changed. It is worth noting that the new information does not have to be negative. Any piece of information that reduces the variation of opinion about a firm will lead to a lower price” (p..250).

2.9 Survival of IPOs

For the purpose of supporting event studies in the UK, Gregory, Tharyan and Christidis (2013) worked on constructing and testing alternative version of Fama-French and Carhart models. In their paper, they have tested some versions of four-factor models trying to explain the cross-section of returns in isolation of the effect of momentum. They find that their model is able to explain the cross-section return of large firms and portfolios more accurately than for the smaller firms. They have tested the basic Fama-French basic model which is basically the CAPM model where they add to it the size and value factors. I used their 4-factor model where they included a winner and loser factor. For the size and value factors (SMB and HML), they formed six portfolios based on market capitalisation and boot to market value in October of each year. With respect to the winner and loser factor, they create another 6 portfolios based on their sizes and the level of momentum. They have also tested some other versions of the model such as the Value weighted factor components model and the Decomposed factor model with both 3 and 4 factors for each one of them.

Fama and French (2004) report a significant decline in the survival of IPOs for the first ten years after IPO from 61.0% for 1973 to only 37.0% for 1991. Likewise, the chances of seasoned firms surviving have declined from 60.6% down to 46.9% over the same period. Seguin and Smoller (1997) inspect the relationship between IPO issue price and firm survival. Schultz (1993) investigates whether unit IPOs survives higher with respect to non-unit IPOs. Neither one of the studies controls for firms' characteristics like ownership characteristics or accounting-based fundamental measures of firm performance. Demers and Joos (2007) develop an IPO failure risk model by incorporating accounting and other deal specific characteristics. They document statistically different failure rates between tech and non-tech IPOs. Other studies examine the different role of auditors and audit reports in predicting post-IPO firm survival (Willenborg and McKeown, 2000; Weber and Willenborg, 2003).

Hensler et al., (1997) examine the effects of several firm characteristics and IPO characteristics on the survival of IPO firms by using an accelerated failure time (AFT) model. The results show that large IPOs, old firms, firms which underprice more and firms with increased level of IPO activity in the aftermarket, firms with higher level of insider ownership survives longer. They also find that survival time decreases with the number of risk characteristics reported in the IPO prospectus. The IPO survives shorter if the firm is in from certain industries such as computer and data, restaurant, wholesale industry, and airline and survives longer if the IPO is from optical or drug industries.

Espenlaub et al., (2012) examine IPO survival the Alternative Investment Market (AIM), where the nominated adviser plays a pivotal role in bringing the IPOs in the market. They find that Nomad reputation significantly affects IPO survival.

Initial public offerings managed by reputable Nomads survive longer almost by around two years than those sponsored by different Nomads. They find that survival rates of AIM IPOs are at par to those of North American IPOs. Ahmad and Jelic (2014) analyses the impact of lockup agreements on the survival of Initial Public Offerings (IPOs) during the time of 1990–2011by using a sample of UK IPOs. Their accelerated failure time (AFT) survival model demonstrates a significant impact of lockup length on the post-IPO survival. The authors report that the IPO firm survives 27% longer if the median of the lockup length increases by a year. Moreover, the delisting rates for IPOs with shorter lockups are significantly higher than the delisting rates for IPOs with longer lockups regardless of the reasons for delisting.

2.10 IPO Tone measures

The content analysis has been used in the IPOs by a number of authors. For example, Hanley and Hoberg (2010) utilising word content examination break down information in the IPO prospectus into its standard and informative parts. They assert that the higher the informative substance, the higher the premarket due diligence that results in more exact offer price and less underpricing. That is because in the higher due diligence on the part of the underwriters diminishes the issuing company's dependence on book building to value the issue. The inverse is valid for standard substance.

Loughrun and McDonald (2013) use the S-1 forms that are the primary SEC recording in the IPO process. The tone of the S-1, as far as its authoritativeness in describing the company's business methodology and operations, ought to influence investors' capacity to value the IPO. Loughrun and McDonald (2013) find that IPOs with large amounts of dubious content have higher first-day returns, total offer price

corrections, and higher volatility. Their results give evidence to the hypothetical models of uncertainty, book building, and prospect theory.

Loughrun and McDonald (2013) demonstrate that word records produced for different subjects misclassify normal words in financial context. Using a large sample of 10-Ks during 1994 to 2008, they show that 75% of the words recognised as negative by the heavily used Harvard Dictionary are words normally not considered negative in financial settings. They came up an alternative negative word list, alongside five categories of words, that better reflect tone in money related context. They relate the word lists to a number of financial variables—filing returns of 10-K reports, the volume of trading, volatility of share return, misrepresentation, material weaknesses, and unexpected earnings.

I used the dictionaries that are constructed by Loughran and Mcdonald to measure the tone of each prospectus. 8 dictionaries in total were constructed in addition to Harvard Negative words dictionary. The Positive and Negative dictionaries contain words that are perceived as positive (e.g. best, innovation) or negative (e.g. delist, underperform) in the finance field. The Constraining dictionary has words such as (commit, limit). The Litigious dictionary has words from the legal field such as (law, indemnity). The Interesting dictionary has some words that make the text sound more interesting such as (compel, insist). The Uncertainty dictionary contains words that would indicate uncertainty and ambiguity such as (assume, might). The Superfluous dictionary has the unnecessary and replaceable words such as (effectuates, whilst). The Modal dictionary has the words that will change the mode of the main verb that follows such as (almost, possibly). They have divided the Modal dictionary into strong, moderate and weak modal words. Besides those, there is a

dictionary that contains the irregular verbs. I calculate the weight of the usage of each dictionary in each of the prospectuses which give me an indication of the sentiment of the text used in the prospectus.

Textual analysis has been utilised as a part of an increasing number of research papers in Accounting and Finance to quantify the tone and sentiment of corporate news releases, Management discussions and analysis (MD&A), yearly reports (10K reports), daily paper articles, and investor message sheets. Examples incorporate Engelberg (2008), Li (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008), Tetlock (2007) Antweiler and Frank (2004). The outcomes in these studies could be summarised as negative words could be a successful method for measuring tone as negative words demonstrate a significant relationship with financial variables.

Different papers look at the tone of different documents in the context of finance related variables. For example, Kothari, Li, and Short (2008) relate the tone of daily paper articles on the cost of their utilised funds, the variability of return, and forecasts by analysts. Henry (2008), Engelberg (2008), and Demers and Vega (2008) relates news discharges with lower firm income, earnings drift, or stock returns. Some papers relate the information content of IPO prospectus with the share returns, price variability and trading volume (e.g., Li (2008, 2009), Feldman et al. (2008), Hanley and Hoberg (2010)).

Chapter 3

3. Methodology and data

In this chapter I go through the methodology I used in this thesis. I discussed briefly the regression models such as OLS, tobit and fixed effect models. Then I discuss tone measures and the textual analysis and how we calculate the tone of each of the prospectuses in our sample. Then in the second section, I explain the data and the how I collected them and their sources.

3.1 Methodology

This thesis uses various techniques and methodologies to test the hypotheses. In particular, I used OLS regressions, Fixed effects model, event studies and textual analysis. I will discuss the above methodologies in brief in the following section.

3.2 Regressions

Ordinary least square regressions have been used in the theses for testing a number of relationships. Since this is a widely used measure, I will just mention the functional form.

$$Y_i = \beta_i X_i + \varepsilon_i \qquad (3.1)$$

Where,

Y_i is the dependent variable, and X_i is a vector of independent variables and ε_i is the error term. I obtain the cluster adjusted robust standard error (Petersen, 2009) to estimate the t statistics.

I also use fixed effects model to address the issue that underwriters are repeat players in the IPO market. The general functional form of fixed effects model is as follows:

$$Y_{i,j,t} = \{\beta_{i,j,t} X_{i,j,t} + C_t + \varepsilon_{i,j,t} \quad (3.2)$$

The regression includes underwriter fixed effects (cj) to account for unobserved heterogeneity at the underwriter level that may be correlated with the explanatory variables. Standard errors are clustered at the bank level to account for heteroscedasticity and serial correlation of errors (Petersen, 2009).

I also used Tobit regression where the data is truncated in nature. For example, for the gross spread regression, only observe the fees within a certain range. So, Tobit model could be better specified as follows:

$$Y_i = \beta_i X_i + \varepsilon_i$$
$$Y^i = \left\{ \frac{= Y_i^* \; if \; Y_i^* \geq c}{= c \; if \; Y_i^* \leq c} \right\} \quad (3.3)$$

The theses also use Logit regression, where the dependent variables are binary in nature.

The simplest idea would be to let π_i be a linear function of the covariates, say

$$Logit(\pi_i) = \acute{X}_\iota \beta \quad (3.4)$$

where β is a vector of regression coefficients.

Exponentiation the above Equation, I found that the odds for the i-th unit are given by

$$\pi_i = \frac{\exp\{\acute{X}_i \beta\}}{1 + \exp\{\acute{X}_i \beta\}} \quad (3.5)$$

3.3 IPO tone measures

I used the word lists developed by (Loughran and McDonald, 2011) in addition to Harvard IV dictionary for negative words. To analyse each prospectus, I used the AntWordProfiler. For each prospectus, I have gathered the following information:

- Total number of words in the prospectus
- The frequency of appearance of each word in the prospectus
- The number of words appeared in the prospectus that belongs to each of the dictionaries.
- Numerical characters have been excluded from the counting

Loughran and McDonald (2011) have introduced nine different dictionaries in addition to the Harvard IV negative words dictionary. There is some overlap between those dictionaries were some words appear in more than one of them. The dictionaries are as follows (Negative, Positive, Uncertainty, Litigious, Constraining, Superfluous, Interesting, Modal, Irregular Verb in addition to Harvard IV Negative).

The largest dictionary by far is Harvard IV. LM Negative dictionary is almost half the size of Harvard IV. The other LM dictionaries sizes range between few hundred words to less than a hundred. Following the same methodology used by Loughran and McDonald (2011), I have calculated the proportional weight of each dictionary as follows:

$$Proportional\ Weight\ of\ a\ Dictionary = \sum \frac{Word_i\ Frequency}{Total\ Nubner\ of\ Words} \quad (3.6)$$

For all the words from this specific dictionary.

I have also calculated the Time Frequency – Inverse Document Frequency (tf.idf) so that the weight of a certain word is in accordance with its importance across all the prospectuses under our study. The calculation of the weight of each word is as follows:

$$W_{i,j} = \begin{cases} \frac{(1+\log(tf_{i,j}))}{(1+\log(a))} log \frac{N}{df_i} & if\ tf_{i,j} \geq 1 \\ 0 & Otherwise \end{cases} \quad (3.7)$$

Where:

$W_{i,j}$: the weight of the i^{th} word in the j^{th} document

$tf_{i,j}$: the row word count of the i^{th} word in the j^{th} document

a : is the average word count in the document

N: Number of documents

df_i: Number of documents containing at least one occurrence of the i^{th} word

I used the sample regressions I used in Chapter 1 and introduce the dictionary weight as dependent variables

3.4 Event studies

For the market model, I ran an OLS regression using the daily return for each company as the dependent variable and index return as the independent variable as follows:

$$R_s = \alpha + \beta R_i + \varepsilon_i \quad (3.8)$$

Where:

R_s : Daily return on share price

R_i : Daily return on index

α : Intercept

β : Slope

ε_i: Random error

I ran the regression for the period starting from the date of the IPO until the starting date of the event window. After getting the coefficients of the intercept and the slope of each regression, I calculated the expected return during the event window. Then I calculated the daily abnormal return as the excess return over the expected return as follows:

$$AR = R_s - E(R) \tag{3.9}$$

Where:

AR : Abnormal Return

$E(R)$: Expected return using the regression coefficients

Then I calculated the cumulative abnormal return (CAR) for the period as the numerical sum of the daily return during the event study window.

I also examined the same using Fama-French 4-Factor Model.

$$R_s - R_f = \alpha + \beta_1(R_m - R_f) + \beta_2 SMB + \beta_3 HML + \beta_4 UMD + \varepsilon_i \tag{3.10}$$

Where:

R_s : Daily return on share price

R_f : Risk free rate of return

α : Intercept

β : Coefficient

ε_i: Random error

SMB: Small minus Big – The difference in return between small cap companies to large cap companies.

HML: High minus Low – The difference in return between growth companies to value companies.

UMD: Up minus Down – The difference in return between winner companies and loser companies.

First, I ran an OLS regression for each company using the four factors to find the intercept and the coefficients multipliers. Then I did the same as what I did for the market model by calculating the expected return for every day during the event study window. Then I calculate the abnormal return as the difference between the actual return and the expected return.

3.5 Data

I collected all the IPOs that took place on the London Stock Exchange for the purpose of my analysis. In this theses, I am covering the IPO that took place in London Stock Exchange in both of its Main and AIM markets for the period starting from 1999 to 2012. I was following the same criteria used in some of the previous studies in order to come with consistent and comparable results. To gain a comprehensive understanding, in my sample, I have only included the IPOs that are new admissions

and incorporated in the UK. I further excluded another 697 IPOs that are of financial and the investment firms as the behaviour of the firms operating in these sectors differ from the rest of the market.

I have collected the data in this study from a number of sources. First, I identified the IPOs that took place in the covered period from the *London Stock Exchange new issue report*. Once I have the list of IPOs that took place on the London Stock Exchange, I have collected the IPO prospectuses from *Perfect Filings* database. From the prospectuses, I have manually collected the IPO related information. In particular, I collected underwriters' remuneration, gross proceeds, the number of bookrunners, lockup length. Some of the information such as the total amount of money raised in the IPO proceeds, the issuing price and the total number of shares are available both in the prospectus and in LSE website. In the case of discrepancy, I used the data on the LSE website as it is assumed to have been reviewed and entered after the actual IPO took place. From *Datastream*, I collected the incorporation date to calculate the age of the company at the time of the IPO. I also collected some of the post-IPO information such as the size of capital over the three years following the IPO to calculate the companies' growth and the trading prices. The initial list included 1075 IPOs. There were around 103 companies where their information was missing and difficult to obtain so I had to exclude them. The final list included 972 IPOs with total proceeds of £41.7 Billion. Furthermore, four more companies were outliers and it was clear that the provided data was inaccurate. We had to remove them so the final list included 968 IPOs. The following figure provides an idea about the IPO market cycles and total proceeds raised by the sample IPOs.

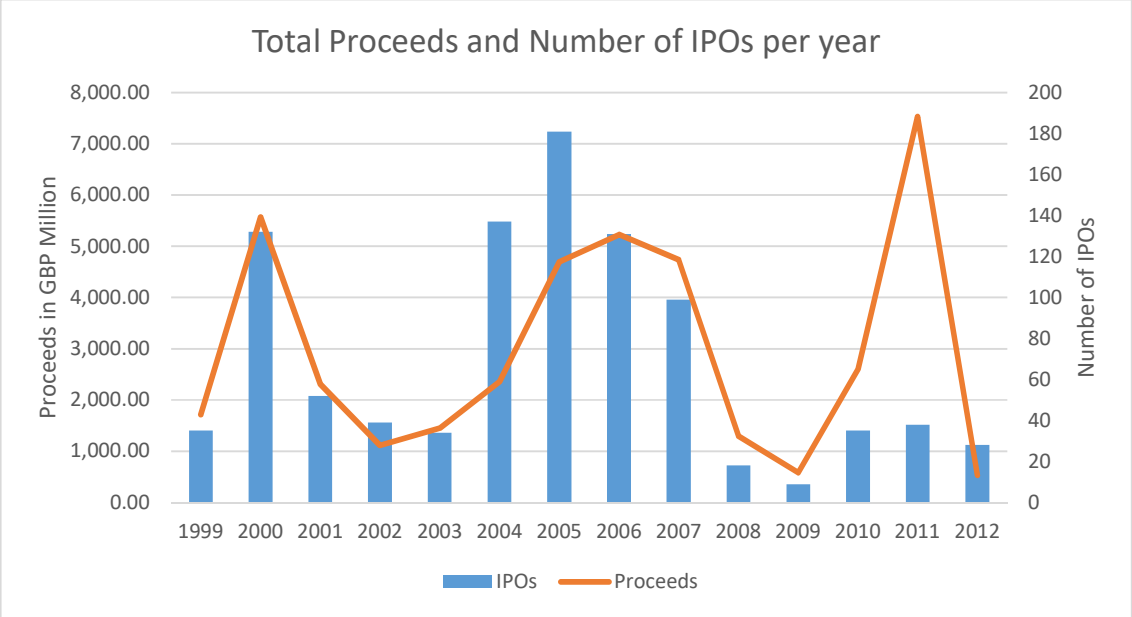

Figure 3.1: Total Proceeds and Number of IPOs per year

Chapter 4

4. Why underwriters charge more in the AIM compared to the Main market?

4.1 Abstract

In this chapter, I examined the factors that contribute to the spread charged by the underwriters for the IPOs that took place in both the Main and AIM markets in the London Stock Exchange over the period from 1999 – 2012. Though there exist few studies on underwriters spread on the Main Markets in different countries there are no studies on the second markets which have grown significantly in the past. I found that the spread charged is in the ranges between 4% - 6.43% with a median of 5%, which is in contrast to the previous US studies which reports a spread of 7%. I found that the economies of scale explain the spread charged in the Alternative Investment Market (AIM) better than the Main Market, implying the spread decreases when the proceed increases. I have tested some other hypotheses such as idiosyncratic risk, volatility, potential growth, underwriters' prestige and rent-seeking. They do not explain spread charged for the Main market while shows more significance for AIM. I did not find evidence that underwriters charge less in the previous year to attract more business in the future. I also found that risky IPOs are underpriced more, meaning that underwriters mitigate risk to a certain extent by underpricing the issue.

Keywords: AIM, Main market, Underwriters spread, IPOs, lockup, underpricing, risk.

JEL code: G15, G24

4.2 Introduction

This is the first empirical chapter. The first section is an abstract of the chapter. The second section is an introduction. Then, the third section is a brief about the institutional background in the UK such as the IPO market, the requirements such as having a nominated advisors for the companies listed in the AIM market and the underwriters fees. The fourth section is a literature review which is followed by the hypotheses of this chapter. On the sixth section, I discussed the date and methodology and then we go through the results and analysis. In this section, the seventh, we go through some descriptive analysis, then the empirical results. Then I discussed use of the spread as a marketing tool. The following part is a robustness check and then I discussed the determinants of underpricing. In the following section I discussed the results then on the ninth section I provide the conclusion of the chapter.

The stock market is a part of the engine that drives the economy of countries. It employs the public equity into public companies to increase the wealth of the investors and hence reflects on the overall economy of the country. In order to have a stable market and to assure the best interest of the investors and the management these markets are highly regulated. IPO has a history that goes back to the early years of the 17th century. Back then in March 1602, when there was a need to raise capital for the Dutch East India Company, the company issued stocks for the partners who decided to join their operations in one single large company. The company has proven that it could benefit from the high capital and was able to get a monopoly over all the trade with Asia (Robins, 2012). Afterwards, with the obvious benefits of the collaboration, the concept of IPO started to prosper and I have seen a number of companies starting to form. London was one of the pioneers in this market where it will be the focus of our study.

The Initial Public Offering (IPO) is a major corporate financing event of a firm's life cycle, and it involves a significant cost. The main purpose of this chapter is

to analyse the costs involved in raising money. There are direct costs and indirect costs of raising money. The direct costs consist of underwriter fees, legal fees, etc. The indirect costs consist of underpricing. Though there are a number of papers which looked into the indirect costs of raising money[2] the number of papers that looked into direct costs is limited.[3] The papers which investigate the direct cost of IPOs have been based on the US market. That is primarily because of the data availability in other markets. The only exception is Torstila (2003) who investigate IPO gross spread from a number of countries.

Chen and Ritter (2000) documented that the IPO gross spread to the issuing firms in the United States is exactly 7% of gross amount of money raised to the investment banking syndicate for IPOs ranging between USD 20 million to USD 80 million. After the paper had been published, this evidence of precisely 7% gross spread triggered an enquiry by the Department of Justice in the United States. Later on, eventually this could not be proved that there was evidence of explicit collusion by the underwriters. A class-action antitrust lawsuit was also started grounded on the evidence shown in Chen and Ritter (2000). However, it was settled privately after a 2007 appeals court decision allowing the claim against the underwriters to proceed. Chen and Ritter (2000) claimed that the clustering of fees at 7% might not provide evidence on explicit collusion, however, asserts that their evidence is in line with implied agreement—or "strategic pricing"—with respect to the underwriters. Papers written later on by Hansen (2001) and Torstila (2003) do not argue that there is implied

[2] See for example Jenkinson and Ljungqvist, (2001), Ibbotson et al., (1994), Brennan and Franks (1997) Loughran and Ritter, 2002; Loughran and Ritter, 2004; Welch, 1992; Hughes and Thakor, 1992; Allen and Faulhaber, 1989; Chemmanur, 1993).
[3] For example, Chen and Ritter (2000), Hansen (2001) and Torstila (2003).

agreement between the underwriters. Hansen (2001) show that the US IPO market is characterised by perfect competition and the entry barrier is low. That could be used as an argument against implicit collusion. Torstila (2003) shows with international data that there is evidence of clustering in gross spreads, however, at a much lower level than 7%. He also argues that clustering does not imply collusive practices.

Abrahamson et al., (2011) conducted a study on European and the US IPO gross spread charged by the underwriters. They find that European IPO fees do not cluster, and only 1% of offerings raising $25 million or more experience gross spread as high as 7%. The European IPO spread averages almost 4%, where the money raised is between $25 million to $100 million range. Their findings show that, certainly, European gross spreads are lower. They find that US IPOs cost 3% more than the European IPOs after controlling for money raised, characteristics of the IPO, multiple bookrunners, and time or country effects. Though the cost is lower for the larger offerings in both regions, their multivariate analysis shows that cost for larger European IPOs got cheaper over their study time while the cost of larger US IPOs has become expensive. They also conduct analysis for the same investment banks who raises money in the US and Europe. Same investment banks charge a significantly higher (lower) fee for IPOs in the US (Europe).

Abrahamson et al. (2011) analyse why same investment bank charges higher fees in the US as compared to the Europe. Armitage (2011) examine the cost of seasoned equity offerings in the United Kingdom and find that mean (median) cost is 5.78% (4.28%). They do not find evidence of economies of scale which is little surprising. Though Abrahamson et al. (2011) examine the large UK IPOs and Armitage (2000) look at the seasoned equity offerings, none of the studies examines

Alternative Investment market (AIM) which has been growing rapidly after launch in 1995 and considered as one of the successful growth markets in the world. The main purpose of this study is to examine in detail the cost of raising money in the United Kingdom Main Market and the AIM and try to understand whether there is any difference between them. If there is any difference in cost of raising equity, why such difference exists.

There exists a significant difference between the IPO markets in the United States and the United Kingdom. Ritter (2003) summarises the differences between the IPO markets of the United States and Europe in terms of price setting, the IPO allocation mechanism, allocations to individual and institutional investors, analyst coverage, etc. There are also differences between the European and US IPO markets in terms of the fees charged by underwriters (Abrahamson et al., 2011). The US lockup length in the US is short (median lockup length is 180 days), and not much information is produced and disclosed during this short lockup period. Also, for the US IPOs, there is a quiet period when analysts affiliated with underwriters are not allowed to issue research reports and recommendations. UK lockup lengths are much longer (median lockup length of 598 days), where a few earnings announcements are normally made during the lockup period. Moreover, unlike the US, there is no quiet period in the UK. As more announcements are made during the longer lockup period, information asymmetry may be less of consideration in the UK market. An empirical question arises, how underwriters charge in such a market where the lockup lengths are longer and more information is produced and disclosed.

The difference is even more striking when I compare the UK Main market and the alternative investment market. The length of lockup in the Main market is

concentrated at 180 days (see, for example, Hoque, 2011), whereas the length varies widely in the Alternative Investment Market. Furthermore, the relationship between underwriters and issuing firms is closer on the Alternative Investment Market than in the Main Market, as the AIM IPO firms need to have a corporate broker to maintain stock exchange listings. Other corporate finance activities, like stake building in target companies, further equity issues and director trading also need to be done through corporate brokers. Underwriters have more power in the UK to set the issue price, as argued by Chambers and Dimson (2009). The UK market serves as a good out of sample test for Abrahamson et al. (2011) and Chen and Ritter (2000), where the IPO market setting is distinct. The different types of institutional settings in the different markets in the UK present an interesting empirical issue, as to whether the fees charged in the two markets in the UK are different.

In this chapter, I first test whether the fees charged in the United Kingdom Main market and Alternative Investment market are the same as it has been shown in some previous studies that show that the US underwriters tend to charge a fixed percentage of 7% of the total proceeds (Chen and Ritter, 2000). In a more recent paper, Abrahamson et al., (2011) compared the observed 7% fees in the US market to the European market, where they reported that the percentage fees charged vary. Although some of the studies argue that a fixed percentage indicate some collusion (Chen and Ritter, 2000). Hansen (2001) argues that the 7% to be an efficient innovation that fits IPO and is not due to collusion. In this chapter, I have collected the data of the IPOs that took place in both the Main and the AIM market of London Stock Exchange. I have found that the underwriters charge ranges between 4% - 6.43% with a median of 5% over the entire sample. It goes in line with the aforementioned study that the European underwriters charge less than their US counterparts.

I also examined the cost of underpricing because underpricing is an indirect cost of raising finance. My results show that for the AIM market, there is a tendency to underprice more. In comparison to the Main Market, this has been noticed more among smaller and new companies. However, this is more noticed in the first trading day as the discrepancy tends to ease where a convergence towards the IPO issue price on the fifth trading day can be seen.

In this chapter, I developed and tested a number of hypotheses that might be affecting the fees charged in the Main market and the alternative investment market. Particularly, I tested economies of scale, riskiness, rent-seeking by the underwriters, attracting future business. Within the coming sections of this chapter, I am going to discuss five hypotheses and will examine them separately throughout the chapter. The base of the hypotheses is the assumption that underwriters charge more in the AIM market compared to the Main market. Then I examine the reason in terms of the economies of scale. Thereafter, I look at the contribution of the risks and whether underwriters use underpricing as a tool to mitigate risks. Then I test the rent-seeking effect. Finally, I examined whether underwriters use lower charges to attract future business.

To test our hypotheses, I ran a number of OLS regression and fixed effect models taking the spread as the dependent variable. I have calculated the spread as the percentage of the difference between the total IPO proceeds and that received by the company to the total proceeds. I identified and tested a number of factors and measured their effect using our regression model. Results show that the factor that has the highest effect on the spread is the size of the proceed received. It has a negative relationship showing that as the total proceeds increases, the spread decreases. That has been

documented by Chen and Ritter (2000) and Abrahamson et al. (2011) where they have talked about the difference between the US market and European market. Although that there is a clustering in the US market around 7%, Corwin and Schultz (2005) has noticed a tendency to show the effect of the dis-economy of scale in the US market. The other factor that has shown to have a significant effect are the year dummies. That shows that the year which the IPO took place contributes to the total spread charged. That can be related to the financial situation and their aftermath as I cover the area where the dot-com bubble took place and financial crises of 2008. The reputation of the underwriter and the rent-seeking identified by the length of the lockup period contribute slightly to the spread charged. I have used the lockup period as the proxy to the presence of rent-seeking. Brav and Gompers (2003) have studied the possibility of using the lockup to gain access to future further issuing. They did not find enough evidence to support this hypothesis.

I have tested some other factors that shown to have a very little or no effect on the spread. Carter (1992) has mentioned that underwriters look for low-risk IPO as this will lead to future business. Higher anticipated risk results in higher spread according to Chen and Mohan (2002). Idiosyncratic risk is a factor that is getting more prevalent as the age of companies pursuing IPO in recent years as per (Fink, Fink, Grullon, and Weston, 2005). To identify any hidden risk that could have caused the spread to be widened, I have calculated the idiosyncratic risk for every share over the first year from the IPO. I have also examined the volatility of the share price during the first year. The results show that both of them are insignificant in the regressions. I have also tested the age of the companies prior to the IPO. That is also insignificant in the regressions. I have further examined potential growth over the 2 – 3 years after the

IPO in terms of the growth of the market capitalization. This factor also found to be insignificant in the regressions.

I have run the regression over different sub-groups of the sample based on the size of the proceeds and the market. Results show that smaller IPOs tend to follow the model more than larger IPOs. That is mainly the IPOs of £16.7M of proceeds or less which makes about 71% of the sample. Most of these IPOs come from the AIM market. That shows that the factor of size proceeds loses its significance as the size gets larger.

I have done a range of robustness checks. Since there are a number of investment banks operating in the IPO market, I control for the investment bank level heterogeneity by using underwriter fixed effects in the regression. My results are qualitatively the same when I used fixed effects models. Secondly, I winsorize the spread at 5% level to minimise the effect of outliers on the fees charged. Again the results are similar to our OLS results. Third, I used Tobit regression because our dependent variable gross spread is truncated at a certain percentage. The results from Tobit regression are similar to that of OLS. Finally, I specify an alternative non-linear model. I find that the non-linear term is significant. In summary, our basic results survive the robustness checks.

Price is one of the main factors in supply and demand. As a result, the spread charged by the underwriter has an effect on future business. As per James (1992), underwriters charge less if they expect further issues by the firm. However, when studying the market share, Dunbar (2000) noticed that underpricing is the main factor that affects market share rather than spread. Krigman, Shaw, and Womack (2001) have noticed that price is the main factor that causes firms to switch underwriters for their

further issues. In this chapter, I have examined whether the price charged in the previous years' lead more business for the firms. I have found some evidence consistent with that notion.

The rest of the chapter is organised as follows. Section two discusses the institutional background in the United Kingdom. Section three discusses the literature and develops the hypotheses. In section four I explain the methodology and the data sources. Section five presents the empirical results. Finally, section 6 concludes.

4.3 Institutional Background in the United Kingdom

4.4 The IPO Market in the UK

London has been a centre of the global finance for centuries. Financial markets can be traced back the late seventeenth century in Jonathan's Coffee-house where they used to list the prices of commodities. It has been developing in terms of practice and laws until it reached a level of maturity at 1801 when a regulated stock exchange emerged. The regulations have been developing and reviewed since then. The stock exchange witnessed one of its first bubbles during the "Railway Mania" in 1845. The Bank of England took action and raised interest rates, which in the result, increased the cost of capital and hence slowed down the momentum, and eventually burst the bubble. As the number of traders increased in number, there was a need for a more spacious building to accommodate them. In 1854, a new stock exchange building was constructed. With the amount of trading and the exchange of stocks, a new Deed of Settlement came into force in 1876. The business was regularly going till it was disrupted by the Great War back in 1914. The stock exchange market was closed for almost six months. However, in the Second World War, the stock exchange market was only closed twice for much shorter periods. The first time was for six days in 1939

while in the second time, in 1945, it was closed only for one day. In 1986, another major event took place when the market was deregulated. That changed the whole structure of the market and resulted in more companies joining the market and a number of companies to go through mergers and acquisitions ("London Stock Exchange History", 2013).

In comparison to other major exchange markets, the LSE comes fourth in terms of market capital size. It comes after New York Stock Exchange (NYSE), National Association of Securities Dealers Automated Quotations (NASDAQ), and the Tokyo Stock Exchange with a market capitalisation of 3,396 billion USD (World Federation of Exchanges, 2012).

The highly regulated market made it difficult for smaller and less mature companies to benefit from the available public equity and other benefits from the companies' shares float. A secondary market was introduced with a more flexible regulatory system. In 1995, the Alternative Investment Market (AIM) market was launched. The main purpose was to create a market for small and growing companies across the world. The admission criteria required to join the AIM market are less strict than those that are required to join the Main market. A summary of the difference between these criteria can be found in the following table as described by London Stock Exchange.

In 1997, LSE started the Stock Exchange Electronic Trading Service. That in return attracted more investors and opened the door to larger sectors of the public to join the market. In 2007, a merger between London Stock Exchange and Borsa Italiana took place creating the London Stock Exchange Group ("London Stock Exchange History", 2013).

Currently, London Stock Exchange provide a different number of services that ranges from the primary markets of both the Main and AIM market where the Initial public offers (IPO) take place to the secondary markets where stocks are traded. They also provide data services such as real data on stock prices in addition to a number of reports and research papers they issue on a regular basis.

Table 4.1: The differences between AIM and MM in terms of admission requirements and continuing obligations

	AIM	**MM**
1. Admission requirements	No requirement for minimum percentage of float No age requirement No minimum market capitalisation Admission documents not pre-vetted by Exchange or UKLA Flat rate admission fee: £4,535 till 2008, and from 2009 fees are charged based on size, min 6,720 and max 75,810.	Minimum 25% of the shares need to be floated Normally 3-years of published account required Minimum market capitalisation of £700,000 Pre-vetting of admission documents by the UKLA Admission fees based on size: min £6,708 and max £388,173.
2. Continuing obligations		
(i) Further issuance costs	No issuance costs.	Similar sliding scale fees like initial issuance, companies get 10% discount compared to IPO cost
(ii)Nominated advisers	Nominated adviser required at all times	No nominated advisers required, but sponsors needed for certain transactions
(iii) Annual fees	Flat rate annual fee: £5,350 plus NOMAD fees.	Sliding scale annual fees: e.g., £4,410, £10,063, £43,470 respectively for up to £50m, up to £500m, >£500m market cap stocks
(iv)Corporate transactions	Shareholder approval is required if the transaction value is higher than the value of the company, simpler documentation required.	Shareholder approval necessary for transactions of much lower value, complex documentation required
(v) Related party transactions	Shareholder approval for related party transactions not required – an announcement to the market that the transaction is fair and reasonable is sufficient	Shareholder approval required for related party transactions
(vi) Corporate Governance	No prescriptive corporate governance Requirements. Combined Code does not formally apply, but companies encouraged to comply	Firms have to comply with or explain non-compliance with the Combined Code and comply with other relevant Listing Rules
(vii) Disclosure requirements	Less prescriptive requirements on nature of financial information to be disclosed	Firms have to comply with more stringent disclosure requirements set out in Listing, Disclosure and Transparency Rules

Source: (Dukas and Hoque, 2016)

4.5 Nominated Advisors (NOMADS)

Having a nominated advisor (NOMAD) is a requirement for any company seeking to be listed on the Alternative Investment market. As per the LSE, A NOMAD is the primary regulator of an AIM company. A NOMAD should meet strict eligibility requirements set by LSE. Part of the listing process in AIM market, a company should assign a NOMAD who will support them to meet the listing requirements and grant them the final approval. NOMADS can also be appointed as the underwriter of the IPO. NOMADs are primarily smaller investment banks or corporate finance advisories rather than larger investment banks.

NOMADs play different roles ranging from being advisors, regulators, middlemen and supervisors. The different roles of NOMADs and the fact that they charge fees for their services could create a conflict of interest. As advisors, NOMADs provide guidance on the rules of the market and thus reduce the cost of the regulating firms. Also, the supervisory role makes the NOMADs liable for any wrongdoing undertaken by the companies they regulate. Moreover, NOMADs are subject to losing their reputation if they fail to prevent their companies from committing such violations.

A company seeking to be listed on the Main Market is required to meet a set of requirements related to the free float rules, minimum market capitalization and to show an established trading record. In other words, for a company to be eligible to be listed on the Main Market should have a minimum of 25 percent of the free float, a market capitalization of more than £700,000 and to show trading record for at least three years.

Despite the differences in regulations between AIM and Main Market, both markets share the same trading mechanism. Shareholders rights are protected under the UK legal system that applies to both markets. In additions, UK Listing Authority (UKLA) requires all companies to publish their accounting information no matter if they fall under UKLA regulations. In contrast to the US market, this prevents the companies to "go-dark".

According to Jenkinson and Ramadorai (2013), the main distinction between the two markets is that the Main Market companies have to abide by a higher degree of compliance, and have obligations to show higher disclosure and transparency in comparison to AIM. However, there are differences in taxation between AIM and Main Market.

4.6 Underwriters fees

With the number of regulations and the level of requirements, entry to LSE is not a straightforward task. Even for the AIM, where the level of requirements is less, companies need to meet certain requirements and comply with a number of requirements and standards to be listed. The investment bankers help companies to prepare for listing on the market. The underwriters' engagement starts when a company decides to peruse the option to go public. The underwriter works with the company for a period that can extend over a number of years until it reaches a level where it will comply with the regulatory requirements. In addition, they will work with them to cover other aspects in terms of company structure, marketing, exposure, etc. Then the underwriter will prepare the prospectus for the IPO and begin the process of raising the capital from the public. The above-mentioned tasks by the investment banks fall under the direct cost to the company seeking the IPO. Nevertheless, another

cost that affects the company is the cost of underpricing. On the trading day, if the closing price is higher than the issue price, and then this would indicate that the issue price set by the underwriter was low, which is known as underpricing. Smaller and newer companies suffer more from this cost as it is more difficult to evaluate their companies. Yet, they could tolerate that since it will attract investors into their venture.

The underwriters' fees charged to companies perusing IPO varies in amount, percentage and type of payment. Most of the underwriters charge a retainer at the beginning of the engagement and then, a yearly fee until the date of the IPO. In our research, I have calculated this as part of the fees charged by the underwriter. IPOs with higher proceeds are expected to attract higher attention and hence higher total cost and more expensive in terms of the absolute value of money. However, as a percentage of the total proceeds generated by the IPO, the percentage tends to drop as the proceeds increases. Larger IPOs usually are undertaken by a syndicate managing the IPO as multi-bookrunners. The bookrunners in such a case can share the fees in terms of percentage or sometimes, they divide the market between them. Some of the underwriters are compensated by allocating them a number of shares exercisable after a lockup period from the IPO date. I have also included this type of compensation into the total fees charged by the underwriter calculating it using the issue price. However, some underwriters issue warrants and rights to issue shares at the predefined price after a certain period from the IPO date. This type of compensation was not included in the total fees calculated in this chapter. This is worthy of mentioning that the information on the fees charged by the underwriter have been collected manually from the IPO prospectuses as it is not legally binding for the underwriters to mention the exact fees charged as a total amount or as a percentage of the proceeds (Abrahamson et al., 2011).

4.7 Literature Review

The literature on IPOs is very rich. Most of the papers try to understand the IPO underpricing and long-run underperformance. A number of puzzles still remain in the IPO literature: underpricing, long-run underperformance and use of lockups. However, less attention has been given to the cost of raising money. For the literature that covers the underwriters' fees, a good number of them are focusing on the issue of the clustering of the spread fees especially in the US market around 7%. Some studies compare the same findings from the US markets to European markets (e.g., Abrahamson et al., 2011). In this section, I will review some of the previous literature and develop a number of hypotheses to be tested in this chapter. The hypotheses that I am going to develop and then test will focus on the difference in the spread charged by different markets in the United Kingdom. I will focus on the spread which is the difference between the total proceeds and the amount that is received by the IPO company at the end of the IPO. The spread is calculated as a percentage of the total proceeds.

The cost of the IPO has direct and indirect parts. The direct cost is the fees that are charged by the underwriter and is usually consists of some fixed costs such as retainers, legal, printing, etc. in addition to a percentage of the proceed. I refer to the fees charged by the underwriter as the spread. The indirect cost occurs when the issue price happens to be underpriced. In this case, the company loses the difference between the issue price and the fair value of the share. I will examine if anticipated risk plays a part in underpricing.

The underwriters' fee is referred to as the spread. The fees components varies between fixed fees and some rights granted to the underwriters. For the sake of this

study, we calculate the spread as a percentage of the total proceed raised in the IPO. One of the papers that bring fee charged by the underwriters in the limelight is the paper the "The 7 Percent Solution" by Chen and Ritter (2000). In this paper, they discuss the factors that cause the underwriter's fees in the US market to be significantly higher than those of other countries. They also argue that for more than 90% of the deals with proceeds between $20 – 80 million have a spread of exactly 7%. This fact contradicts the dis-economies of scale where the increase of the amount of the proceeds should result in a decrease in the spread charged. In our sample, I will examine if the dis-economies of scale is a factor that contributes to the spread charged by the underwriters. This is also noticed in Abrahamson et al. (2011). The spread charged by the underwriters in the European market was decreasing as the proceeds increases. This is too observed in the US market even with the presence of the clustering around the 7% spread. Corwin and Schultz (2005) have noticed in their sample that the spread is negatively related to offering size. In addition, they have also reported that it is positively related to the aftermarket standard deviation. They defined the aftermarket standard deviation as the estimate of the standard deviation of a continuously compounded daily return from the 21st day after the IPO until the 125th day.

Hansen (2001) investigates the implications of collusion in the IPO market as advocated by Chen and Ritter (2000). Hansen (2001) examines this in detail in three phases. In the first phase, he presents evidence from few tests that focus on market structure, whether there is implicit agreement, and whether the spread contains abnormal profit. He checks the nature of IPO market—monopoly or perfect competition. If there is a high concentration in the IPO market, then it may be worthy of investigation because the idea of implicit collusion is present in the highly

concentrated market. This was done in line with Dutta and Madhavan (1997) who show that this is not the case in their implicit collusion model because the dealer market they analyse is a competitive market as dealers are the same size with the same market share. However, the IPO investment bank market structure is different in the United Kingdom as well as in the rest of the world. For example, in 2012 Credit Suisse has a market capitalisation of 40.2 Billion GBP, Collin Stewart has a market cap of 226.3 Million GBP. Hansen (2001) second test investigates entry into the IPO market. Perfect competition is characterised with low entry barriers. If there are low entry barriers, it is hard to imagine any sort of collusion. Hansen (2001) finds evidence consistent with low barriers to entry and low and unchanging concentration in the US market over the 7% era that is against collusion. In the third test, Hansen (2001) apply, in the context of IPOs, the Christie et al. (1994) experiment for NASDAQ dealer collusion. Use of 7% contracts does not vanish after the collusion allegation review is publicised, though there was a federal investigation launched. The fourth test examines whether 7% contains abnormal profit. He develops a benchmark model and estimates the coefficients based on the non-7% spread that is paid in the US market. Using those estimated coefficients, he finds that the 7% spreads do not contain abnormal profit. In sum, his findings do not suggest any form of collusion, either implicit or explicit. On the other hand, Chen and Ritter (2000) simply argue that 7% contains abnormal profit as it is well above the spreads paid in foreign IPOs. However, their claim does not take into considerations the institutional and legal differences in IPO markets between the US and international markets such as the IPO procedure, the IPO contract, underwriter quality. Recently, Abrahamson et. al., (2011) analyse why the same underwriters charge less for the European IPOs compared to the US IPOs.

Abrahamson et. al., (2011) make comparisons in IPO gross spreads between the US and Europe. Since the IPO markets are fundamentally different in these two regions previous studies, previous studies have not been able to make comparisons (e.g. Chen and Ritter, 2000). Historically, the process of raising equity finance in the United States was unique. While the process of the US IPOs for a long time have been managed using the technique called book building, whereby investment banks campaign and gather officially nonbinding but serious indications of interest from institutional investors in the pre-IPO stage, IPOs in Europe used the straightforward fixed-price method or Dutch-auction process which are less time demanding and cheaper in terms of direct costs.[4] Industry participants cite institutional difference is the reason why gross spreads were higher in the United States than in Europe. More recently, nonetheless, the method of raising money has changed. Ljungqvist, Jenkinson, and Wilhelm (2003) indicate that at the end their sample period (July 1999) where Europe is two-third of their sample, approximately 80% of all non-US offerings used the book building method. This trend of book building in the Europe has continued by and large. In the past decade, the majority of European IPOs have used book building, except for a very few small IPOs. Since the European IPO techniques have converged with those in the US, Abrahamson, et al. (2012) claim that IPO Spread across US and Europe are now comparable.

Since the European and US IPOs use similar methods to raise IPOs in the recent times, in order to make a meaningful comparison between the US and European IPOs, Abrahamson, et al. (2012) use 10 years of data from 1998 to 2007 of European IPOs which use book building method. The only study that uses data across the Europe

[4] see, for example, Ritter (2003) and Jenkinson and Ljungqvist (2001).

is Torstila (2001), but this study covers the 1986-1999 when a number of different offering techniques are used to make any meaningful comparison hard between the US and European IPOs. Abrahamson et al. (2012) draw a data on European IPOs which matches a data set of US IPOs spanning the same time frame of 10 years. Moreover, as the world has become a global village leading investment banks compete for IPO business on both the continents—the US and European markets. Hence, Abrahamson et al. (2012) are capable of comparing IPO gross spreads across the US and Europe. In particular, they examine how and to what extent a certain investment bank charges for its practically same services at US and Europe.

Abrahamson et al. (2012) find that the 7% gross spread had become even more pervasive in the United States than when Chen and Ritter (2000) reported. In their sample period (1998-2007), 95.4% of US IPOs in the range of $25 million to $100 million had gross spreads of exactly 7%. The comparable number as reported in Chen and Ritter (2000) after adjusted for the inflation was 84%. Abrahamson et al. (2011) report that gross spreads of exactly 7% became more common among larger US IPOs during their sample period than in Chen and Ritter (2000) sample period. While Abrahamson et. al. (2011) find that 77% of all their sample offerings ranging $100 million-$250 million charge exactly 7%, Chen and Ritter (2000) showed nearly no IPOs over $150 million charges a 7% gross fee. Third, Abrahamson et. al., (2011) show that there is no clustering for European IPO gross spread, and only a few IPOs (1 percent) raising $25 million or higher get charged gross spreads of a maximum of 7%. The average fee for European IPOs is little higher than 4% for the IPOs raising $25 million to $100 million. Certainly, they find that there is a difference of 3 percent between the gross fees of European IPOs and the US IPOs. European IPOs are cheaper as compared to the US ones after controlling for IPO characteristics, the size of IPO,

multiple or single bookrunner, and time or country effects. They also show that gross spreads are lower for the larger offerings in both Europe and the US, that is, economies of scale are in place in the underwriting market. However, their multivariate analysis shows that there is a tendency to increase the fees for the larger US IPOs over their sample period, while European IPOs have shown a trend of decline in their fees. Finally, they show that the same underwriter charges less for an IPO of an in the Europe as compared to the United States for an IPO with similar characteristics.

One might argue that institutional and country differences between the two markets explain the 3% difference between the gross spread of US and European IPOs. Abrahamson et al. (2012) examine five of such claims: legal costs, retail offerings, litigation risk, sell-side analysts, and the possibility that lower underpricing might offset higher fees. They do not find any evidence that these factors explain the lower level of gross spread in the Europe as compared to the United States. The only evidence that supports their claim is that they find that European IPOs have a lower level of underpricing compared to the US IPOs.

Chen and Mohan (2002) assert that underwriters have basically two ways to mitigate the risk of the issuing firm. The first one is explicit which underwriter spread is and the second one is implicit which is underpricing the issue. If an underwriter may well charge a spread which is adequate to reward for all the risk they bear, then underpricing turn out to be less dominant because mispricing may harm in underwriter reputation and/or economic loss. The finding that underwriter reputation is negatively correlated with underpricing is consistent with the safeguarding of investment banks reputation capital (Carter and Manster, 1990). Nonetheless, there are two major obstacles against charging a high spread as compensation for the risk taken by the

underwriters: competition and regulation (e.g., Chen and Mohan, 2002). As a result, underpricing is widely observed (Chen and Ritter, 2000). Competition and regulation, nonetheless, still leave room for underwriters to shift the cost from spread to underpricing and vice versa (Chen and Mohan, 2002).[5] Thus, underpricing and spread works as substitutes for underwriters. However, in an alternative scenario, for a risky IPO, it is possible that due to regulations if a ceiling-level spread needs to be adhered it could be compensated by a higher level of underpricing to reward for underwriting a very risky IPO. Thus, Chen and Mohan (2002) conjecture that the underpricing and spread are complementary mechanisms and the nature of the relationship between these two be determined by the IPO features and negotiating power, the level of competition in the investment banking market and the underwriter's pricing strategies.

Since Chen and Mohan (2002) hypothesise that underpricing and spread could be substitutes or complementary they examine both of them. According to Chen and Mohan (2002) underwriter spread does not only reflects the investment banks risk taking in terms of bringing the new company to the market, but it also related with the underpricing of the issue. As underwriter spread and underpricing are interrelated, and both could be determined as a result of the decisions taken by the underwriter, analysing one of those in isolation without the other is incorrect. Thus Chen and Mohan (2002) examined the spread in a simultaneous equation modelling system. They estimate the spread in terms of the gross proceeds, other expenses, offering price

[5] Chen and Mohan (2002) provide an example: "Imagine a scenario where competition drives the underwriter spread to x%. To bear the full risk of underwriting, the underwriter will, say, underprice the issue by y% to achieve an equilibrium risk premium of (x+y)%. Or, in the case of an influential issuer, the underwriter can only underprice the issue by y%. To make up the loss of risk premium, the underwriter will negotiate an equilibrium spread of x%. In a different scenario, the lack of intense competition allows the underwriter to demand a (x+q)% spread, and a (y-q)% underpricing would be sufficient to bear the full cost of risk" (page 523-524)

and underpricing as the first equation. The second equation was the estimation of the underpricing in terms of standard deviation, underwriter reputation and the spread. They have argued in their system discussion that other expenses and fees are positively related with spread charged. The other expenses and fees in their system refer to the legal and bookkeeping expenses. They argue that the positive relation with the spread indicate that riskier issues are associated with higher spread and higher other expenses. They also examine whether the market for investment banking is segmented or not. The level of underpricing depends on the investment banking market structure, that is, whether the market is integrated or is segmented. In a fragmented market, some of the new issuers could not raise money with some underwriters. Consequently, the underpricing of IPOs reveal the pricing bias that results from the fact that some smaller companies are excluded from the service of large and reputable investment banks.

An important factor that plays a significant role for the investment bankers in the process of bringing a company to the market is the risk associated with the company. Fink et al. (2005) have discussed how the idiosyncratic risk has risen since the 1960s. They have discussed in their paper how the age of the companies going public has fallen drastically as the average age was 40 years old back in the 1960s to an average of 5 years old in the late 1990s. This leads younger companies to have access to cheaper equity. Moreover, they argue in their paper that the positive trend of idiosyncratic risk is fully explained by the proportion of young firms in the market.

Carter (1992) discussed the reason why reputable underwriters seek low-risk IPOs. They argue that bringing a low-risk IPO to the market will make them viable for future business. They have reported in their paper that the likelihood for an underwriter to issue further equity is positively related to its reputation and negatively

related to the IPO gross spread. However, still, the US market show a fixed gross spread for most of the IPOs. Chen and Ritter (2000) show that the 7% spread is consistent with their sample of 1,111 IPOs. As shown by subsequent research (Abrahamson et al., 2011) that 7% is relatively higher compared to their counterpart in the European markets, they have tested the hypotheses that this high spread is due to the level of riskiness of the firm seeking IPO. They did not find any evidence to support this hypothesis. They explained that the 7% persists for risky deals as well as for the relatively easy to value firms. They have argued that the clustering around the 7% is more apparent than it was a decade before the publishing date of Chen and Ritter (2000). They also noticed that the 7% spread is not evident at the IPOs outside the US market. Moreover, there is no significant clustering for other financial markets inside the US such as a bond, convertible bond, and seasoned equity offering.

Another aspect that I will be examining is the possibility of rent-seeking is that underwriters could be using the lockup to benefit from further issuing after the IPO. Brav and Gompers (2003) have examined, amongst some others, the roles of the lockup on IPOs, whether the underwriters extract extra compensation after the IPO use the lockup. They have examined the relationship between the length of the lockup and the reputation of the underwriters. They found that the relationship is negative. The more reputable underwriters, the shorter lockup period they impose. This suggests that reputable underwriters do not use longer lockup periods to get additional compensation.

I will examine lockup period from the company side. Lockup period is used as a tool to overcome some issues such as information asymmetry and moral hazard issue. As a result, the longer the lockup period, the more likelihood of one or both

present. The paper by Karpoff, Lee, and Masulis (2013) examined the lockup agreements in Seasoned Equity Offerings (SEOs). They measured the duration of the lockup period to the issuer information asymmetry measures. They found that likelihood of lockup and its duration are positively related to the information asymmetry measures. In a sample of 1,926 IPOs over the period of 1996-2006, they reported that 81.5 percent of the sample has a lockup period of 180 days. However, with respect to the SEOs, in their sample, they have 2,579 SEOs over the same period with 64.4 percent with 90 days' lockup. Besides the previous studies that attribute the lockup duration to control moral hazard, they argued that lockup reduces the likelihood of overpricing. They used eight information asymmetry proxies. These proxies are firm size, time since IPO, the number of analysts providing firm earnings forecasts, tangible assets, and the number of prior stock offers since IPO, bid-ask spread, return volatility and abnormal accruals. Part of their study was to examine the relation between the information asymmetry with the spread. They found that there is a positive relation between them.

Hoque (2014) conducted further investigation on the relation between the lockup duration and information asymmetry. In his paper, they covered 1,117 IPOs over the period between 1999-2006 in both markets of London Stock Exchange, the Main Market and the Alternative Investment Market (AIM). In his study, he identified the companies with information asymmetry when the company is small (less than £11.92 Million) when underwritten by a non-prestigious underwriter and those that are issued in the AIM market. He has also identified the companies with high moral hazard. He has identified them using director ownership. If director ownership is higher than 0.75%, the company is identified to have moral hazard. He reported that information asymmetry drives underpricing. However, for the subset identified as a

high moral hazard, the standard deviation is reported not to be significant. This means that information asymmetry is not the driver of underpricing when a high moral hazard exists. In their analysis, they found that information asymmetry is not the driver for lockup duration. However, director ownership is significant for the whole sample and for the high moral hazard subset. Hence, they concluded that lockup duration is used to mitigate moral hazard. Therefore, information asymmetry drives underpricing and moral hazard drives lockup duration. Rent-seeking can be a factor in the spread charged as they could indicate the level of information asymmetry and moral hazard. I will examine this by testing the relationship between the lockups periods stated after the IPO to the spread charged. To the best of our knowledge, no previous studies have examined the relationship between rent-seeking and spread charged by the investment banks. I try to fill this gap in the literature.

The last explanation I am seeking to explain in the spread in this chapter is the possibility that underwriters charge less to attract future business. James (1992) examine the pricing of the underwriters' services. They reported in their paper that the underwriters charge less spread if they expect subsequent issuing from the firm. They argue that with a lower spread, the likelihood of the firms to switch to another underwriter for future issuing is less appealing given the switching costs. In other words, the spread is being used as a marketing tool in case if following issuing is expected. In a study by Dunbar (2000), they investigated the factors that affect the investment banks market share in the IPO market. Their study covers IPOs issued in the US market over the period between 1984 and 1995. They investigated the effect of a number of factors on the market share changes for established and less established investment banks in the IPO market. They found in their study that those factors are significant in gaining market share. They mentioned that if shares are underpriced at

the IPO, this will have a negative effect on their market share following the IPO. They have also reported that they tend to have a better market share if they diversify their portfolio in terms of sectors and IPO size rather than industry specialisation. Although they have reported that those factors are significant for the whole sample, they mentioned that these factors are more significant for the well-established bank in comparison to the less established ones. Industry specialisation factor was, economically, the most important factor. It was followed by the analyst reputation and initial IPO return.

Another study by Krigman et al. (2001) investigated the main reason for the issuers to switch their IPO underwriter for their SEOs. In their study, they covered IPOs that took place in the mid-1990s. They argued that 30% of issuers undertaking SEOs within the first 3 years of the IPO switch their lead underwriters. They found that switchers were significantly underpriced compared to non-switcher. Therefore, their dissatisfaction is not attributed to the services offered by the underwriters; rather it is due to the money left on the table and hence lower proceeds. The second main reason was that companies look for higher reputation underwriters that will yield in more exposure in terms of more influential analyst coverage in comparison to the less reputation underwriter.

In our study, I will examine the relationship between the spread charged and the future business by examining the relation between the current year spread with the number of IPOs undertook by the underwriter the following year. However, I have excluded some of the market sectors such as the financial sector. Moreover, I have not included any of the further equity issuances. Yet, this could be compensated by the assumption that the underwriters have some speciality as noted by (Dunbar, 2000).

Moreover, the paper by Christie and Schultz (1994) mentioned that competition could be a contributor to the spread charged, it could be a marketing tool used by the underwriter to attract future business.

4.8 Hypothesis

In this chapter, I am studying and comparing the main market which is a well-established and mature stock exchange market with a high level of regulation compared to the AIM which is an emerging market that attracts smaller, less mature and hence riskier companies. In addition, the AIM market has fewer restrictions for the companies to be listed on them compared to the Main market. The company characteristics are different in the Main market compared to the alternative Investment Market. The main market attracts mature companies with higher age, lower underpricing, higher profitability, higher asset tangibility, lower growth prospects, low ownership concentration, high dividend payments (Vismara, Paleari and Ritter, 2012). Also, Main Market companies make acquisitions at a great pace and do more capital changes.[6] On the other hand, AIM attracts companies with lower age, higher underpricing, lower profitability, lower asset tangibility, higher growth prospects, high ownership concentration, low dividend payments. However, AIM companies issue

[6] The following is a list of non-chargeable transactions for companies admitted to the MM: Capital reorganisation, Sub division of capital, Consolidation of capital, Redenomination, Capitalisation of reserves, The reclassification of shares in order to liquidate a company under a scheme for reconstruction, Establishment and updating of issuance programmes, Block listings for issues of shares under employee share schemes and exercise of options (including issues of shares to directors not under an employee share scheme) with a market capitalisation below £2m, Further issues of shares issued under an existing offer for subscription, Substitution of issuer, Migration between 'securities categories – equity shares' (London Stock Exchange, 2011).

subsequent equity (SEOs) frequently as the companies do not need to pay fees to the stock exchange for SEOs. Initial fee considerations can be vital to the decision to issue equity on Alternative Investment Market. For the London Main Market, incremental admission fees are charged on equity issues. "Admission fees of minimum £6,708 to a maximum of £388,173 are charged based on size in the MM. In comparison, the flat rate admission fee of £4,535 was charged till 2008 in the AIM. From 2009, minimum £6,720 and maximum £ 75,810 fees are charged based on size. The admission and on-going fees can be an important consideration to issuing equity on AIM compared to the MM." (Doukas and Hoque, p 386). Given the different characteristics of both markets, I will start by establishing the fact that the spread charged by the underwriters is different in both markets. Considering those AIM companies are relatively new and the AIM companies require a nominated adviser (who normally is the IPO underwriter) the underwriters will charge more. I propose the following hypothesis:

H 1. Underwriters spread is higher in the AIM compared to the Main market.

I will then start to investigate the reasons behind the differences in both markets as Abrahamson et al. (2011) examined the differences between the European and US market. as Abrahamson et al. (2011) and Corwin and Schultz (2005) have discussed how the spread charged is affected by the size of the proceeds raised in an IPO. Table 4.2 shows that the proceed size is higher in the Main market compared to the AIM market. I will examine if this is due to the concept of the economies of scale where the spread decreases as the proceed increases. I will examine this between the markets and within each market as well. I will look into the spread charged in both the Main and AIM markets in London Stock Exchange. After that, I will further examine our sample by dividing our sample into sub-groups based on the size of the IPO

proceeds and investigate the reasons behind them. The main market is comprised of companies with higher maturity level and larger proceeds expected from their IPOs. This implies that Main market enjoys relatively less risky companies. Larger IPO proceeds mean that those companies could benefit from the economies of scale. There has been much research discussing the underwriting pricing. They have discussed whether pricing varies over different markets and investigating the reasons behind that. The study by (Abrahamson et al., 2011) has found that the spread charged by the underwriters in the US market is almost fixed at 7%. In fact, they argue that it is becoming the norm for IPO rising up to $250 million. However, they have compared this to the European market where they have reported underwriters charge about 3% less than the US market. In this chapter, I would like to examine if the spread charged in the Main market and AIM market differs.

The second hypothesis that I would like to examine is the following:

H 2. Underwriters charge more on AIM because money raised in the IPOs is less (Dis-economies of scale).

Another difference is the level of riskiness between the two markets. Since, AIM attracts younger and newly created firms the riskiness of AIM should be higher than the Main market. There have been criticisms for AIM from different practitioners. Cited in Doukas and Hoque (2016) “For example, Roel Campos, a Commissioner at the US Securities and Exchange Commission, in 2007 was quoted saying “I’m concerned that 30% of issuers that list on AIM are gone in a year. That feels like a casino to me and I believe that investors will treat it as such.” Treanor, Jill “City hits out over US ‘casino’ jibe at AIM” The Guardian 10 March 2007. Similarly, John Thain, chief executive of the New York Stock Exchange (NYSE), criticized AIM for

its lack of regulation and corporate governance standards. Mr Thain, speaking at the World Economic Forum in Davos, Switzerland, stated that AIM "did not have any standards at all and anyone could list." James Quinn, NYSE Chief attacks AIM, The Telegraph, 27 January 2007." (p 38). To protect underwriters' reputation capital, they might be charging more for AIM companies.

The age of the firm is a proxy that has been utilised in a few studies. Normally it is believed that older firms are less risky. Another alternative that has been used in the literature is the underwriter reputation. (e.g., Carter and Manaster (1990) and Johnson and Miller (1988)). If the firms hire more reputable underwriters, the underwriters are thought to take on some risks for the firms. The reputable underwriters will do so to protect their reputation capital. Also, highly reputed underwriters will have done better due diligence to endorse the quality of the firm truthfully. Both of these proxies has been used in a number of empirical studies to examine the impact of risk on underpricing. Most of the results find a significant relationship between these risk proxies and underpricing. However, the explanatory power of these proxies are not very high as the r-squared values of these regressions are rather low.

I will look into two measures of risk. I will test the idiosyncratic risk which is the risk specified to the specific company going into the IPO. Another risk measure is the volatility of the share return after listing. However, the estimation of ex-ante risk for IPOs is considerably more troublesome in light of the fact that there is no authentic price data before the IPO to calculate risk. Most of the studies use the standard deviation of returns after the IPO as a proxy for risk before the IPO. By tradition, the standard deviations of the returns for the first twenty days utilised in many studies.

Since this ex-ante risk measure has demonstrated practically little explanatory power, some researchers conclude that risk does not fundamentally impact returns. Johnson and Miller (1988) reason that the standard deviation of post-IPO returns is a poor measure of ex-ante risk. Hence, a great part of the current research makes utilisation of totally distinctive risk proxies such as underwriter reputation as developed by Carter and Manaster (1990). I will also examine if the underwriter will mitigate the anticipated risk by underpricing to give compensation to the subscribers on the short-run. The third hypothesis that I am going to test is the following:

H 3. (a) Underwriters charge more on AIM because AIM IPOs are riskier.

Rock (1986) contends that underpricing is a result of the effect of the risk assumed by uninformed investors. The issues need to be underpriced to reward the uninformed investors and to reduce the chances that informed investors cannot take advantage of uninformed investors. Beatty and Ritter (1986) develops Rock's model further and find that underpricing increases with the riskiness of the IPO. Carter and Manaster (1990) develop another extension of Rock's model exhibiting that as the risk of an issue increases, informed demand will expand, compounding the adverse selection issue and the required underpricing. Generally, investors cannot differentiate less risky firms from high-risk firms. In any case, the IPOs can involve highly reputable investment banks to take on some of the risks and provide investors with the confidence that they are not risky firms, which permits them to underprice less. Booth and Smith (1986) asserts that the underwriter puts its reputational capital at stake which works as a covenant and shares price reflect all withheld information about the firms' performance. Underpricing gives both assurance and compensation to the utilisation of the underwriters' reputational capital. Tinic (1988) recommends that

underpricing is a type of protection to ensure underwriters against potential due diligence legitimate liabilities. Grinblatt and Hwang (1989) claim that underpricing is a signal by a more informed issuer to indicate firm value and the variance of expected returns to less informed investors. Thus, underwriters could mitigate the risk of the firms by underpricing more.

H 3(b) Alternatively, underwriters mitigate risk by lowering the issue price, i.e. underpricing more.

The underwriters usually become corporate brokers, a prerequisite for every organization recorded on the London Stock Exchange (LSE), who provide business related advice, new issue related services, co-ordinate with institutional investors and related services, and liaise with the London Stock Exchange and UK Listing Authority on administrative issues confronting listed organizations (FSA Listing Rules, 2007). They likewise execute equity related transactions, including insider trading, share repurchases, stake building in target organisations. Also, to keep their listing, AIM organisations are obliged to have a named investment bank that completes all the previously stated administrations. This infers that the association with the underwriters closer in the UK, in comparison to the US where firms often switch underwriters based on their performance during the IPO process, and to look for more prestigious underwriters and analyst coverage (e.g., Dunbar, 2000, and Ljungqvist, Marston, and Wilhelm, 2009). The close relationship with the underwriters in the AIM could also mean that there is a rent-seeking potentiality by the underwriters.

Rent-seeking by the underwriters have been investigated by Brav and Gompers (2003) in the context of IPO lockups in the US and Hoque and Lasfer (2015) in the context of IPO lockups in the UK. From my data, I have noticed that the lockup period

is slightly higher in AIM market than it is in the Main market. The main purpose of having a lockup period is to show that owners are committed to the business and hence will raise the investors' confidence to participate in the IPO. With higher expected risk and moral hazard from the smaller companies in AIM market, lockup period is expected to be higher. I will examine if the underwriters are proposing shorter lockup periods in return of higher spread. The fourth hypothesis that I am going to examine is the following:

H 4. Underwriter charge more in the AIM because of rent-seeking.

DuCharme, Rajgopal and Sefcik (2001) likewise look at marketing role during the process of IPOs. Their 'marketing theory' explores whether IPO underpricing is higher in the consumer oriented company such as business-to-consumer (B2C) web organisations. They also analyse for these set of firms whether IPO underpricing lead to increase in revenues after the IPO. Their evidence with respect to the marking hypotheses is mixed. They find that underpricing is significantly higher for B2C web organisations with respect to other web organisations, however, they do not find a significant relationship between underpricing and after IPO revenues. In another study, Aggarwal, Krigman, and Womack (2002) additionally contend that very high underpricing pulls in media consideration and makes attention for the issuing firm. On the relationship that underpricing pulls in media consideration and makes profitable attention, Demers and Lewellen (2003) expect an increase in web activity taking after the IPO. They find that increase in the web in the month after the IPO is significantly connected with IPO returns, and the impact is economically significant.

In a competitive market, many factors can be used to gain better market share. Pricing is one of the most important factors. I will examine if the future market share

is affected by the number of IPOs undertaken on the year before and I will test the significance of the spread charged by the underwriters on attracting future business. The fifth hypothesis that I are going to examine is the following:

H 5. Underwriters charge less to attract future business.

4.9 Data and methodology

4.10 Data

I have collected all the IPOs that took place on the London Stock Exchange for the purpose of our analysis. In this chapter, I am covering the IPO that took place in London Stock Exchange in both of its Main and AIM markets for the period starting from 1999 to 2012. I am following the same criteria used in some of the previous studies in order to come with consistent and comparable results. To gain a comprehensive understanding, in our sample, I have only included the IPOs that are new admissions and incorporated in the UK. I further excluded another 697 IPOs that are of financial and the investment firms as the behaviour of the firms operating in these sectors differ from the rest of the market.

I have collected the data in this study from a number of sources. First, I identified the IPOs that took place in the covered period from the London Stock Exchange new issue report.[7] Once I have the list of IPOs that took place on the London Stock Exchange, I have collected the IPO prospectuses from Perfect Filings database. From the prospectuses, I have manually collected the IPO related information. In particular, I collect underwriters' remuneration, gross proceeds, the number of bookrunners, lockup length. Some of the information such as the money raised from

[7] http://www.londonstockexchange.com/

the IPO proceeds, issuing price and the total number of shares are available both in the prospectus and on the LSE website. In the case of discrepancy, I used the data on the LSE website as it is assumed to have been reviewed and entered after the actual IPO took place. From Datastream, I collected the incorporation date to calculate the age of the company at the time of the IPO. I also collected some of the post-IPO information such as the size of capital over the 3 years following the IPO to calculate the companies' growth and the trading prices.

The initial list included 1075 IPOs. There were around 106 companies where their information was missing and difficult to obtain so I had to exclude them. The final list included 968 IPOs with a total proceed of £41.7 Billion.

Table 4.2 shows a summary of the number of IPOs and their sizes over the sampling period. I noticed from the table that the total proceeds for all the IPOs on our sample is about £41.7 Billion. The number of IPOs that fit the sampling criteria from AIM market is much higher than the IPOs from the Main market. With 806 IPOs from the AIM market make more than 83% of our sample. I also noticed that the number of IPOs was high during the Dot Com bubble then afterwards with the burst of the bubble the number of IPOs dropped. Afterwards, the number of IPOs has been increasing over the years till it was hit by the financial crisis in the year 2008 where the number of IPOs has dropped again. Nevertheless, the number of IPOs started to show increase the following years after the crisis. The total proceeds of the IPOs in the AIM market captured in our sample show an amount of about £10.1 Billion over 13 years. Yet, this only makes about 25% of the total IPO proceeds. The IPO proceeds for the AIM market ranged from a low £39K up to £307 Million with an average of £12.6 Million per IPO.

Table 4.2: IPO Sample Summary

	Main				AIM				The whole sample			
Year	Main IPOs	Total Proceeds	Average Proceeds size	Average Proceeds Median	AIM IPOs	Total Proceeds	Average Proceeds size	Average Proceeds Median	Total IPOs	Total Proceeds	Average Proceeds	Average Proceeds Median
1999	11	1,644.72	149.52	24.65	24	71.12	2.96	2.34	35	1,715.84	49.02	3.68
2000	52	5,079.53	97.68	38.22	80	493.19	6.16	3.88	132	5,572.72	42.22	9.24
2001	6	2,123.05	353.84	32.49	46	180.76	3.93	2.10	52	2,303.81	44.30	2.35
2002	9	1,039.80	115.53	29.19	30	75.57	2.52	1.86	39	1,115.37	28.60	2.64
2003	4	1,335.30	333.82	209.91	30	112.92	3.76	1.57	34	1,448.22	42.59	2.38
2004	15	1,354.98	90.33	63.10	122	1,000.45	8.20	3.40	137	2,355.43	17.19	4.29
2005	15	2,678.85	178.59	114.47	166	2,016.46	12.15	4.07	181	4,695.31	25.94	4.47
2006	14	2,963.51	211.68	149.71	117	2,264.79	19.36	4.82	131	5,228.30	39.91	6.24
2007	18	3,115.85	173.10	80.45	81	1,622.31	20.03	5.65	99	4,738.16	47.86	12.77
2008	2	940.47	470.24	470.24	16	355.37	22.21	12.17	18	1,295.84	71.99	12.83
2009	1	55.82	55.82	55.82	8	523.22	65.40	47.29	9	579.04	64.34	55.82
2010	8	1,942.42	242.80	197.86	27	659.12	24.41	20.48	35	2,601.54	74.33	33.16
2011	5	7,140.79	1,428.16	342.54	33	391.15	11.85	3.26	38	7,531.94	198.21	4.32
2012	2	132.00	66.00	139.93	26	398.50	15.33	6.93	28	530.50	18.95	7.40
Total	162	31,547.08	194.74	73.15	806	10,164.94	12.61	3.93	968	41,712.02	43.09	5.58

This table shows a summary of the IPO dataset used in our study. Total proceeds are the amount of money raised in the IPO; the average proceeds size is the mean proceeds, and the average proceeds median is the average money raised. Total proceeds, average proceeds size and average proceeds median are in GBP million. This information is shown per year and for the Main, AIM and the whole sample.

The sample includes 162 IPOs from the Main market where the total proceeds are more than £31 Billion which makes about 75% of the total proceeds. The IPO sizes ranged from £1.58 Million to £6 Billion with an average of £194.7 Million per IPO. I have taken the effect of inflation into consideration. I have discounted all the amounts in our data sample to 2012 GBP value. I have used CPI index from the Office for National Statistics.

I have also classified the underwriters' status to be either a prestigious or non-prestigious. I followed (Derrien and Kecskes, 2007). "I classify a broker as "prestigious" if it is a global investment bank.[8] In instances in which prestige is not obvious, I consult the 1997 to 2003 editions of Thomson's Extel Survey" (Derrien and Kecskes, 2007). However, I have consulted the 2013 Thomson's Extel Survey to check the classification further as Derrien and Kecskes (2007) sample ends in 2003.

4.11 Methodology

To test our hypotheses, I first use univariate tests such as correlation analysis, plots for spreads and tabulate several univariate tables. Then, I proceed to the multivariate tests where I apply ordinary Least Square (OLS) regressions in a linear model. After that, I conduct several robustness checks including fixed effects regression, running the OLS while winsorizing spread, Tobit model and also considered a non-linear model. The models are presented in detail when I discuss my empirical results.

To test if the spread charged in the previous year is used to attract more business and how it would affect the future business I ran a couple of OLS models.

---

[8] I follow Derrien and Kecskes (2007) and include in prestigious underwriters global investment banks such as ABN AMRO (including Hoare Govett), Cazenove & Co., Credit Lynnais Securities, Dresdner Kleinwort Wassertein, HSBC Securities, Credit Suisse, Investec Hendersen Crosthwaite securities, KBC Securities, Peel Hunt, Lehman brothers, Nomura International, Schroder Salomon Smith Barney, SG securities, UBS, West LB, Merrill Lynch International, Goldman Sachs.

Both of the models examine the effect of the number of IPO taken, the average proceeds raised in them and the average spread charged by the underwriter in the previous year to the number of IPOs and then to the average spread charged in the current year.

Then I examine the determinants of underpricing because spread charged is the direct cost of money raised and underpricing is an indirect cost of IPO. Hence, the spread charged and IPO underpricing could be substitutes or complementary. I try to shed light on this issue by running an OLS regression on IPO underpricing.

I have not controlled for the industry in our models. This could cause a heterogeneity problem. However, I have segmented our sample based on the size of the IPO which will mitigate the heterogeneity when running the regression models for the segment. Yet, heterogeneity could still be present when regress the whole sample.

4.12 Results and analysis

In this section will analyse the collected data and will test the hypothesis identified earlier.

4.13 Descriptive Statistics

4.13.1.1 Main and AIM markets statistics

During the period that I am covering in this chapter, the IPO market in the UK and globally has gone through ups and downs in terms of number of IPOs and the total amount of money raised. In this section, I will examine whether the results I get are in line with some of the previous studies such as Abrahamson et al., 2011; Armitage, 2000; Chen and Mohan, 2002; H. C. Chen and Ritter, 2000 and some of the previous studies who argued that unlike the US market, the spread in the European market is not clustered.

I will compare the results across our sample between the Main market and AIM market and discuss the reasons behind the differences in terms of the nature of the markets and the markets components and the laws governing each of them.

Table 4.2 shows a summary of the data of the whole IPO data sample under study over the period from 1999 to 2012. In this table, I have divided the IPOs based on the value of the proceeds received. To produce comparable results to the paper of Abrahamson et al. (2011), I have identified four ranges from our sample in addition to the whole sample and that of Main and AIM market. I have used an average exchange rate of $1.5 per £1. The sets and the subsets based on the ranges are the whole sample, Main Market, AIM market, Small with less than £16.7 Million, Medium with proceeds between £16.7 Million – £66.7 Million, Large between £66.7 - £333.4 and Very Large with proceeds more than £333.4. The number if IPOs in each of the above subset is 968, 162, 806, 686, 169, 93, and 20, respectively (in the same order).

In Table 4.3 I show some statistics across the years, I am covering. I show the number of IPOs that took place in each year and average spread charged by the underwriters in terms of mean and median. I also show the average number of bookrunners per IPO which gives an idea of the frequency of using the multi-bookrunners across the years too. I then show the return at the end of the first trading day and at the end of the fifth trading day when the price is more stable. In panels B, C, and D, I have shown the information for the different ranges. In order to compare the results from both markets, I have generated two more tables where Table 4.4 shows the data for the Main market and Table 4.5 shows the data for AIM market. In order to remove the effect of the Dot Com bubble, I have also shown the results excluding the years 1999 and 2000 for all panels in the three tables.

From Table 4.3, I notice that the median spread charged by the underwriters ranges between 4% - 6.69% with a median of 5% over the entire sample. The mean of the spread over the whole sample, on the other hand, ranges from 4.40% to 8.84%. However, looking at both Table 4.4 and Table 4.5, shows that the main attribution to this number comes from the AIM market as it has more IPOs. The median spread for the Main market shown in Table 4.4 ranges between 2.23% - 6.25% with an all-year median of 4%. Table 4.5 shows that that the median spread ranges between 3.76% - 7.5% with a median over all the years of 5.08%. The mean of multi-bookrunners is lowest at 1 in 2001 and highest at 1.50 at the year 2011, with an overall average of 1.12 for the entire sample.

Table 4.3: IPOs Proceeds and spread charged by the Underwriters for Main and AIM Markets

Panel A									
Year	Count	Mean of multi-bookrunners	Spread (Mean) %	Spread (Median) %	Median Proceeds (Millions)	D1 Return (Mean) %	D1 Return (Median) %	D5 Return (Mean) %	D5 Return (Median) %
1999	35	1.11	4.40	4.00	3.68	25.13	6.25	62.63	11.67
2000	132	1.02	4.28	4.00	9.24	32.60	10.25	84.53	11.27
2001	52	1.00	6.42	5.00	2.35	25.38	8.18	17.54	7.76
2002	39	1.05	6.34	5.40	2.64	3.01	7.22	-2.36	6.11
2003	34	1.06	8.84	6.20	2.38	14.13	6.31	16.16	9.16
2004	137	1.06	6.39	5.00	4.29	17.95	9.15	122.97	9.38
2005	181	1.06	6.29	5.00	4.47	30.13	11.54	71.07	11.54
2006	131	1.14	6.52	5.07	6.24	15.09	10.00	46.14	10.19
2007	99	1.12	5.70	5.00	12.77	10.04	8.00	11.05	6.43
2008	18	1.17	5.71	4.90	12.83	9.90	6.19	18.78	6.48
2009	9	1.22	5.38	4.00	55.82	7.50	3.50	53.05	8.50
2010	37	1.41	4.99	5.00	33.16	12.20	0.00	19.56	8.60
2011	38	1.50	7.19	6.69	4.32	-7.16	0.00	8.49	3.65
2012	30	1.23	8.11	6.43	7.40	-1.85	0.00	15.13	9.08
All Years	972	1.11	6.05	5.00	5.58	18.88	8.09	55.44	9.20
All years Ex. 1999 - 2000	805	1.12	6.42	5.00	5.35	16.36	7.95	50.35	9.09
Panel B: 16.7 < Proceeds < 66.7									
All Years	168	1.10	4.48	4.39	30.63	17.26	7.92	6.38	7.74
All years Ex. 1999 - 2000	131	1.12	4.61	4.43	31.89	11.62	6.67	-0.45	7.50
Panel C: 66.7 < Proceeds < 333.4									
All Years	96	1.50	3.85	3.75	119.49	-0.53	2.84	-6.81	2.53
All years Ex. 1999 - 2000	77	1.57	3.88	3.72	120.98	1.06	3.00	-3.73	3.00
Panel D: Proceeds > 333.4									
All Years	22	2.18	3.31	3.00	643.65	-1.19	0.97	44.48	0.45
All years Ex. 1999 - 2000	18	2.39	3.25	2.73	612.64	-3.86	0.15	-6.37	-0.47

This table shows a summary of the IPOs proceeds and the fees charged and whether the shares were underpriced or not in day 1 and 5 of trading for the whole sample (both the Main and AIM Markets together). I have divided the table into 4 panels based on total proceeds. In panel A of this table, I have shown all the IPOs from our sample. Panel B shows IPOs with a proceed between £16.7M - £66.7M, Panel C shows IPOs with a proceed between £66.7M - £333.4M and Panel D shows the IPOs with a proceed larger than £333.4M. At the bottom of each panel, I have excluded the period 1999-2000 to eliminate the effect of the .com bubble. This table also shows the mean of the multi-bookrunner. It is the arithmetic mean of the number of the bookrunners for the IPOs during the mentioned period. The spread is the cost charged by the underwriters as a percentage of the gross proceeds. The spread is shown as mean and median in this table. The table also shows the median of the absolute amount of the proceeds in millions. The table shows the underpricing factor for Day 1 and Day 5 of trading as mean and median (D1 Return, D5 Return prospectively). The underpricing factor is calculated as the return on trading. If IPO was underpriced, the underpricing factor is positive, if it was overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0.

Since underpricing is an indirect cost of IPO, I also compute the underpricing by following Abrahamson et al., (2011). In particular, I have calculated the overpricing and underpricing of the shares' issue price using the return of both the first and the fifth trading days. This factor is simply the return of a share having been purchased at the IPO and sold at the closing price of the (first/fifth) trading day. When I look at the whole sample in Table 4.3, I can see that the mean return for the first trading day ranges between -7.16% and 32.60%, with an average of 16.36% for the whole sample. However, median ranges between 0.00% and 11.54%, with an average median of 7.95%. Underpricing based on day 5 even shows a higher level of underpricing. I notice that the mean return for the fifth trading day ranges between -2.36% and 122.97%, with an average of 50.35% for the whole sample. However, median ranges between 3.65% and 11.67%, with an overall median of 9.09%.

The average spread charged for the whole market for the whole sample and then for the medium, large and very large subsets are 6.05%, 4.48%, 3.85%, and 3.31%, respectively. When looking at Table 4.4 for the Main market, I find that the average spread charged for the in the same order above are 4.04%, 4.37%, 3.87% and 3.31%. Finally, from Table 4.5 for AIM market the average spread for the different subsets are 6.47%, 4.55%, 3.81% while there were no IPO in our sample with proceeds that exceed 333.4M in the AIM market. This explains how the average spread for both the whole sample and the Main market sample with proceeds more than 333.4M are the same. In comparison to the paper by Abrahamson et al., (2011), for the European markets, they reported 3.43% while I reported 4.04% for the Main market. The Main market shows to be closer those results and more in line than the AIM market results of 6.47% or even for the whole sample of 6.05% as it is affected by the high number of IPOs in the AIM market. They have reported in their paper the following spreads:

4.22%, 3.76%, and 2.60% for the ranges ($25Million - $100 Million, $100 Million - $500 Million and over $500 Million). These results are comparable to the results I got from the Main market which are 4.37%, 3.87% and 3.31% respectively. So, in other words, the larger the size of the IPO the lower the gross spread becomes. This provides an early indication of the economies of scale hypotheses that I propose in the Hypotheses section.

The table shows that the level of underpricing was much lower following the financial crisis in 2008. When I look at the median, the lowest values are recorded in the year 2009 onwards with a value of 3.50% in the year 2009 and 0.00% in the following three years. The effect is even more explicit when looking at the mean. Although that for the year 2010 the average return was 12.20%, the average return for the following two years shows a loss of 7.16% and 1.85% consecutively. Yet, spread charged seems to be increasing during those years. I have come to the same conclusion when I look at Table 4.4 and Table 4.5. This also could be attributed to the low number of IPOs during the last years in comparison to the years before the financial crisis. The mean of the underpricing factor for both the first, the fifth trading days, and the median of the fifth trading day show that for the AIM market, there is a higher tendency to underprice the shares which result in higher returns. The mean of the underpricing factor shows more tendency to overprice if I look at the first trading day. This tendency seems to ease where it comes closer to the fair price as I approach the fifth trading day.

The positive relationship between the underpricing and spread charged can also be noticed in the paper by Abrahamson et al., 2011. Looking at the European markets, I can see that for the years 1999 and 2000, average return jumped from 8.0% in the year 1998 to 11.4% and 11.8%. The spread charged also increased on average

from 3.98% in the year 1998 to 4.09% and 4.24% in the following two years. On the other hand, this is not noticed on the US market as the spread charged seems to be unchanged.

Table 4.4: IPOs Proceeds and spread charged by the Underwriters for Main Market

Panel A: Main Market									
Year	Count	Mean of multi-bookrunners	Spread (Mean) %	Spread (Median) %	Proceeds (Millions) (Median)	D1 Return (Mean) %	D1 Return (Median) %	D5 Return (Mean) %	D5 Return (Median) %
1999	11	1.27	3.73	3.75	24.65	21.32	18.23	120.38	14.72
2000	52	1.06	4.11	4.00	38.22	22.82	9.33	28.30	5.85
2001	6	1.00	4.30	4.56	32.49	10.06	7.96	13.39	13.45
2002	9	1.22	4.06	4.39	29.19	1.45	8.19	-12.17	3.33
2003	4	1.50	3.71	4.13	209.91	4.20	0.26	6.70	5.25
2004	15	1.40	4.15	4.25	63.10	0.12	5.00	-2.25	4.93
2005	15	1.53	3.79	3.75	114.47	59.21	7.67	-0.63	2.50
2006	14	1.86	4.02	3.99	149.71	13.51	13.01	13.66	13.01
2007	18	1.39	4.02	3.65	80.45	9.78	4.58	0.55	3.04
2008	2	1.00	6.25	6.25	470.24	-12.17	-12.17	-12.17	-12.17
2009	1	2.00	4.00	4.00	55.82	0.00	0.00	11.11	11.11
2010	10	1.90	3.16	2.88	197.86	1.78	0.00	8.24	13.52
2011	5	4.20	2.98	2.23	342.54	-22.75	0.00	-23.36	-0.94
2012	4	2.25	6.95	4.00	139.93	0.24	0.00	3.10	3.04
All Years	166	1.45	4.04	4.00	73.15	15.95	6.52	17.57	5.09
All years Ex. 1999 - 2000	103	1.66	4.04	3.78	109.55	11.90	5.13	1.17	4.93
Panel B: 16.7 < Proceeds < 66.7									
All Years	67	1.09	4.37	4.50	31.02	35.60	10.50	16.90	10.00
All years Ex. 1999 - 2000	35	1.14	4.60	4.50	41.59	31.67	9.52	2.63	11.11
Panel C: 66.7 < Proceeds < 333.4									
All Years	63	1.65	3.87	3.75	139.63	-0.47	2.55	-4.84	2.50
All years Ex. 1999 - 2000	44	1.84	3.94	3.57	152.31	2.35	2.43	1.42	3.46
Panel D: Proceeds > 333.4									
All Years	22	2.18	3.31	3.00	643.65	-1.19	0.97	44.48	0.45
All years Ex. 1999 - 2000	18	2.39	3.25	2.73	612.64	-3.86	0.15	-6.37	-0.47

This table shows a summary of the IPOs proceeds and the fees charged and whether the shares were overpriced or underpriced in day 1 and 5 of trading as a total for the Main Market. I have divided the table into 4 panels based on total proceeds. In panel A of this table, I have shown all the IPOs from our sample. Panel B shows IPOs with a proceed between £16.7M - £66.7M, Panel C shows IPOs with a proceed between £66.7M - £333.4M and Panel D shows the IPOs with a proceed larger than £333.4M. At the bottom of each panel, I have excluded the period 1999-2000 to eliminate the effect of the .com bubble. This table also shows the mean of the multi-bookrunner. It is the arithmetic mean of the number of the bookrunners for the IPOs during the mentioned period. The spread is the cost charged by the underwriters as a percentage of the gross proceeds. The spread is shown as mean and median in this table. The table also shows the median of the absolute amount of the proceeds in millions. The table shows the underpricing factor for Day 1 and Day 5 of trading as mean and median (D1 Return, D5 Return prospectively). The underpricing factor is calculated as the return on trading day. If IPO was underpriced, the underpricing factor is positive, if it was overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0.

Table 4.5: IPOs Proceeds and spread charged by the Underwriters for AIM Market

Panel A									
Year	Count	Mean of multi-bookrunners	Spread (Mean) %	Spread (Median) %	Proceeds (Millions) (Median)	D1 Return (Mean) %	D1 Return (Median) %	D5 Return (Mean) %	D5 Return (Median) %
1999	24	1.04	4.71	4.16	2.34	26.87	4.77	36.16	6.53
2000	80	1.00	4.39	4.00	3.88	38.96	12.02	121.08	17.30
2001	46	1.00	6.70	5.54	2.10	27.37	8.21	18.08	7.76
2002	30	1.00	7.02	7.04	1.86	3.48	6.18	0.58	6.42
2003	30	1.00	9.53	6.57	1.57	15.45	6.68	17.42	9.16
2004	122	1.02	6.67	5.19	3.40	20.14	9.95	138.36	10.53
2005	166	1.02	6.51	5.30	4.07	27.50	12.63	77.55	12.89
2006	117	1.05	6.82	5.40	4.82	15.28	9.38	50.03	10.16
2007	81	1.06	6.08	5.00	5.65	10.10	8.75	13.39	7.61
2008	16	1.19	5.64	4.78	12.17	12.66	8.01	22.65	7.88
2009	8	1.13	5.55	3.76	47.29	8.43	3.63	58.30	8.50
2010	27	1.22	5.67	5.30	20.48	16.06	0.00	23.76	8.06
2011	33	1.09	7.83	7.50	3.26	-4.80	0.00	13.32	8.16
2012	26	1.08	8.28	6.77	6.93	-2.17	0.00	16.99	9.65
All Years	806	1.04	6.47	5.08	3.93	19.49	8.48	63.24	10.00
All years Ex. 1999 - 2000	702	1.04	6.77	5.36	4.07	17.01	8.33	57.57	9.68
Panel B: 16.7 < Proceeds < 66.7									
All Years	101	1.11	4.55	4.31	30.44	5.10	5.63	-0.59	6.98
All years Ex. 1999 - 2000	96	1.11	4.62	4.35	30.44	4.30	5.52	-1.57	6.65
Panel C: 66.7 < Proceeds < 333.4									
All Years	33	1.21	3.81	4.00	101.88	-0.65	4.00	-10.58	3.00
All years Ex. 1999 - 2000	33	1.21	3.81	4.00	101.88	-0.65	4.00	-10.58	3.00
Panel D: Proceeds > 333.4									
All Years	0	0	0	0	0	0	0	0	0
All years Ex. 1999 - 2000	0	0	0	0	0	0	0	0	0

This table shows a summary of the IPOs proceeds and the fees charged and whether the shares were overpriced or underpriced in day 1 and 5 of trading as a total for the AIM Market. I have divided the table into 4 panels based on total proceeds. In panel A of this table, I have shown all the IPOs from our sample. Panel B shows IPOs with a proceed between £16.7M - £66.7M, Panel C shows IPOs with a proceed between £66.7M - £333.4M and Panel D shows that there are no IPOs with a proceed larger than £333.4M. At the bottom of each panel, I have excluded the period 1999-2000 to eliminate the effect of the .com bubble. This table also shows the mean of the multi-bookrunner. It is the arithmetic mean of the number of the bookrunners for the IPOs during the mentioned period. The spread is the cost charged by the underwriters as a percentage of the gross proceeds. The spread is shown as mean and median in this table. The table also shows the median of the absolute amount of the proceeds in millions. The table shows the underpricing factor for Day 1 and Day 5 of trading as mean and median (D1 Return, D5 Return prospectively). The underpricing factor is calculated as the return on trading day. If IPO was underpriced, the underpricing factor is positive, if it was overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0.

4.13.1.2 Bookrunners statistics

In our sample, I have identified 150 bookrunners. I have classified 35 of them as Prestigious while the remaining 115 bookrunners are identified as Non-Prestigious. I have followed Derrien and Kecskes, 2007 in identifying the prestigious bookrunners. I have identified the global investment firms as prestigious in case if the bookrunners status is not obvious, Iconsult the Thomson's Extel Survey. For my sample, I have noticed that although the prestigious bookrunners make up only 23% of the total list of the bookrunners, they underwrote more than 43% of the IPOs in our sample. Moreover, the prestigious bookrunners were dominant in the Main market where they were present in more than 88% of the IPOs. In the AIM market, however, they were overtaken by the non-prestigious where the prestigious made only about 39% of the IPOs.

Table 4.6, Panel A shows the top 10 prestigious bookrunners ranked on the number of IPOs undertaken during the sample period. In this list, I can see that most of the companies have undertaken more IPOs in the AIM market compared to IPOs undertaken in the Main market. This bias towards the AIM market can be explained by the fact that the IPOs from the AIM market make about 83% of the whole sample. This is the case for seven of the top ten bookrunners except for Merrill Lynch and Credit Suisse that have undertaken more IPOs in the Main market. Investec Bank seems to have almost divided the number of IPOs between the two markets. However, the bookrunners that have undertaken more of the Main market IPO show higher average fees in the Alternative Investment market. This is obvious as this is due to the higher IPO proceeds in the Main market than the AIM market.

Table 4.6: Top 10 Prestigious bookrunners ranking

N	Bookrunner	Total IPOs	Main IPOs	AIM IPOs	Total Average Fees (Million)	Main Average Fees (Million)	AIM Average Fees (M)	Total Average Proceeds (Million)	Main Average Proceeds (Million)	AIM Average Proceeds (Million)	Total Average Spread %	Average Spread Main %	Average Spread AIM %
Panel A: Underwriters ranking based on number of IPOs													
1	KBC PEEL HUNT LTD (UK) Limited	54	4	50	455.77	1,081.53	405.71	12.61	28.64	11.33	5.05	4.23	5.05
2	Seymour Pierce Limited	49	1	48	429.72	3,057.99	374.96	9.51	116.36	7.29	2.66	7.14	2.66
3	Collins Stewart	40	7	33	925.62	1,287.55	848.85	20.52	30.35	18.44	4.15	6.16	4.15
4	Evolution Securities Limited	36	7	29	766.05	1,269.38	644.56	17.32	29.33	14.43	3.80	5.65	3.80
5	WH Ireland Limited	34	0	34	200.59	-	200.59	2.96	-	2.96	0.00	10.14	0.00
6	Brewin Dolphin	33	3	30	401.43	1,261.82	315.39	9.36	37.46	6.55	3.13	7.10	3.13
7	Numis Securities Limited	28	0	28	1,428.85	-	1,428.85	33.37	-	33.37	0.00	5.46	0.00
8	BofA Merrill Lynch	27	25	2	16,066.11	17,158.74	2,408.21	544.58	583.84	53.82	3.74	4.30	3.74
9	Investec Bank plc	25	11	14	1,171.15	1,593.70	839.15	25.50	37.73	15.89	4.17	5.50	4.17
10	Credit Suisse	24	23	1	11,578.49	11,958.85	2,830.30	463.47	480.38	74.48	3.77	3.80	3.77
Panel B: Underwriters ranking based on number of Total Average Proceeds													
1	BNP PARIBAS	1	1	0	105389.943	105,389.94	-	6022.28	6022.28	-	1.75	1.75	0.00
2	Société Générale	1	1	0	105389.943	105,389.94	-	6022.28	6022.28	-	1.75	1.75	0.00
3	Barclays Capital	2	2	0	54688.4164	54,688.42	-	3110.81	3110.81	-	1.88	1.88	0.00
4	BofA Merrill Lynch	27	25	2	16066.1117	17,158.74	2408.21138	544.58	583.84	53.82	3.78	3.74	4.30
5	MORGAN STANLEY and CO. INTERNATIONAL PLC	18	15	3	15278.085	16997.6026	6680.49704	595.12	676.15	189.98	3.87	3.88	3.83
6	Citi	15	13	2	14649.5905	15,876.86	6672.308	640.78	714.55	161.33	3.42	3.33	4.00
7	Credit Suisse	24	23	1	11,578.49	11958.8495	2,830.30	463.47	480.38	74.48	3.78	3.77	3.80
8	UBS Investment Bank	22	20	2	10,222.43	10,938.08	3,066.00	433.70	466.99	100.83	3.89	3.98	2.95
9	J.P. Morgan Cazenove	20	16	4	9,726.64	11,239.26	3676.18115	235.74	273.37	85.25	3.95	3.62	5.28
10	Goldman Sachs International	19	19	0	9,440.48	9,440.48	-	259.45	259.45	-	3.91	3.91	0.00

Panel C: Underwriters ranking based on number of Total Average Proceeds with more than 5 IPOs

1	BofA Merrill Lynch	27	25	2	16066.1117	17,158.74	2408.21138	544.58	583.84	53.82	3.78	3.74	4.30
2	MORGAN STANLEY and CO. INTERNATIONAL PLC	18	15	3	15278.085	16,997.60	6680.49704	595.12	676.15	189.98	3.87	3.88	3.83
3	Citi	15	13	2	14649.5905	15,876.86	6672.308	640.78	714.55	161.33	3.42	3.33	4.00
4	Credit Suisse	24	23	1	11578.4935	11,958.85	2830.30488	463.47	480.38	74.48	3.78	3.77	3.80
5	UBS Investment Bank	22	20	2	10222.4345	10938.0776	3066.00336	433.70	466.99	100.83	3.89	3.98	2.95
6	J.P. Morgan Cazenove	20	16	4	9726.64388	11,239.26	3676.18115	235.74	273.37	85.25	3.95	3.62	5.28
7	Goldman Sachs International	19	19	0	9,440.48	9440.47852	-	259.45	259.45	-	3.91	3.91	0.00
8	Deutsche Bank	14	9	5	7,610.31	8,710.65	5,629.69	203.82	237.92	142.44	4.62	4.74	4.40
9	HSBC	9	7	2	4,179.90	3,409.87	6874.99748	168.10	169.99	161.48	4.13	3.39	6.73
10	Cazenove and Co. Ltd	10	8	2	3,422.67	4,109.75	674.34	89.18	104.75	26.93	4.04	4.32	2.92

This table shows the top prestigious bookrunners. Panel A ranks the bookrunners based on the number of IPOs. Panel B ranks them based on the total average proceeds. Panel C ranks them based on the number of IPOs but excluding bookrunners who have undertaken less than 5 IPOs to overcome the bookrunners who only participated in the Mega IPOS. The table shows The total IPOs. Average Fees charges, Average proceeds for the IPOs and the average spread charged by the bookrunners as a percentage of the total proceeds for the whole sample, then for the Main and AIM markets.

In Table 4.6, Panel B, the bookrunners are ranked based on the total average proceeds. Table 4.6, Panel B shows more bookrunners undertaking IPOs in the main market. At the top 3, I can see bookrunners that have been present only once in our sample such as BNP PARIBAS and Société Générale or only twice like Barclays Capital. For the first two, they were both in the huge Glencore Intl plc IPO with proceeds of £6bn. However, they were part of 8 bookrunners who shared the £105M. Yet, as I am ranking based on the average fees, they showed at the top. To overcome this issue, Table 4.6, Panel C shows the top ten bookrunners that have undertaken five or more IPOs during the sampling period. This table shows all the bookrunners that have undertaken IPOs in both the Main and the AIM markets except for Goldman Sachs International that has only undertaken Main market IPOs based on our sample. Thus, most of the IPOs undertaken by the bookrunners in this list are from the Main market.

Table 4.7, Panel A shows the top non-prestigious bookrunners ranked on the number of IPOs. Among the top ten bookrunners, I can see that all of them have undertaken IPOs from the AIM market. However, the table shows that only three of them have also undertaken IPOs from the Main market. Nevertheless, those three companies have only taken one, two, and four IPOs from the Main market. Moreover, the average proceeds of those IPOs ranged between £14m and £39m.

In Table 4.7, Panel B, I have ranked the non-prestigious bookrunners on the total average fees. This resulted in showing some bookrunners that have less frequently appeared in our sample. Eight bookrunners out of the top ten have undertaken IPOs from the Main market while only six have undertaken IPOs from AIM market. Again, I am faced with the same problem of showing bookrunners with only a few lucrative IPOs.

Table 4.7: Top 10 Non-Prestigious bookrunners ranking

N	Bookrunner	Total IPOs	Main IPOs	AIM IPOs	Total Average Fees (Million)	Main Average Fees (Million)	AIM Average Fees (Million)	Total Average Proceeds (Million)	Main Average Proceeds (Million)	AIM Average Proceeds (Million)	Total Average Spread %	Average Spread Main %	Average Spread AIM %
Panel A: Underwriters ranking based on number of IPOs													
1	Teather and Greenwood Limited	26	0	26	295.03	-	295.03	7.19	-	7.19	5.67	0.00	5.67
2	Charles Stanley Securities	23	0	23	269.86	-	269.86	4.40	-	4.40	6.49	0.00	6.49
3	Daniel Stewart and Company plc	19	0	19	216.02	-	216.02	3.28	-	3.28	7.63	0.00	7.63
4	William de Broe Plc	18	4	14	385.27	574.21	331.28	7.84	14.18	6.02	5.45	4.09	5.83
5	Cenkos Securities plc	17	0	17	1474.78	-	1474.78	31.00	-	31.00	6.42	0.00	6.42
6	Arbuthnot Securities Limited	16	1	15	406.94	1,372.29	342.58	7.83	28.49	6.45	6.05	4.82	6.13
7	Hichens, Harrison and Co. plc	16	0	16	121.36	-	121.36	2.14	-	2.14	7.25	0.00	7.25
8	Noble and Company Limited	15	0	15	319.93	-	319.93	4.93	-	4.93	7.91	0.00	7.91
9	J M Finn and Co	15	0	15	174.61	-	174.61	2.84	-	2.84	7.40	0.00	7.40
10	Panmure Gordon and Co	14	2	12	968.22	2,231.79	757.63	21.64	39.50	18.66	5.00	5.96	4.84
Panel B: Underwriters ranking based on number of Total Average Proceeds													
1	Schroder Salomon Smith Barney	1	1	0	17095.49	17,095.49	-	569.85	569.85	-	3.00	3.00	0.00
2	Lazard and Co. Limited	3	3	0	14201.32	14,201.32	-	452.49	452.49	-	3.67	3.67	0.00
3	Bear Stearns	1	0	1	12309.27		12309.27	307.73	-	307.73	4.00	0.00	4.00
4	Warburg Dillon Read	1	1	0	4414.30	4,414.30	-	98.10	98.10	-	4.50	4.50	0.00
5	Renaissance Securities (Cyprus) Limited	6	2	4	4023.05	7884.23	2092.46	112.12	252.7	41.82	4.56	3.61	5.03
6	Pathfinder Prospectus	1	1	0	3575.12	3,575.12	-	102.15	102.15	-	3.50	3.50	0.00
7	Fairfax I.S. PLC	6	1	5	3,332.84	4539.76	3,091.45	73.38	108.17	66.42	4.69	4.20	4.79
8	RBC Capital Markets	2	1	1	2,776.49	5,134.09	418.90	116.88	228.18	5.59	4.88	2.25	7.50
9	Mirabaud Securities LLP	10	1	9	2,723.40	8,143.31	2121.19	46.57	162.87	33.65	5.94	5.00	6.05
10	FOX-PITT, KELTON LIMITED	1	0	1	2,161.95	-	2,161.95	259.45	259.45	-	3.91	3.91	0.00

Panel C: Underwriters ranking based on number of Total Average Proceeds with more than 5 IPOs

1	Renaissance Securities (Cyprus) Limited	6	2	4	4,023.05	7,884.23	2,092.46	112.12	252.70	41.82	4.56	3.61	5.03
2	Fairfax I.S. PLC	6	1	5	3,332.84	4,539.76	3,091.45	73.38	108.17	66.42	4.69	4.20	4.79
3	Mirabaud Securities LLP	10	1	9	2,723.40	8,143.31	2,121.19	46.57	162.87	33.65	5.94	5.00	6.05
4	ORIEL SECURITIES LIMITED	9	1	8	2,101.67	5,701.73	1,651.67	54.25	162.91	40.67	4.51	3.50	4.64
5	Canaccord Genuity Limited	6	0	6	1,892.49	-	1,892.49	34.75	-	34.75	5.88	0.00	5.88
6	Liberum Capital Limited	8	0	8	1,736.26	-	1,736.26	41.16	-	41.16	4.88	0.00	4.88
7	Cenkos Securities plc	17	0	17	1,474.78	-	1,474.78	31.00	-	31.00	6.42	0.00	6.42
8	Altium Capital Limited	10	2	8	1,163.59	1,875.95	985.50	29.73	50.52	24.53	4.13	3.63	4.25
9	Panmure Gordon and Co	14	2	12	968.22	2,231.79	757.63	21.64	39.50	18.66	5.00	5.96	4.84
10	Fox-Davies Capital Limited	6	0	6	807.52	-	807.52	13.60	-	13.60	7.26	0.00	7.26

This table shows the top Non-prestigious bookrunners. Panel A ranks the bookrunners based on the number of IPOs. Panel B ranks them based on the total average proceeds. Panel C ranks them based on the number of IPOs but excluding bookrunners who have undertaken less than 5 IPOs to overcome the bookrunners who only participated in the Mega IPOS. The table shows The total IPOs. Average Fees charges, Average proceeds for the IPOs and the average spread charged by the bookrunners as a percentage of the total proceeds for the whole sample, then for the Main and AIM markets.

To overcome this issue as I did with the prestigious bookrunners, in Table 4.7, Panel C, I have regenerated this table showing bookrunners that have at least undertaken five IPOs. For both tables, I can notice that they are dominated by IPOs from the AIM market. It shows some bookrunners undertaking one or two IPOs from the Main market.

4.13.1.3 Spread Clustering

(Abrahamson, et al., 2011) argued in their paper that underwriters in the US market show an average fee of about 7% of the proceeds while the underwriters in the European markets, in contrast, tend to charge less. In Table 4.8, Panel A I have shown the percentage of IPOs were bookrunners charge 7% in comparison to those who charge differently in both the Main and AIM markets. The table shows that the percentage of underwriters who are charging fees of 7% varies across the years by only 4.22% over the whole sample. This result is driven by the IPOs that joined the AIM. Only one IPO in the Main market charged around 7%, while 40 IPOs in the AIM charged around 7%. These results also highlight the fact that underwriters in the AIM charge more than in the Main market. See Table 4.8.

The information is also presented in Figure 1 showing the number of bookrunners who are charging 7% compared to those who are charging differently in both the Main and AIM markets.

Table 4.8: Number of IPOs charging about 7% and 5% of the net proceeds

Year	Total IPO	Total AIM	Total Main	Total different from 7%	Total about 7%	AIM about 7%	Main about 7%	Percentage of IPOs charging about 7%
Panel A: Number of IPOs charging about 7% of the net proceeds (Between 6.7 and 7.3)								
1999	35	24	11	33	2	2	0	5.71
2000	132	80	52	130	2	2	0	1.52
2001	52	46	6	50	2	2	0	3.85
2002	39	30	9	36	3	3	0	7.69
2003	34	30	4	32	2	2	0	5.88
2004	137	122	15	130	7	7	0	5.11
2005	181	166	15	172	9	9	0	4.97
2006	131	117	14	129	2	2	0	1.53
2007	99	81	18	97	2	2	0	2.02
2008	18	16	2	18	0	0	0	0.00
2009	9	8	1	9	0	0	0	0.00
2010	37	27	10	32	5	5	0	13.51
2011	38	33	5	36	2	1	1	5.26
2012	30	26	4	27	3	3	0	10.00
All Years	972	806	166	931	41	40	1	4.22
Panel B: Number of IPOs charging about 5% of the net proceeds (Between 4.7 and 5.3)								
1999	35	24	11	31	4	2	2	11.43
2000	132	80	52	114	18	5	13	13.64
2001	52	46	6	45	7	4	3	13.46
2002	39	30	9	36	3	0	3	7.69
2003	34	30	4	33	1	1	0	2.94
2004	137	122	15	113	24	19	5	17.52
2005	181	166	15	155	26	23	3	14.36
2006	131	117	14	110	21	17	4	16.03
2007	99	81	18	80	19	18	1	19.19
2008	18	16	2	15	3	2	1	16.67
2009	9	8	1	8	1	1	0	11.11
2010	37	27	10	27	10	9	1	27.03
2011	38	33	5	30	8	8	0	21.05
2012	30	26	4	26	4	3	1	13.33
All Years	972	806	166	823	149	112	37	15.33

This table has 2 panels. Panel A shows the number of IPOs per year and for each market. Then it shows how many of these IPOs have a spread of 7% and the percentage of underwriters charging 7% Panel B shows the number of IPOs per year and for each market. Then it shows how many of these IPOs have a spread of 5% as this is the median of our sample. Then it shows the percentage of underwriters charging 5%.

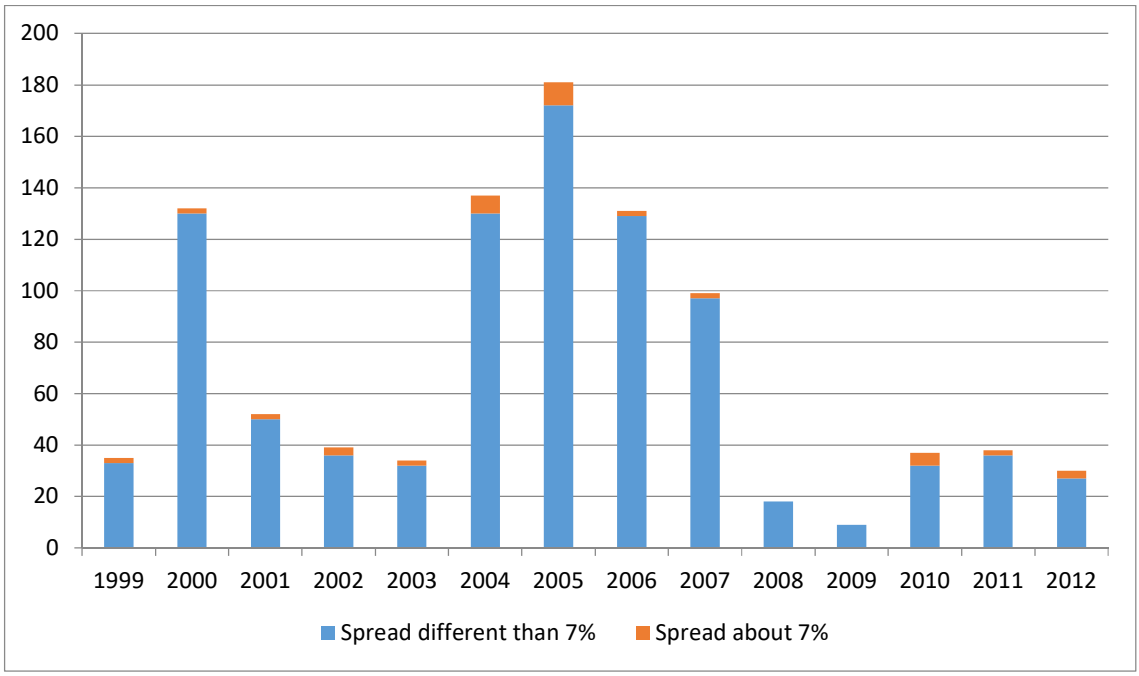

Figure 4.1: Showing number of IPOs charging about 7% of the net proceeds (Between 6.7 and 7.3)

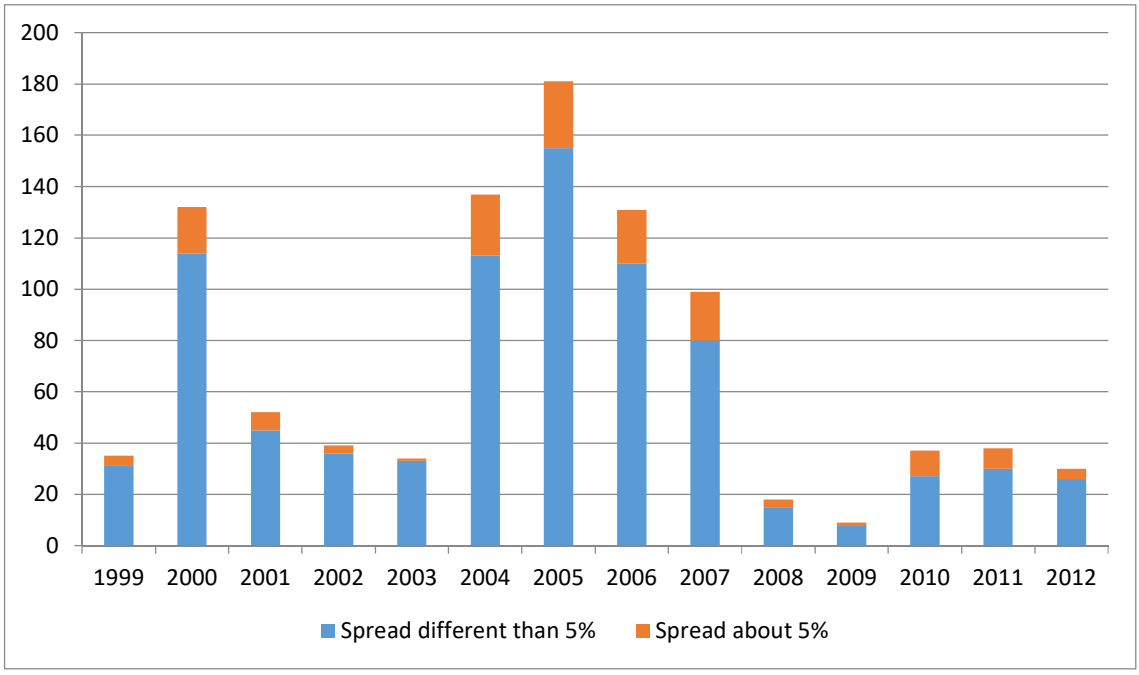

Figure 4.2: Showing number of IPOs charging about 5% of the net proceeds (Between 4.7 and 5.3)

As our sample showed that the median of the fees charged by the underwriter is 5%, I have reproduced the above table and figure for the fees of 5%. Table 4.8, Panel B shows the percentage of IPOs where the underwriters charged 5%. As you can see from the table, for most of the years except for 2002 and 2003, the IPOs with 5% fees charged are more than 10%. This percentage seems to be increasing over the years as it reached about 27% of the IPOs of 2010 it also makes about 15.33% of the whole sample compared to 4.22 of underwriters charging 7%. Figure 4.1 and Figure 4.2 show the same information in a bar chart.

This goes in line with the findings reported in the paper by (Abrahamson, et al., 2011) where they have reported that there is a concentration of a spread of 7% for the IPOs conducted in the US market. Yet, no such concentration is present in the European market.

Figure 4.3 shows that the charged fees have a downgrading slope as the value of the proceeds increase with most of the values at 5% or less. This is shown more clearly when I take the log of the proceeds as shown in Figure 4.4. As I took the log of the IPO proceeds in millions, and since there are a number of IPOs with proceeds less than £10 million, part of the values is shown in the negative part of the chart. I noticed that the value of the charged percentage is scattered over for Log (proceeds £M) that are less than 0.5. As I move along the X-axis towards Log (proceeds £M) of 1.5 the charged fees drop down to around 5%. For the area that comes after that, I can see that the charged fees drop even more to below 5%. These give support for the economies of scale hypothesis.

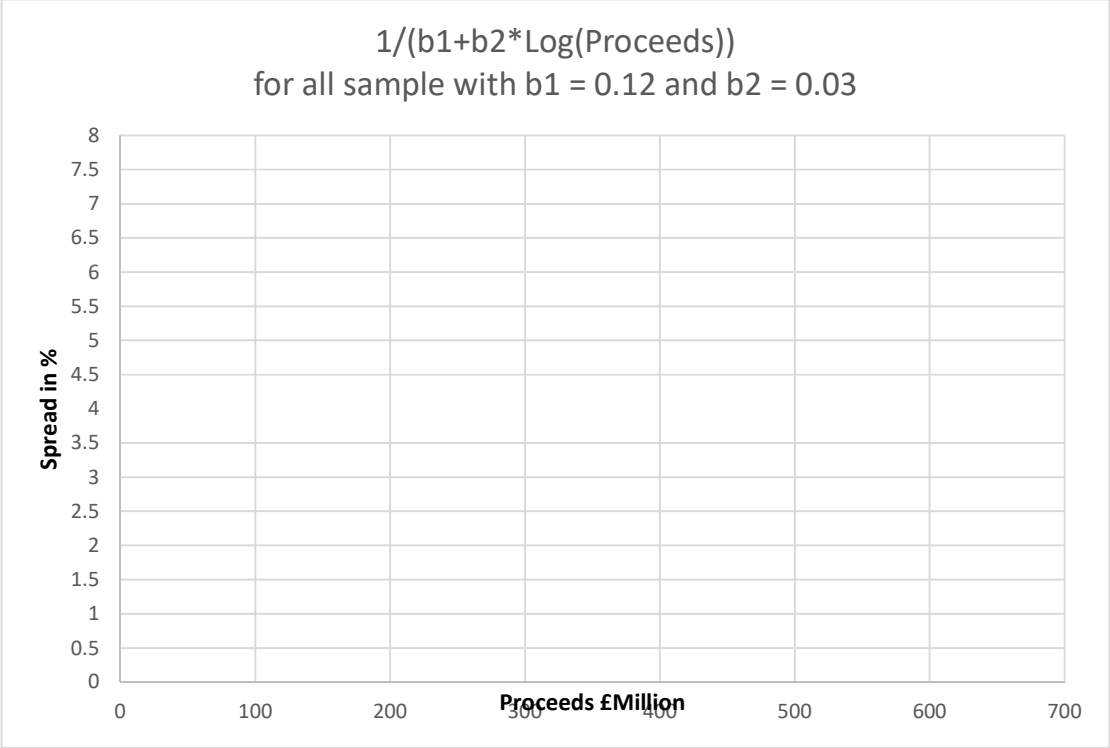

Figure 4.3: Curve Model showing the relation between Proceeds and the Spread

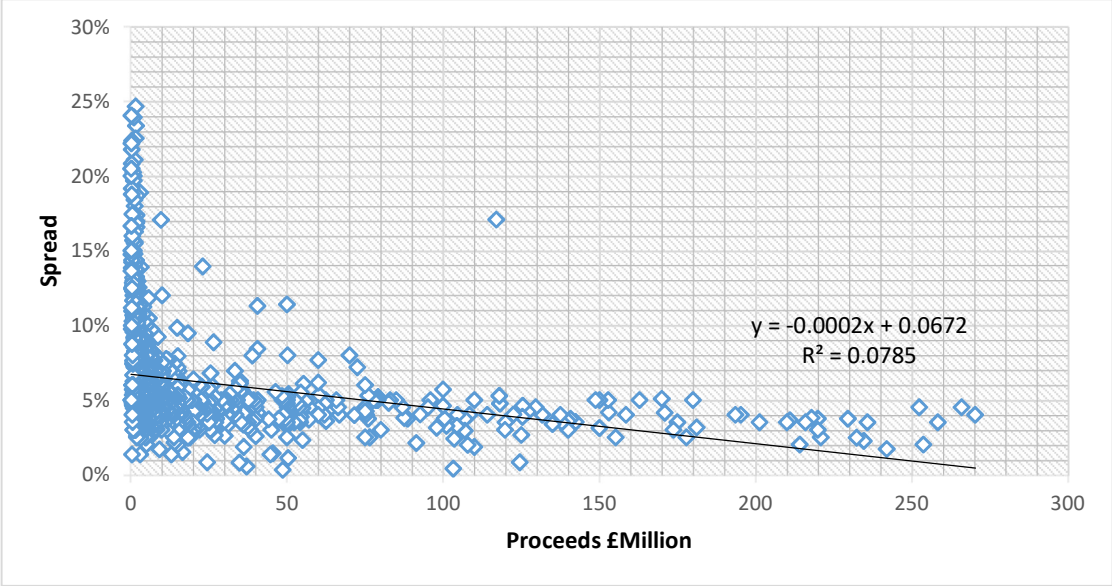

Figure 4.4: Spread percentage to proceeds (All Sample)

To further investigate these findings with the whole sample, I have generated the same figures for Main and AIM market. The same findings can be noticed in the Main and AIM markets. Figure 4.5 shows the scattering of spread charged versus proceeds in £Millions. Figure 4.6 for the Main market shows the same information but versus Log (proceeds £M). I notice that the spread is highly scattered at the lower values of log (proceeds £M) and then it starts sloping downwards until it reaches about 5% at log (proceeds £M) of 50. Then spread is slightly ranging below 5%.

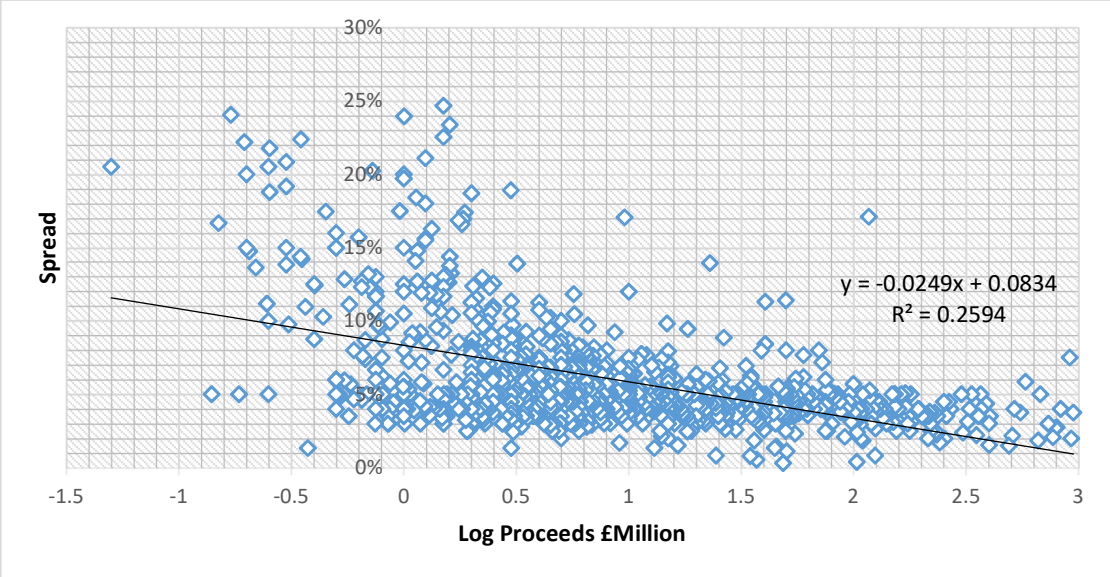

Figure 4.5: Spread percentage to log proceeds (All Sample)

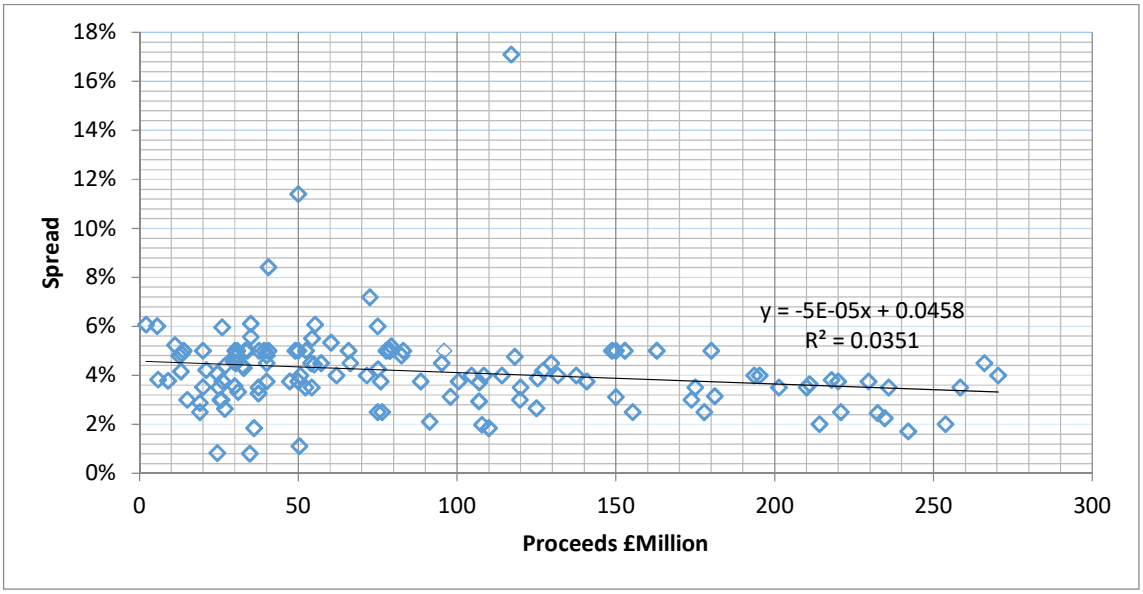

Figure 4.6: Spread percentage to proceeds (Main Market)

Figure 4.7 shows the information for AIM market with proceeds in £Millions. Figure 4.8 shows it in correspondence to log (proceeds £M). From this figure, dis-economies of scale are also prevalent here as the downward slope indicate a decrease of the spread charged as the proceed increases. I noticed that there a convergence towards 5% as the value of log (proceeds £M) increases. What is noticeable here is that the higher fees charged in the AIM are driven by the very small IPOs which raised money in the AIM

To a certain extent, these findings agree with the findings reported by Abrahamson et al., 2011. In their paper, they reported a clustering of the spread at 7% at the US market where no such clustering is present in European markets. From our sample, I have noticed that the spread scatter widely for the smaller IPOs, and then the spread charged start to converge towards 5% and falls below it in larger IPOs. Hence, I conclude that there is no clustering of fees for the UK IPOs as Abrahamson et al. (2011) shows no clustering for European IPOs.

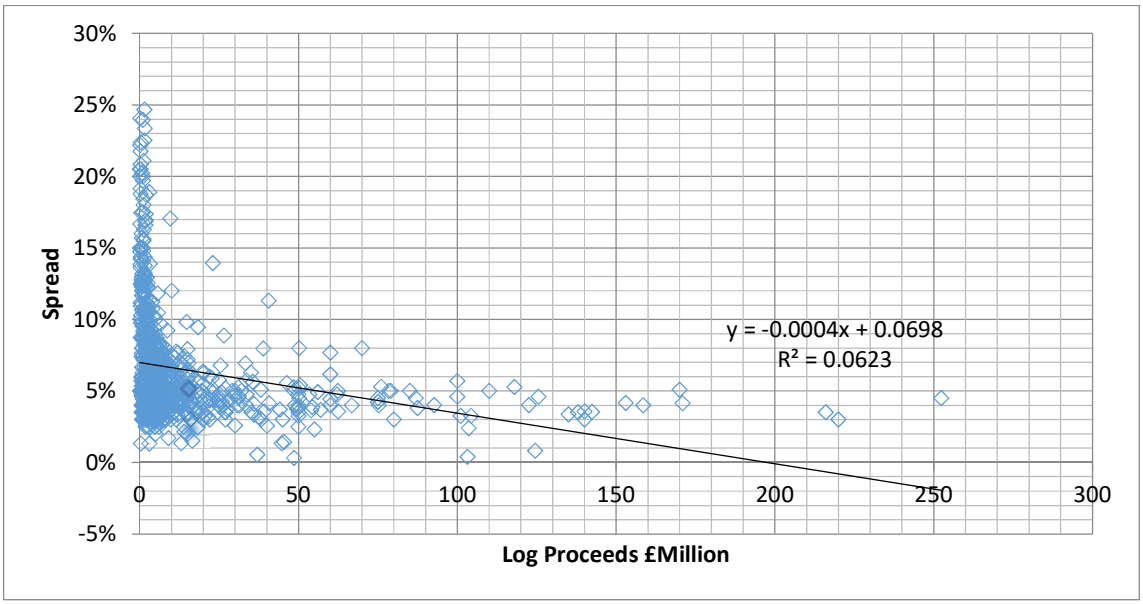

Figure 4.7: Spread percentage to log proceeds (Main Market)

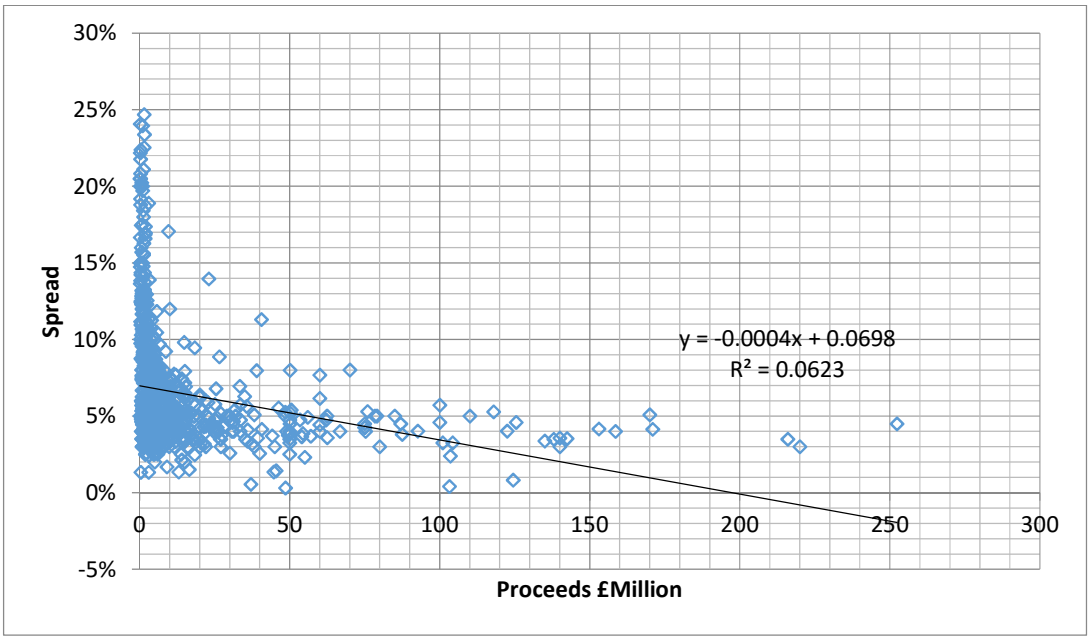

Figure 4.8: Spread percentage to proceeds (AIM Market)

4.14 Empirical Results

4.14.1.1 Correlation analysis and Univariate regressions

To examine whether there is any correlation among the variables, first, I conducted a correlation matrix and then I ran a series of univariate regressions. Table 4.9 presents the results of correlation analysis. It shows that there a high significance correlation between the spread and log of proceeds, market dummy, multi-bookrunners dummy, the number of bookrunners, prestigious underwriters and the log of the lockup period. Moreover, there is significant yet less than the aforementioned variables with the age of the company on the day of the IPO. I noticed the log(proceeds) has a correlation with the same variables as the spread. In addition to them, the table also shows a correlation with high significance between log(proceeds) and both the age of the company at the day of the IPO and log(Lockup). By looking at the first column, it shows that the log of proceeds is negative -0.499 and it is statistically significant. It means that higher the proceeds, the lower the spread. The market dummy is positively correlated with the spread, and the coefficient is 0.2425. Multi-bookrunner, the number of bookrunners, prestigious and age are negatively correlated, and idiosyncratic risk, return volatility, lockup length are positively correlated with the spread. Potential growth and underpricing are positively correlated, but the coefficient is very small. Another aspect of the correlation matrix is that it detects any problem of multicollinearity in the models. For example, multi-bookrunner and number of bookrunners are highly positively correlated, as expected, at 0.8189. This implies that I cannot include both of these variables (multi-bookrunners and number of bookrunners) in the same regression model; otherwise, it will create the problem of multicollinearity.

Table 4.9: Correlation Table

Variables	Spread	Log Proceeds	Market Dummy	Multi-bookrunner	Number of bookrunners	Prestigious	Idiosyncratic Risk	Return Volatility	Age at IPO in Days	Potential Growth	Log Lockup period	Underpricing
Spread	1											
Log Proceeds	-0.5***	1										
Market Dummy	0.243***	-0.603***	1									
Multi-bookrunner	-0.119***	0.397***	-0.333***	1								
Number of bookrunners	-0.127***	0.387***	-0.331***	0.819***	1							
Prestigious	-0.12***	0.37***	-0.37***	0.171***	0.168***	1						
Idiosyncratic Risk	0.02	-0.004	0.015	-0.01	-0.009	-0.031	1					
Return Volatility	0.028	-0.206***	0.039	-0.052	-0.05	-0.042	-0.006	1				
Age at IPO in Days	-0.072*	0.149***	-0.161***	0.064*	0.048	0.062	-0.063*	-0.063*	1			
Potential Growth	0.006	0.016	-0.04	0.022	0.023	0.035	-0.023	-0.008	0.009	1		
Log Lockup period	0.109***	-0.221***	0.249***	-0.283***	-0.236***	-0.115***	-0.001	-0.026	-0.071*	-0.05	1	
Underpricing	0.008	-0.172***	0.021	-0.102***	-0.091***	-0.052	-0.062*	0.261***	-0.057	0.012	-0.003	1

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation between the variables I am testing in this chapter. The spread is the amount charged by the underwriter as a percentage of the IPO proceeds. Log proceeds are the log of the proceeds raised from the IPO. The Market dummy takes a value of 0 for the Main market and 1 for the AIM market. Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. The number of bookrunners is the actual number of bookrunners in each IPO. Prestigious is a dummy with a value of 1 if the underwriter is prestigious and 0 if it is not. Idiosyncratic risk shows the firm specific risk. The return volatility is the standard deviation of the share return from the IPO date for a year. The Age at IPO shows the age of the company at the day of the IPO and is calculated in days. Potential growth is a factor showing if the company increased the size of its capital at the years following the IPO. Log lockup period where the lockup period is measured in days.

To examine the effect of the different variables on the spread, I have run a univariate regression for all the variables with the spread. I have run the regression using the whole sample, then for each market separately. A summary of the univariate regression is shown in Table 4.10.

In Table 4.10 I showed the univariate analysis for the whole sample in column A. Then I showed it for the Main market in column B and finally for AIM market in column C. I noticed that the log proceeds variable is negatively related to the spread charged and showing the highest significance when looking at the whole sample in column A and the AIM market in column C. They both show a negative coefficient. The variable that comes next in terms of significance when looking at the whole sample is the market dummy. It has a positive relationship with the spread. Looking at column C, the second variable in terms of significance is lockup period. The coefficient shows a positive relation with the spread. This means that the longer the lookup period, the more the cost charged.

Some of the variables showed insignificance across all three panels. Those are Age of the company, idiosyncratic risk, Return Volatility, Day 1 and Day 5 returns and potential growth. The univariate analysis of AIM market in panel C shows more significance than the Main market which means that the main market is less sensitive to those variables compared to the growing AIM market.

Table 4.10: Univariate OLS Regression with Spread

	A: Whole sample		B: Main Market		C: AIM Market	
Independent Variable	**Coefficient (t Statistic)**	**Adjusted R^2**	**Coefficient (t Statistic)**	**Adjusted R^2**	**Coefficient (t Statistic)**	**Adjusted R^2**
Log (Proceeds	-2.386***	0.2486	-0.582**	0.0302	-2.978***	0.2350
MCPI Adjusted	(-17.92)		(-2.45)		(-15.76)	
Market (Main or AIM)	-2.417***	0.0578				
	(-7.77)					
Prestigious	-0.893***	0.0134	0.134	-0.0056	-0.296	0.0001
	(-3.75)		(0.32)		(-1.05)	
Multi-bookrunner	-1.643***	0.0131	-0.606**	0.0188	-0.583	-0.0004
	(-3.72)		(-2.02)		(-0.82)	
Number of bookrunners	-1.08***	0.0150	-0.402***	0.0384	-0.621	-0.0002
	(-3.97)		(-2.73)		(-0.92)	
Rent-Seek	2.454***	0.0107	1.095	0.0078	1.195	0.1802
	(3.39)		(1.5)		(1.34)	
Age in Days	1E-4*	0.0038	-8.06E-6	-0.0071	-1.3E-4	0.0008
	(-1.92)		(-0.28)		(-1.2)	
St. Deviation	3E-5	-0.0008	-0.001	-0.0012	-3.7E-5	-0.0009
	(-0.52)		(-0.9)		(-0.53)	
Idiosyncratic	5E-5	-0.0009	-0.002	-0.0061	-6.6E-5	-0.0010
	(-0.31)		(-0.14)		(-0.42)	
Day 1 Return	-0.009	-0.0010	-0.015	-0.0060	-0.024	-0.0012
	(-0.17)		(-0.13)		(-0.41)	
Day 5 Return	0.006	-0.0010	-0.02	-0.0062	-0.006	-0.0012
	(0.17)		(-0.18)		(-0.18)	
Potential Growth	0.001	-0.0011	-0.007	-0.0050	0.003	-0.0010
	(0.18)		(-0.52)		(0.5)	

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows a summary of the univariate regression between the Spread charged by the underwriters in terms of a percentage of the proceeds and the above listed factors. The market factor is a dummy with 1 for Main market and 0 for AIM market. Prestigious is a dummy with 1 if a prestigious underwriter and 0 if a non-prestigious underwriter. Rent-seek is the log of lockup period in days. The STDev is the standard deviation of share price during the first year of trading. Idiosyncratic is calculated for the first year of trading. D5Return and D1Return are the return on the fifth and the first day of trading if purchased at the IPO. The potential growth is the Slope of the Market Capitalization of each company during the first 2 – 3 years of trading. Panel A, B and C shows the results for the whole sample, Main market and AIM market respectively.

4.14.1.2 Multiple-regression model

I now turn the attention to testing my hypotheses in a multivariate setting, where I test a number of hypotheses controlling for a number of factors. In this section, I will examine the relation between fees charged in terms of spread and a number of different variables. In the paper by Abrahamson et al., 2011, they used an OLS model to test the effect of a number of variables over the spread. The independent variables in their model are whether the company is from the high-tech sector, whether the company was backed by venture capitalists, whether privatisation was involved, whether the IPO involves multiple bookrunners, year dummies, and country dummies. In my study, I will use some of the variables used by Abrahamson et al., 2011 that apply to the market I am studying which are proceeds, Multi-bookrunners. Since UK market has two segments the Main market and the AIM, I use a market dummy (Main or AIM) instead of the country dummy and year dummy. In addition, I will control for the effects of the Prestigious Bookrunner, Number of bookrunners, and lockup period. I will run OLS regression for the following model:

$$\begin{aligned} Spread_i = {} & \alpha + \beta_1 Log\ Proceeds_i + \beta_2 Market_i \\ & + \beta_3 MultiBookrunner_i + \beta_4 Prestigious_i \\ & + \beta_5 D1\ Underpricing_i + Risk_i \\ & + \beta_6 LogLockupPeriod_i + \left(\sum_{j=1999}^{2012} \beta_j Year_j \right) + \varepsilon_i \end{aligned} \tag{4.1}$$

The first variable is the log of the proceeds. Besides being one of the variables used in previous studies, it is one of the main contributors to the spread charged, as it is the main goal for the bookrunners to maximise their return while minimising the risk. It shows the significance of the level of success anticipation by the investors. If the coefficient of log proceeds is negative, then, it means that there is a negative

relationship between the amount of proceed and the spread charged which will support the hypothesis of economies of scale. As I am comparing the Main market and AIM market, I have introduced the Market dummy where its value is 0 for the Main market and 1 for AIM market. I noticed earlier that the spread is higher in the AIM market compared to the Main market, I would expect the coefficient to have a positive relation with the spread which will support the economies of scale hypothesis too as the size of the proceeds is lower on the AIM market compared to the Main market. The level of confidence of the coefficient will be an indication of the difference in characteristics between both markets. Multi-bookrunners presence indicates that strong due diligence is in place. I have used a value of 1 when multi-bookrunners are present and 0 otherwise. The coefficient is expected to be positive as to have multi-bookrunners will involve more coordination and hence more cost. The prestigious variable will indicate the level of the cost of hiring a prestigious underwriter if the coefficient is positive. If negative, then this will indicate the level of competition between the prestigious companies. The underpricing variable will indicate the costs of finding the fair value of the share. It is expected to be negative, as the lack of effort to get a fair price would lead to a lower spread.

I have also examined the effect of rent-seeking on the spread in terms of the lockup period. I ran a regression between the log of the lockup period and the spread. This proved to have some significance with a correlation of 0.1085 and a Beta of 2.45. If the coefficient's sign is negative, this would indicate rent-seeking either from the company. It could indicate that company is offering higher compensation to the underwriters to reduce the lockup period. On the other hand, it could indicate that the underwriter is incentivising the company with lower spread to accept longer lockup period and hence longer engagement.

In our study, I have considered a number of other variables and examined their statistical significance in the model. To test whether underwriters charge more for in the AIM because the riskiness is higher in the AIM, I introduce two risk proxies (one at a time). In terms of risk, I have looked into two variables. I have considered the idiosyncratic risk and the volatility of the return on the share price. To identify possible risks that could have contributed to the underwriter fees, I identified the idiosyncratic risk for every share price over the first year after the IPO. I calculated the idiosyncratic risk as the standard error using the least square method between the share return on the issue price and the index price. Then I annualised the result. I then tested the correlation with Spread. With a correlation of 0.0198, Beta has shown insignificant, and the F test of the estimate is also insignificant, I concluded that the idiosyncratic risk is insignificant in our model and hence has no effect on fees charged by the underwriters. I also tested the volatility of the return on the share price as an indicator of a hidden risk factor by using the standard deviation of the share prices over the first year. The correlation was slightly higher with 0.0273, and both beta and F test for the estimate showed to be insignificant and hence this variable was not included too. Since both our risk factors proved insignificant and hence I cannot conclude that higher charges in AIM are due to higher risk in terms of daily volatility. Moreover, when I look at Table 4.10, I also noticed that the coefficient for both the Standard Deviation and Idiosyncratic are low.

I have introduced another variable that is the number of bookrunners. Since the previous studies show a negative relationship between the presence of multi-bookrunners and spread charged, I have decided to examine this variable even further by looking at the number of bookrunners. I will observe whether the coefficient will follow the last coefficient or not. If it does, then this could mean that with more

bookrunners there will be more complication and then more cost. If otherwise, then it could indicate that cooperation leads to more efficiency. However, with a correlation of 0.82 between the multi bookrunner dummy and the number of bookrunners, in addition to the fact the model presented better results with the multi-bookrunner dummy, I have decided not to use the number of bookrunners variable.

In this chapter, I have identified a number of variables. I extracted some of the variables directly from the above-mentioned data sources. Besides that, I have estimated three other variables that are an idiosyncratic risk, return risk, and potential growth.

The idiosyncratic risk is the risk associated directly with a particular company in the separation of the market systematic risk. To calculate the idiosyncratic risk, I started by calculating the daily return on the market index as an indicator of the systematic risk. I have used "FTSE All Share" Index for the main market and "FTSE AIM All Share" index for the AIM market. Then I calculated the daily return on the share price starting from the 1st trading day for 1 year. I then ran a regression between the two series as shown in the following equation:

$$Return_t = \alpha + \beta IndexReturn_t + \varepsilon_t \quad (4.2)$$

The difference between the index return and company return is captured in the standard error. The average daily idiosyncratic risk is estimated using the standard error. To annualise, I multiply it by the square root of 260. Some of the companies stopped trading in less than 260 days. In this case, I multiply it by the square root of the number of observations.

To estimate the return risk, I measured the volatility of the return of the share for the first trading year. As the volatility of the return is captured in the standard

deviation of the return sample, I used the standard deviation as the estimate of the return risk. I have estimated the potential growth of the company at the time of the IPO. I used the size of capital over the 3 years following the IPO. I calculated the estimate of the capital size over the time and used the slope β as the indicator for the potential growth.

To examine the spread charged by the underwriters and how a different number of variables would affect it. To do this, I have identified a number of independent variables, and then I used the ordinary least square model regression (OLS) to measure the level of how would they affect. I also used the OLS model to examine the variables that would affect the underpricing. The underpricing that I have calculated is the return of a share at the end of the 1st trading day if purchased at the IPO price. This shows whether the price was overpriced, underpriced or fairly priced.

Another factor that I considered to proxy for risk is to include in the spread model the company's age at the time of the IPO. I have calculated the age of the company in days starting from the incorporation date until the day of trading. I ran the regression for this variable over 715 IPO from our sample where the data were available. This also proved insignificant. There was a correlation of -0.072, and beta of-1.15×10^{-6} with Spread, so it was not included in the model (Results are not reported to save space).

In contrast to the risk and age of company where both showed low significance, I examined the possibility of potential growth of the company. Potential growth would be expected to have a negative relationship with the fees charged by the underwriter. I measured the potential growth of the companies by observing the change in their market capitalization over 2 – 3 years following the IPO. I used the slope of the log as

the potential growth. I ran a simple linear regression between the spread and potential growth variable and got a correlation of 0.005 and a Beta of 7.94×10^{-6} which also proved insignificant and hence did not include it in the model.

I also examined the return on the first and the fifth day of trading to check the effect of the over/underpricing on the fees charged and this too proved insignificant as univariate. However, I have included the return of first trading day or on other words, Underpricing. This is calculated as the return of stock having been bought at the IPO using the issue price and being sold at the end of the 1st/5th trading day after it becomes public.

Table 4.11: Spread Regression for different sizes of IPO proceeds and for AIM and Main

		Whole sample						
		All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
Control Variables	Log Proceeds	-3.183***	-4.553***	-0.362	-1.642	1.601	-0.389	-3.588***
		(-17.83)	(-15.88)	(-0.41)	(-1.59)	(0.76)	(-1.38)	(-17.8)
	Multi-bookrunner	0.584	0.957	0.538	0.012	1.045	-0.3	1.296**
		(1.38)	(1.05)	(1.08)	(0.03)	(0.86)	(-0.82)	(2.08)
	Prestigious	0.793***	1.076***	-0.124	0.599	2.251	0.259	0.963***
		(3.58)	(3.88)	(-0.37)	(1.34)	(0.89)	(0.62)	(3.97)
	Underpricing D1	-0.004**	-0.005**	-0.002	-0.006	0.183	-0.001	-0.005**
		(-2.15)	(-2.13)	(-0.6)	(-1.21)	(1.25)	(-0.33)	(-2.15)
	Log(Lockup)	0.681	1.7**	-0.227	0.095	-6.192	0.706	1.223
		(1.06)	(2.03)	(-0.23)	(0.09)	(-1.19)	(0.86)	(1.62)
	Market Dummy	-1.69***	-0.487	-0.213	-0.344			
		(-4.68)	(-0.52)	(-0.53)	(-0.82)			
Year Dummies	2000	0.674	0.675	0.718	0.766	-0.299	0.408	0.718
		(1.15)	(0.91)	(0.66)	(0.78)	(-0.13)	(0.73)	(0.96)
	2001	1.702**	1.856**	1.41	1.901	0.413	0.53	2.08**
		(2.54)	(2.32)	(0.85)	(1.06)	(0.16)	(0.64)	(2.59)
	2002	1.541**	1.754**	1.284	0.368	-0.746	0.372	1.979**
		(2.17)	(2.02)	(0.94)	(0.26)	(-0.34)	(0.51)	(2.27)
	2003	3.943***	4.301***	0.773		0.071	0.333	4.572***
		(5.35)	(4.85)	(0.47)		(0.03)	(0.34)	(5.24)
	2004	2.119***	2.216***	1.595	0.573		0.62	2.591***
		(3.64)	(3.07)	(1.39)	(0.55)		(0.91)	(3.64)
	2005	2.262***	2.451***	1.449	0.546	0.923	0.331	2.769***
		(3.97)	(3.47)	(1.25)	(0.56)	(0.49)	(0.5)	(3.97)
	2006	2.992***	3.406***	1.339	1.416	-1.062	0.646	3.586***
		(5.08)	(4.64)	(1.14)	(1.42)	(-0.55)	(0.96)	(4.99)
	2007	2.756***	3.36***	1.027	0.92	1.662	0.504	3.493***
		(4.51)	(4.21)	(0.9)	(0.94)	(0.68)	(0.8)	(4.65)
	2008	3.145***	2.882**	1.307	1.353	7.331*	2.741**	3.372***
		(3.51)	(2.33)	(0.92)	(0.99)	(2.31)	(2.16)	(3.24)
	2009	2.841**	2.735	0.492	0.994		0.497	3.727***
		(2.47)	(1.35)	(0.29)	(0.82)		(0.29)	(2.85)
	2010	3.03***	3.986***	1.193	0.396	1.88	-0.093	3.946***
		(4.06)	(3.43)	(0.97)	(0.37)	(0.87)	(-0.12)	(4.31)
	2011	3.831***	4.425***	1.724	1.544	-1.632	0.047	4.404***
		(5.27)	(4.87)	(1.22)	(1.16)	(-0.68)	(0.05)	(5.12)
	2012	5.111***	5.762***	0.803	8.482***		6.394***	5.584***
		(6.5)	(5.99)	(0.52)	(6.22)		(5.14)	(6.11)
Constant		5.741***	2.039	4.612	5.724	9.707	2.415	2.345
		(3.24)	(0.82)	(1.49)	(1.62)	(0.84)	(0.98)	(1.13)
	Adjusted	0.3295	0.3251	-0.0729	0.3296	-0.1473	0.1345	0.3286
	Observations	967	686	169	92	20	161	806
	Prob > F	0	0	0.9885	0.0001	0.6465	0.0024	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Regressions where each one is presented in one of the columns. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The fifth regression with proceeds larger than £333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.
The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(lockup) is the log of the lockup period, and Idiosyncratic shows the idiosyncratic risk for the shares in the sample. Market dummy is 0 for Main market and 1 for AIM market.

The F test in Table 4.11 indicates high significance for the whole sample, for IPOs with less than 16.7M and AIM market. Taking into consideration that the number of IPOs from the AIM market makes more than 70% of our sample, I can conclude that the level of significance is mainly coming from the AIM market which indicates a difference in the characteristics between the two markets. Most of the variables tested showed a high level of significance in those aforementioned segments except for the two bookrunners variable. Multi-bookrunner is showing to be significant on AIM market. This is because multi-bookrunners are seen on most of the larger IPOs rather than those in AIM market. The F test for the Main market shows a significance of more than 95%. Yet, Adjusted R^2 is 0.14 and none of the variables of the model shows high significance. Furthermore, F test shows that the model is insignificant for IPOs with proceeds between 16.7M - 66.7M and Greater than 333.4M.

When I look at 66.7M to 333.4M segment, I can notice that Adjusted R^2 is 0.33. This segment has the top IPOs from the AIM market and the bottom IPOs from the Main market with more than third of the IPOs are from the AIM market. This can be considered as emerging companies that have reached a level of maturity to be part of the Main market yet it has a potential to grow there or at the top of AIM market and has the chance to migrate to the main market. From Table 4.11, I can notice that the Log proceeds has about 90% significance. The negative relationship between log proceeds and the spread indicates that the higher the proceeds, the lower the spread and hence I conclude the presence of dis-economies of scale. Moreover, the coefficient of the log proceeds variable is higher – as an absolute value – for the Main market and larger IPO segments, which support the presence of the dis-economies of scale too.

Moreover, the number of bookrunners variable is showing significance at 95 level. This could indicate that more bookrunners are interested in this segment of the market. Multi-bookrunners were in more than 33% of the IPOs in this segment compared to less than 7% of the whole sample. In addition, the Multi-bookrunner variable has high significance only when looking at the AIM market. However, the significance is proven to be poor otherwise.

The coefficient of the prestige variable is of high significant for the whole sample, less than 16.7M and AIM market. It is positively related to the spread that means that prestigious underwriters charge more than their non-prestigious counterparts do. However, as this coefficient is significant in the AIM market and the smaller IPO segment, this shows that for smaller companies, they have to pay more to make their IPO more appealing to the prestigious underwriters. In addition, not being significant for the Main market and the larger IPOs shows that underwriters do not compete on cost in the competition for larger IPOs.

Underpricing variables show a negative relationship with the proceeds with small coefficients compared to the other variables. This means that the more the IPO is underpriced, the less the spread charged however with a small multiplier. Therefore, I conclude that underwriters mitigate risk to some extent by lowering the issue price. Thus I accept H3(b). The lockup period variable is showing significance only for IPOs with proceeds less than 16.7M. This variable too has a positive relationship with the spread. This indicates that companies with smaller IPO will have to pay relatively more to prolong the underwriters' engagement and hence no presence of rent-seeking. As the lookup is used as a tool to mitigate moral hazard and to compensate for the information asymmetry, the longer lockup indicates more risk and hence higher

spread, especially for smaller IPOs. Thus I reject the null of H4 and conclude that there is some evidence of rent-seeking.

Table 4.11 shows that the year dummies have a high significance on the value of the fees charged. By reviewing the coefficients of the years' dummies, I can notice that the fees charged have been increasing since 1999 (the year 1999 is not shown in the table as it is the reference variable for the years' dummies).

As a result, the model is a better representation of the AIM market than the Main market. It shows that the Main market is less sensitive to the factors that attribute to the spread charged at the AIM market. They AIM IPOs at our sample make about 71% of the whole sample under study. This complies with the aforementioned results about smaller size IPO as most of the IPOs that took place in the AIM market fall under the category of less than 16.7M.

4.15 Spread as a marketing tool

With lower cost, the demand is expected to increase. Here I will examine our fifth hypothesis to see if the underwriters charge lower spread in order to attract more business. I ran the following two regression models:

$$\begin{aligned} No.of\ IPOs_t = {} & \alpha + \beta_1 No.of\ IPOs_{t-1} \\ & + \beta_2 Log\ Average\ Proceeds_{t-1} \\ & + \beta_3 Average\ Spread_{t-1} + \varepsilon_i \end{aligned} \tag{4.3}$$

$$
\begin{aligned}
&Average\ Spread_t \\
&\quad = \alpha + \beta_1 No.of\ IPOs_{t-1} \\
&\quad + \beta_2 Log\ Average\ Proceeds_{t-1} \\
&\quad + \beta_3 Average\ Spread_{t-1} + \varepsilon_i
\end{aligned}
\tag{4.4}
$$

The first model is an estimator of the number of IPOs per underwriter based on the number of IPO, the log of the average proceeds and the average spread charged in the previous year. I first ran a pooled simple OLS regression for the whole sample then I applied the fixed effects model for the bookrunners. I have checked tested the model for multicollinearity and it is not affecting the results. As you can see from Table 4.11, the number of IPOs conducted in the previous year has a positive relationship with the IPOs conducted the current year. This means the more IPOs a bookrunner can undertake in a year, the more the number of IPOs in the following year. However, the effect of the number of IPOs in a year has a significance, yet, a minimal effect on the average spread charged on the following year. The average proceeds coefficient is insignificant except for the number of IPOs when applying the fixed effects model. The average spread charged from the previous year has a positive relationship with the number of IPOs undertaken in the following year. This means that the spread charged is not used as a marketing tool by the underwriters and hence I will reject hypotheses H5. The positive relationship could be explained by the fact that the book-runners are emerging to be more reputable facing less competition and would charge higher spread. (Robins, 2012) have reported that underwriters that are more reputable charge higher spread on the long-run. This could give an explanation for our findings.

Table 4.12: The effect of Year t-1 level of fees on Year t

		No. Of IPOs (t)	Average Spread (t)	No. of IPOs (t) (FE)	Average Spread (t) (FE)
Control Variables	No. of IPOs (t-1)	0.408***	0.005***	0.236***	0.004***
	Average Spread (t-1)	6.083***	0.267***	7.292***	0.132***
	Log Average Proceeds (t-1)	6.083	0.000	-0.052**	0.000
Constant		0.238***	0.008***	0.376***	1.085***
Adjusted R^2 / Overall R		0.240	0.183	0.233	0.1808
Observations (Groups)		1950	1950	1950 (150)	1950 (150)
Prob > F		0.0000	0.0000	0.0000	0.0000

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the effect of a number of factors in a previous year over the number of IPOs and the Average Spread charged on the consequence year. The factors of the previous year are the number of IPOs, Average Spread and Log Average Proceeds. The factors of the last year are denoted by t-1. The tables show 2 types of regressions; the first 2 columns are pooled regression while the last 2 columns are using Fixed Effects for the bookrunners. Adjusted R^2 is the standard value for the pooled regression and the overall value for the Fixed Effects regression. Observations shows the groups for the Fixed Effects regression. Prob > 0 indicates the significance of the model.

4.16 Robustness Checks

I have identified 115 underwriters in the model. Some underwriters may charge more than other underwriters. In other words, our OLS model is unable to capture the heterogeneity existing at the investment bank level. Hence, I reiterate our models using the fixed effect specifications to examine whether the results of OLS still hold.

$$\begin{aligned} Spread_i = {} & \alpha + \beta_1 Log\ Proceeds_i + \beta_2 Market_i \\ & + \beta_3 MultiBookrunner_i + \beta_4 Prestigious_i \\ & + \beta_5 D1\ Underpricing_i + Risk_i \\ & + \beta_6 LogLockupPeriod_i \left(\sum_{j=1999}^{2012} \beta_j Year_j \right) + \varepsilon_i \end{aligned} \tag{4.5}$$

To put more emphasis on the bookrunners and to capture its specific effect, I have repeated the regression of the same model using Fixed Effects model for the bookrunners. The results are shown in Table 4.12. The results are similar to those from Table 4.11.

Table 4.13: Spread Regression for different sizes of IPO proceeds and for AIM and Main Using Fixed-Effects for Bookrunners

		Whole sample					
		All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Main	AIM
Control Variables	Log Proceeds	-3.458***	-5.123***	-1.239	-2.31**	-0.703*	-3.957***
		(-15.23)	(-14.59)	(-1.04)	(-2.69)	(-1.79)	(-15.15)
	Multi-bookrunner	0.356	0.921	0.62	0.432	-0.44	1.21*
		(0.76)	(0.84)	(0.93)	(1.21)	(-1)	(1.67)
	Prestigious	-0.18			0.482	0.11	
		(-0.08)			(0.56)	(0.06)	
	Underpricing D1	-0.003	-0.003	-0.002	-0.003	0.001	-0.004
		(-1.4)	(-1.06)	(-0.68)	(-0.68)	(0.27)	(-1.39)
	Log(Lockup)	1.601**	3.233***	-0.463	0.112	1.506	2.196**
		(1.97)	(2.85)	(-0.34)	(0.11)	(1.5)	(2.15)
	Market Dummy	-1.144***	0.387	-0.551	0.217		
		(-2.61)	(0.35)	(-0.95)	(0.59)		
Year Dummies	2000	0.242	0.363	1.832	0.773	0.207	0.12
		(0.37)	(0.43)	(0.86)	(0.71)	(0.3)	(0.14)
	2001	1.158	1.508*	2.446	1.446	0.481	1.408
		(1.54)	(1.67)	(0.96)	(0.96)	(0.51)	(1.51)
	2002	0.655	0.821	2.305	-0.462	0.151	0.897
		(0.84)	(0.86)	(1)	(-0.36)	(0.17)	(0.92)
	2003	3.06***	3.604***	1.181		0.774	3.533***
		(3.77)	(3.65)	(0.45)		(0.71)	(3.55)
	2004	1.592**	1.804**	2.548	1.065	0.72	1.961**
		(2.38)	(2.16)	(1.2)	(0.95)	(0.89)	(2.32)
	2005	1.689**	2.3***	2.378	0.009	-0.165	2.221***
		(2.53)	(2.76)	(1.1)	(0.01)	(-0.21)	(2.63)
	2006	2.078***	2.709***	2.775	1.002	0.467	2.624***
		(3.05)	(3.16)	(1.25)	(0.92)	(0.59)	(3.04)
	2007	2.116***	3.297***	2.062	0.546	0.064	2.909***
		(2.96)	(3.46)	(0.94)	(0.47)	(0.08)	(3.21)
	2008	2.071**	1.856	2.969	2.486	2.347	2.227*
		(2.08)	(1.28)	(1.17)	(1.64)	(1.62)	(1.87)
	2009	1.392	3.845	0.772	0.592	-0.069	2.714
		(1.02)	(1.32)	(0.25)	(0.46)	(-0.04)	(1.57)
	2010	2.078**	4.418***	1.883	-0.39	-0.723	2.963***
		(2.41)	(3.18)	(0.83)	(-0.34)	(-0.74)	(2.7)
	2011	3.611***	5.231***	2.349	0.805	-0.378	4.403***
		(4.28)	(4.81)	(0.99)	(0.63)	(-0.34)	(4.27)
	2012	2.738***	3.403***	2.289	14.048***	6.166***	3.128***
		(3)	(2.97)	(0.92)	(8.77)	(4.37)	(2.84)
Constant		4.276*	-1.829	5.636	6.986**	1.401	1.137
		(1.75)	(-0.57)	(1.29)	(2.34)	(0.4)	(0.41)
Adjusted		0.31	0.3077	0.025	0.4273	0.2	0.3169
Observations		967	686	169	92	161	806
Prob > F		0	0	0.991	0	0.0043	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 6 Regressions using Fixed Effects for the bookrunners where each one is presented in one of the columns. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.
The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of the lockup period, Idiosyncratic shows the idiosyncratic risk for the shares in the sample. Market dummy is 0 for Main market and 1 for AIM market.

However, the coefficient of the Log Proceeds has increased in absolute value. So I accept the economies of scale hypothesis. Looking into the underwriter status variables; I can notice that prestigious variable shows less significance in this table. With respect to the bookrunner variables, it shows more significance than in Table 4.11 for the same segment (66.7M – 333.4M). This time they both are significant. In summary, I find that log of proceeds is negatively related, lockup length is positively related, and the market dummy is positively related when using the fixed effects model. This implies that economics of scale, rent-seeking are valid explanations of higher spread charged and AIM companies are riskier.

Table 4.13 shows the same results as above after I have winsorized the spread variable with a 5% from both ends in order to eliminate the effect of outliers. In this process, I look into the values of the spread. The values that are above the 95th percentile will be adjusted to be the same as the 95th percentile and the data that are below the 5th percentile will be adjusted to be the same as the 5th percentile. I noticed from the table that the results did not change much from the original results in Table 4.10. From the above tables, lockup period variable when proved significant has a positive relationship with the spread charged. This can be noticed in Table 4.11 for IPOs less than £16.7 Million, and AIM market. In Table 4.13 it can be noticed for the whole sample, Less than £16.7 Million and AIM market. Therefore, I can conclude that lockup period variable contributes to the higher spread in AIM market rather in the Main market and hence the increase in the cost for the smaller IPOs reflects the higher cost of a longer engagement from the underwriter and not due to the rent-seeking.

Table 4.14: Spread Regression for different sizes of IPO proceeds and for AIM and Main Using 5% Winsor for the spread

		Whole sample						
		All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
Control Variables	Log Proceeds	-2.574***	-3.526***	-0.32	-1.521*	1.177	-0.436*	-2.574***
		(-18.07)	(-15.53)	(-0.39)	(-1.86)	(0.71)	(-1.78)	(-18.07)
	Multi-bookrunner	0.378	0.746	0.447	0.018	0.784	-0.262	0.378
		(1.12)	(1.03)	(0.97)	(0.05)	(0.82)	(-0.83)	(1.12)
	Prestigious	0.538***	0.78***	-0.221	0.397	1.982	0.195	0.538***
		(3.04)	(3.56)	(-0.71)	(1.12)	(0.99)	(0.54)	(3.04)
	Underpricing D1	-0.004**	-0.005**	-0.002	-0.005	0.148	-0.001	-0.004**
		(-2.35)	(-2.3)	(-0.65)	(-1.17)	(1.29)	(-0.4)	(-2.35)
	Log(Lockup)	0.366	1.147*	-0.462	-0.011	-5.156	0.358	0.366
		(0.72)	(1.73)	(-0.49)	(-0.01)	(-1.26)	(0.5)	(0.72)
	Market Dummy	-1.255***	-0.213	-0.166	-0.177			-1.255***
		(-4.35)	(-0.29)	(-0.45)	(-0.53)			(-4.35)
Year Dummies	2000	0.567	0.484	0.889	0.57	-0.222	0.425	0.567
		(1.21)	(0.82)	(0.88)	(0.73)	(-0.12)	(0.87)	(1.21)
	2001	1.479***	1.598**	1.401	1.655	0.553	0.474	1.479***
		(2.76)	(2.52)	(0.91)	(1.16)	(0.27)	(0.66)	(2.76)
	2002	1.573***	1.82***	1.315	0.119	-0.692	0.283	1.573***
		(2.77)	(2.64)	(1.04)	(0.11)	(-0.41)	(0.44)	(2.77)
	2003	2.917***	3.182***	0.76		0.153	0.28	2.917***
		(4.96)	(4.53)	(0.5)		(0.09)	(0.33)	(4.96)
	2004	1.783***	1.862***	1.56	0.362		0.5	1.783***
		(3.84)	(3.25)	(1.46)	(0.44)		(0.85)	(3.84)
	2005	1.882***	2.026***	1.468	0.375	1.054	0.315	1.882***
		(4.14)	(3.62)	(1.37)	(0.48)	(0.71)	(0.55)	(4.14)
	2006	2.624***	3.006***	1.297	1.123	-0.91	0.593	2.624***
		(5.58)	(5.17)	(1.19)	(1.42)	(-0.6)	(1.01)	(5.58)
	2007	2.36***	2.91***	1.065	0.672	1.481	0.47	2.36***
		(4.84)	(4.61)	(1)	(0.86)	(0.76)	(0.86)	(4.84)
	2008	2.715***	2.55***	1.215	0.988	6.683*	2.726**	2.715***
		(3.79)	(2.6)	(0.92)	(0.91)	(2.68)	(2.47)	(3.79)
	2009	2.448***	2.944*	0.526	0.579		0.408	2.448***
		(2.67)	(1.84)	(0.34)	(0.6)		(0.28)	(2.67)
	2010	2.66***	3.557***	1.247	0.282	2.007	0.07	2.66***
		(4.47)	(3.87)	(1.09)	(0.33)	(1.18)	(0.1)	(4.47)
	2011	3.381***	3.846***	1.66	1.487	-0.739	0.49	3.381***
		(5.82)	(5.35)	(1.27)	(1.4)	(-0.39)	(0.57)	(5.82)
	2012	3.856***	4.307***	0.728	6.517***		4.697***	3.856***
		(6.14)	(5.65)	(0.5)	(6)		(4.34)	(6.14)
Constant		5.982***	2.975	-0.32	-1.53*	1.18	-0.44*	-2.89***
		(4.22)	(1.52)	(-0.39)	(-1.86)	-0.71	(-1.78)	(-18.02)
Adjusted		0.3408	0.3234	-0.0744	0.3236	0.0189	0.0966	0.3349
Observations		967	686	169	92	20	161	806
Prob > F		0	0	0.9903	0.0001	0.5504	0.0163	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Regressions where each one is presented in one of the columns. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The fifth regression with proceeds larger than £333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.
The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(lockup) is the log of the lockup period, Idiosyncratic shows the idiosyncratic risk for the shares in the sample. Market dummy is 0 for Main market and 1 for AIM market.

In order to examine the possibility of spread being bounded either from the top or from the bottom, I ran a Tobit regression. If the underwriters were charging different from the fair price due to some imposed constraints or limitation, Tobit regression would have come with the fair value. Tobit regression did not indicate any type of censoring. The results from Tobit and OLS are qualitatively the same.

Finally, as Figure 4.4, Figure 4.6 and Figure 4.8 are showing the spread in relationship to the log of proceeds variable have a nonlinear relationship, I have tested a cubed regression using the log of proceeds.

$$Spread = \alpha + \beta_1 \log(proceeds) + \beta_2(\log(proceeds))^2 + \beta_3(\log(proceeds))^3 + \varepsilon_i \quad (4.6)$$

When I used the whole sample, the coefficients $\beta_1, \beta_2, \beta_3, \alpha$ were -5.20, 2.34, -0.38, and 8.25 respectively. The value of F is zero and the Adjusted R^2 is 0.3178. This shows that proceeds only is the main factor that affects the spread. To further investigate and understand which part of our sample is mostly affected by the spread, I have repeated the regression the different markets and proceeds sizes. I found that the model is significant mainly for AIM market and IPOs less than 16.7M. This conclude that proceeds are the main factor in determining the spread for small IPOs while has less significance for larger IPOs.

Table 4.15: Spread Regression for different sizes of IPO proceeds and for AIM and Main Using Tobit Model

		Whole sample						
		All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
Control Variables	Log Proceeds	-3.183*** (-18.02)	-4.553*** (-16.11)	-0.362 (-0.44)	-1.642* (-1.79)	1.601 (1.7)	-0.389 (-1.47)	-3.588*** (-18.01)
	Multi-bookrunner	0.584 (1.39)	0.957 (1.06)	0.538 (1.15)	0.012 (0.03)	1.045 (1.93)	-0.3 (-0.87)	1.296** (2.11)
	Prestigious	0.793*** (3.62)	1.076*** (3.94)	-0.124 (-0.39)	0.599 (1.51)	2.251 (1.98)	0.259 (0.66)	0.963*** (4.01)
	Underpricing D1	-0.004** (-2.18)	-0.005** (-2.16)	-0.002 (-0.64)	-0.006 (-1.36)	0.183** (2.81)	-0.001 (-0.35)	-0.005** (-2.18)
	Log(Lockup)	0.681 (1.08)	1.7** (2.06)	-0.227 (-0.24)	0.095 (0.11)	-6.192** (-2.67)	0.706 (0.92)	1.223 (1.64)
	Market Dummy	-1.690*** (-4.73)	-0.487 (-0.53)	-0.213 (-0.57)	-0.344 (-0.92)			
Year Dummies	2000	0.674 (1.16)	0.675 (0.92)	0.718 (0.7)	0.766 (0.88)	-0.299 (-0.28)	0.408 (0.78)	0.718 (0.98)
	2001	1.702** (2.56)	1.856** (2.36)	1.41 (0.91)	1.901 (1.19)	0.413 (0.36)	0.53 (0.68)	2.080*** (2.63)
	2002	1.541** (2.19)	1.754** (2.05)	1.284 (1)	0.368 (0.3)	-0.746 (-0.77)	0.372 (0.54)	1.979** (2.29)
	2003	3.943*** (5.41)	4.301*** (4.92)	0.773 (0.5)		0.071 (0.07)	0.333 (0.37)	4.572*** (5.3)
	2004	2.119*** (3.68)	2.216*** (3.11)	1.595 (1.48)	0.573 (0.62)		0.62 (0.97)	2.591*** (3.68)
	2005	2.262*** (4.01)	2.451*** (3.52)	1.449 (1.34)	0.546 (0.63)	0.923 (1.09)	0.331 (0.53)	2.769*** (4.02)
	2006	2.992*** (5.14)	3.406*** (4.71)	1.339 (1.21)	1.416 (1.59)	-1.062 (-1.24)	0.646 (1.02)	3.586*** (5.05)
	2007	2.756*** (4.56)	3.36*** (4.27)	1.027 (0.95)	0.92 (1.05)	1.662 (1.51)	0.504 (0.86)	3.497*** (4.71)
	2008	3.145*** (3.54)	2.882** (2.36)	1.307 (0.98)	1.353 (1.11)	7.331*** (5.17)	2.741** (2.3)	3.372*** (3.28)
	2009	2.841** (2.5)	2.735 (1.37)	0.492 (0.31)	0.994 (0.93)		0.497 (0.31)	3.727*** (2.88)
	2010	3.03*** (4.1)	3.986*** (3.48)	1.193 (1.03)	0.396 (0.42)	1.88 (1.95)	-0.093 (-0.13)	3.946*** (4.37)
	2011	3.831*** (5.32)	4.425*** (4.94)	1.724 (1.3)	1.544 (1.3)	-1.632 (-1.53)	0.047 (0.05)	4.404*** (5.18)
	2012	5.111*** (6.57)	5.762*** (6.08)	0.803 (0.55)	8.482*** (6.98)		6.394*** (5.48)	5.584*** (6.18)
Constant		5.741*** (3.27)	2.039 (0.83)	4.612 (1.59)	5.724* (1.82)	9.707 (1.88)	2.415 (1.04)	2.345 (1.14)
Pseudo		0.0768	0.0749	0.0127	0.1568	0.3847	0.0673	0.0759
Observations		967	686	169	92	20	161	806
Prob > chi2		0	0	0.9824	0	0.0191	0.0009	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Tobit Regressions where each one is presented in one of the columns. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The fifth regression with proceeds larger than £333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.
The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of the lockup period, and Idiosyncratic shows the idiosyncratic risk for the shares in the sample. Market dummy is 0 for Main market and 1 for AIM market.

Table 4.16: Spread Nonlinear Regression for different sizes of IPO proceeds and for AIM and Main Using Cubed Model

	Whole sample						
	All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
log(proceeds)	-5.207***	-5.125***	-21.309	232.346	41.965	-1.171	-5.17***
	(-16.66)	(-9.48)	(-0.1)	(0.54)	(0.13)	(-0.4)	(-15.1)
(log(proceeds)^2	2.325***	2.554***	15.188	-107.942	-7.754	0.488	2.56***
	(7.1)	(4.15)	(0.11)	(-0.54)	(-0.07)	(0.32)	(5.22)
(log(proceeds)^3	-0.383***	-0.613	-3.543	16.57	0.221	-0.106	-0.512**
	(-3.96)	(-0.89)	(-0.12)	(0.54)	(0.02)	(-0.42)	(-2.32)
Constant	8.26***	8.2***	14.25	-161.41	-57.95	5.28***	8.2***
	(52.24)	(37.51)	(0.13)	(-0.53)	(-0.17)	(2.95)	(43.53)
Adjusted R2	0.3178	0.2534	-0.0177	-0.0213	0.0464	0.0204	0.2855
Observations	968	686	169	92	21	162	806
Prob F > 0	0	0	0.9943	0.7764	0.2991	0.0999	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Regressions using nonlinear cubed regression model to examine the relationship between the spread and log proceeds. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.

Another nonlinear regression that I have tested is the curve regression using the following model:

$$Spread = \frac{1}{b_1 + b_2 \log(proceeds)} + b_3\, Underpricing + b_3\, Return\, Stdev + \varepsilon_i \quad (4.7)$$

Table 4.17: Spread Nonlinear Regression for different sizes of IPO proceeds and for AIM and Main Using Curve Model

	Whole sample						
	All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
b1	0.119***	0.118***	0.231***	0.119	0.252	0.183***	0.118***
	(47.5)	(40.06)	(4.06)	(0.87)	(0.37)	(6.7)	(43.74)
b2	0.026***	0.026***	-0.004	0.028	0.022	0.015**	0.026***
	(24.51)	(19.52)	(-0.2)	(0.97)	(0.19)	(2.33)	(22.04)
Underpricing	-0.005**	-0.006**	-0.002	-0.005	0.015	-0.001	-0.006**
	(-2.38)	(-2.27)	(-0.71)	(-0.77)	(0.86)	(-0.39)	(-2.48)
Day 1 Return Volatility	-0.089***	-0.09**	-0.027	-0.032	0.322	-0.004	-0.088**
	(-2.83)	(-2.38)	(-0.32)	(-0.33)	(1.2)	(-0.05)	(-2.5)
Adjusted R^2	0.8146	0.8091	0.8702	0.8254	0.8122	0.8491	0.8127
Observations	968	686	169	92	21	162	806

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Regressions using nonlinear curve regression model to examine the relationship between the spread and log proceeds. The Table consists of three vertical sections. The first Section is for the whole sample (Main/AIM Markets). Under this section, I have run 5 regressions. The first one for the whole sample. The second regression is for the IPOs with a proceed less than £16.7 million. The third regression with proceeds between £16.7 – 66.7 million. The fourth regression with proceeds between £66.7 – 333.4 million. The second section shows the regression of the IPOs from Main market only. The third section shows the regression of the IPOs from AIM market only.

I can notice from Table 4.16, an Adjusted R^2 value of more than 0.80 appears across all the 7 regressions with a high significance for both b_1, b_2 at all the regression except when the proceeds are between 66.7M and 333.4M. This shows that the main contributor to the spread is the size of the proceeds.

Figure 4.3 shows a high effect of the proceeds on the spread charged by the underwriters. The effect is higher with the smaller IPOs. As the proceeds increase, the effect eases and becomes flatter.

4.17 Determinants of Underpricing

As the underwriters are not addressing the extra risk in the IPOs by increasing the spread, I will examine the second part of our third hypothesis. I will check if the underwriters will underprice the share price more for the riskier IPOs. From Table 4.9 I can notice that the idiosyncratic risk and the potential growth have a small correlation factor. However, the volatility of the share return shows higher correlation factor of a positive 0.2608.I ran the following Underpricing regression model:

$$Underpricing_i = \alpha + \beta_1 LogProceeds_i + \beta_2 Market_i + \beta_3 MultiBookrunner_i + \beta_4 Prestigious_i + \beta_5 ShareReturnVol_i \left(\sum_{j=1999}^{2012} \beta_j Year_j \right) + \varepsilon_i \quad (4.8)$$

I ran the regression for the whole sample, each market, and the different proceeds' sizes. A summary of the regression in Table 4.17.

Overall, the model was significant for the whole market, AIM market, and the small IPOs. From Table 4.17, the model shows to be significant for the whole sample, AIM market and the small IPOs. As you can see, the factor of the share return volatility is positive and significant for the whole sample, AIM market, and the small IPOs. This shows that the riskier IPOs tend to be underpriced more.

Log proceeds and Return Volatility variables showed to have a high significance on the model too. Log Proceeds variable has a coefficient of -12.98 for the whole sample. The negative value means that there is a negative relationship between the IPO size and underpricing. On other words, the cost of underpricing is higher with smaller IPOs. Moreover, as an absolute value, the Log Proceeds coefficient is higher for the small IPOs than the whole AIM market.

The coefficient for the Market dummy is 18.44 where AIM is 1 and Main is 0. This value eases the effect of Log Proceeds variable coefficient. The evidence mentioned above support this hypothesis. The AIM market has smaller IPOs in terms of proceeds and companies in this market are exposed to more risk. I found that IPOs from the AIM market are underpriced more than the IPOs from the Main market. This

could be introduced by the underwriters to mitigate the risk. This evidence supports the hypothesis H3(b).

Table 4.18 Underpricing Regression for different sizes of IPO proceeds and for AIM and Main

		Whole sample						
		All	Less than 16.7M	16.7M - 66.7M	66.7M - 333.4M	Greater than 333.4M	Main	AIM
Control Variables	Log Proceeds	-12.977***	-14.491***	-10.329	-37.046	48.27	-17.649	-12.194***
		(-3.82)	(-3.12)	(-0.25)	(-1.48)	(1.87)	(-1.22)	(-3.8)
	Market Dummy	18.437***	13.115	35.107*	12.108			
		(2.74)	(0.87)	(1.96)	(1.17)			
	Multi-bookrunner	-10.89	-6.885	-8.438	-5.518	-1.99	-22.716	-8.884
		(-1.4)	(-0.48)	(-0.36)	(-0.54)	(-0.11)	(-1.18)	(-0.92)
	Prestigious	-2.289	-7.062	8.535	18.627	25.11	13.821	-4.057
		(-0.55)	(-1.58)	(0.55)	(1.66)	(0.56)	(0.61)	(-1.06)
	Return Volatility	4.078***	4.429***	-0.943	-2.28	-5.55	0.641	4.270***
		(6.88)	(7.7)	(-0.21)	(-0.83)	(-0.73)	(0.15)	(7.98)
Year Dummies	2000	11.976	12.476	19.498	13.997	10.41	-4.264	18.468
		(1.09)	(1.04)	(0.38)	(0.6)	(0.29)	(-0.14)	(1.57)
	2001	11.532	11.038	-2.968	25.325	-31.14	-15.644	15.203
		(0.9)	(0.85)	(-0.04)	(0.56)	(-0.85)	(-0.34)	(1.19)
	2002	-12.715	-15.314	-39.169	20.582	-6.09	-21.21	-12.228
		(-0.94)	(-1.09)	(-0.61)	(0.61)	(-0.18)	(-0.52)	(-0.88)
	2003	0.237	-3.062	-18.489		-20.69	-7.104	1.607
		(0.02)	(-0.21)	(-0.24)		(-0.69)	(-0.13)	(0.12)
	2004	5.576	2.964	11.599	0.303		-19.563	8.611
		(0.5)	(0.25)	(0.22)	(0.01)		(-0.52)	(0.76)
	2005	18.889*	12.592	46.162	20.072	-5.22	47.117	16.907
		(1.74)	(1.09)	(0.85)	(0.81)	(-0.19)	(1.26)	(1.52)
	2006	7.864	5.298	5.631	17.946	-17.42	4.575	8.628
		(0.7)	(0.44)	(0.1)	(0.71)	(-0.59)	(0.12)	(0.75)
	2007	4.234	-2.269	5.367	33.095	-22.21	-4.775	5.742
		(0.37)	(-0.18)	(0.1)	(1.35)	(-0.81)	(-0.14)	(0.48)
	2008	3.385	1.735	27.03	24.473	-10.37	-29.472	8.345
		(0.2)	(0.09)	(0.41)	(0.73)	(-0.3)	(-0.42)	(0.51)
	2009	11.015	-12.824	-3.885	31.824		-6.16	14.513
		(0.51)	(-0.39)	(-0.05)	(1.03)		(-0.07)	(0.7)
	2010	13.201	26.543	15.704	22.991	-13.92	-2.659	16.919
		(0.93)	(1.41)	(0.27)	(0.86)	(-0.47)	(-0.06)	(1.16)
	2011	-14.652	-20.659	14.478	8.44	-54.96*	-12.627	-13.651
		(-1.07)	(-1.4)	(0.22)	(0.25)	(-2.06)	(-0.25)	(-1)
	2012	-7.506	-14.339	16.195	22.14		-18.669	-6.586
		(-0.5)	(-0.92)	(0.22)	(0.65)		(-0.27)	(-0.45)
Constant		7.298	10.808	1.093	43.277	43.28	43.527	5.192
		(0.68)	(0.94)	(0.02)	(0.83)	(0.83)	(1.06)	(0.47)
	Prob > F	0	0	0.8061	0.8047	0.7461	0.8473	0
	Adj R Square	0.096	0	0	-0.0385	0	0	0.1371

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 7 Regressions that are estimating the underpricing level. The underpricing is the return on the share having been bought on the IPO and sold at the closing of the first trading day, where each one of them is presented in one of the columns. It shows the results for the whole sample, IPOs from the Main market, IPOs from the AIM market, and different sizes of IPOs from both markets. The sizes of proceeds are shown in Millions of UK Pounds. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for Main market and 1 for AIM market. Multi-bookrunner takes a value of 1 if there is more than one bookrunner in the IPO and 0 if only one. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Return Volatility shows the standard deviation of the return on the share from the IPO date for a year.

4.18 Discussion of the results

AIM market is a new sector for the smaller and younger companies seeking access to the public equity. AIM market has been growing since 1995. For the IPOs in our sample, the money raised in the AIM market is about one-third the size of the money raised in the Main market over the same period and applying the same filters. There is no doubt that there are a number of differences in the characteristics of the two markets. In this section, based on the statistics of the data I have and using the regression model analysis results, I would like to verify and validate that the spread charged by the underwriters differs between the two markets.

From Table 4.3, Table 4.4 and Table 4.5, I can see that the spread charged differs. For the whole sample, the spread charged is (Mean = 6.42%, Median = 5.00%). When looking at the AIM market only, I find the results to be close as IPOs from the AIM market make about 71% of the whole sample. AIM market reported higher spread (Mean = 6.47%, Median = 5.08%). When I compare this to the Main market I can notice that the spread in this market is lower (Mean = 4.04%, Median = 4.00%).

Moreover, when I look into the spread regression model, I can see from Table 4.3, Table 4.4 and Table 4.5 that the market dummy is always positive and highly statistically significant when including the whole sample. This is also true for the univariate analysis in Table 4.10. The factor of the market dummy in the univariate table is 2.42 with 99% level of significance. From the above, I can accept the first hypothesis that underwriters spread is higher in the AIM compared to the Main market.

Economies of scale are noticed in many economic contracts and many aspects of business activities. The increase in production is often accompanied by a reduction in cost. I will go through our findings from our statistical analysis and examine if the

economies of scale apply to the spread charged as the proceeds increase. From Table 4.5, the spread charged for the IPOs in the AIM market is (Mean = 6.47%, Median = 5.08%). As I have discussed earlier, the size of proceeds is much higher in the Main market where I can see the spread is lower than the AIM market. The spread charged in the main market is (Mean = 4.04%, Median = 4.00%) as shown in Table 4.4.

In addition, when I look at Figure 4.4, Figure 4.6 and Figure 4.8 showing the spread to the log proceeds, I can notice that there is a downward sloping in all the three of them. This shows that the spread decreases as the proceed increases. Log of proceeds has the highest significance and is negative in the whole sample and in AIM. They both show a negative coefficient, which supports Hypothesis H2 about the dis-economy of scale. The absolute value of log proceeds has increased when I apply the FE model, the Tobit model and the non-linear model which support the hypothesis of the economies of scale.

How do underwriters deal with the IPOs in AIM? I addressed this issue by using several data. When underwriter evaluates the company seeking the IPO in order to set a fair share price, even with the due diligence in place, it is difficult to cover all the risk factors especially with small and young companies as those in the AIM market. With that in mind, I would examine the first part of our third hypothesis. Would the underwriter charge more in AIM market due to the associated higher risk?

I turn my attention to the fact that whether underwriters charge more for risky IPOs. The risk factors that I are testing here are the idiosyncratic, the volatility of the share return over the first year and on the contrast, I am testing the potential growth. From Table 4.10 I can see that the univariate test between the spread and the three factors mentioned above all show insignificant. Moreover, from Table 4.9, I can see

that the correlation between the spread and the three mentioned factors are minimal. However, I can look at the lockup period as an indirect indicator of the risk level as it is expected to have longer lockup period for riskier companies. The univariate test in Table 4.10 shows a positive 2.45. Also, from Table 4.11 and Table 4.12 I can see that the lockup period has higher significance for AIM and the smaller IPOs. This shows that risk in AIM IPOs has a minimal effect on the spread charged. Hence, I cannot accept the hypothesis H3(a) that the underwriters charge more in the AIM because the IPOs are riskier.

The insignificance of the risk proxies could imply that the proxies are not suitable to represent riskiness of the IPOs. I have examined two measures of risk. I have examined the idiosyncratic risk which is the risk specific to the specific company going public. Another risk measure is the volatility of the share return after listing. However, the estimation of ex-ante risk for IPOs is considerably more troublesome in light of the fact that there is no authentic price data before the IPO to calculate risk. The standard deviation of secondary returns after the IPO has been much of the time utilised as a proxy for ex-ante risk. By tradition, the standard deviations of the returns are utilised. This ex-ante risk measure has demonstrated practically little explanatory power, causing some prior researchers to conclude that risk does not fundamentally impact returns. Later studies (e.g., Johnson and Miller (1988) reason that the standard deviation of post-IPO returns is a poor measure of ex-ante risk; a great part of the current research makes utilisation of totally distinctive risk proxies (e.g., Carter and Manaster (1990). Nonetheless, I will also examine if the underwriter will mitigate the anticipated risk by underpricing to give compensation to the investors in the short-run.

To shed further light on the rent-seeking potential by the underwriters, I have examined the relationship between the spread and the lockup period. If the relation is negative, then this could mean that the underwriter offers a lower cost to the company to have a longer lockup period. Longer lockup period would allow the underwriters to benefit from any further issuing. Moreover, it could also indicate that the underwriter is offering the company shorter lockup period for a higher spread. A shorter lockup will allow the owners a faster exit from the company.

From Table 4.9, the correlation between the spread and the log of the lockup period is positive. In addition, looking at Table 4.10, I can notice that the coefficient of the log of the lockup period is positive and significant for the AIM market and the small IPOs. The same can be noticed from the iterations in Table 4.11 and Table 4.12 where the results did not change much from the original results in Table 4.10.

The above findings contradict the rent-seeking hypothesis and hence I will reject the fourth hypothesis that the underwriters charge more in the AIM market because of the rent-seeking. On the contrary, the positive relationship suggests that the longer the engagement, the higher the spread.

4.19 Conclusion

In this chapter, I have analysed the role of the underwriters and fees charged for managing the IPO. I have covered the IPOs that took place in both the Main market and the AIM market of London Stock Exchange over the span of 13 years from 1999 to 2012. After excluding the financial companies, the companies that are not incorporated in the UK and those with missing information, our sample had 972 IPOs. I examined the fees charged by the underwriters in terms of percentage and found that it ranges between 4% and 6.43% with a median of 5%. I have also calculated the

underpricing of the shares at the IPO and found that underpricing is less on the Main Market compared to the AIM market. However, IPOs issued in the AIM market showed that they are underpriced when calculated on the first day of trading. When I calculated it for the fifth day of trading, I have noticed that gap get narrower and this shows that prices were closer to the fair price.

I have also explored the fees charged by the bookrunners in both markets and found that fees vary between the IPOs. Nevertheless, as the proceeds increase, the fees tend to get lower as a percentage of the proceeds. The regression model showed that the main attribute to the fees charged is the value of the proceeds. Another factor that affects the fees was the year is which the IPO took place in. I have examined a number of factors such as rent-seeking in terms of the length of the lockup period. This proved to have little significance. I have also examined a number of other factors such as the age of the company at the time of the IPO, some risk factors such as idiosyncratic risk, and the potential growth. They all proved insignificant. I have also examined the pricing of the shares as an early indicator of the underwriters' proof of their due diligence and hence the post-IPO performance. I found that the underwriters provided fair prices when looking at the median of the share return on the first and fifth trading days. I have also run regressions to identify the factors that affect the pricing of the shares. I find that the model better describes AIM market and the smaller IPOs.

I investigated five hypotheses. The first hypothesis was that underwriters charge more in percentage terms in the AIM market compared to the Main market. I showed that the spread charged in AIM market has a mean of 6.47% and a median of 5.08% compared to a mean of 4.04% and a median 4.00% of for the Main market. The second hypothesis was to investigate the reason for the higher charges in the AIM

market where I tested the economies of scale from three different aspects. First, by examining the regression model for the spread, I found that log proceeds have a negative relationship with the spread. Second, when looking at the fees charged distribution, I noticed a convergence to 5% that continues below it as the value of the proceeds increases. Third, when studying underpricing, I found that the cost of underpricing is higher for smaller IPOs.

Table 4.19: Hypothesis analysis Summary

Hypothesis (Conclusion)	Supporting evidences
Hypothesis 1. Underwriters spread is higher in the AIM compared to the Main market. (Accept)	- Statistically, from Tables 1, 2, and 3, both the mean the median for the AIM market is higher than the Main market.
Hypothesis 2. Underwriters charge more on AIM because money raised in the IPOs is less (Dis-economies of scale). (Accept)	- From Table 6, the coefficient of the log(proceeds) is negative - From the regression model, the coefficient of the log(proceeds) is negative too. - Figures 1.4, 1.5, 1.6, 1.7, 1.8 and 1.9 all show a negative sloping - The return model shows that AIM IPOs are more underpriced which is results in a higher spread.
Hypothesis 3. (a) Underwriters charge more on AIM because AIM IPOs are riskier. (Reject) (b) Alternatively, underwriters mitigate risk by lowering the issue price, i.e. underpricing more. (Accept)	- (a) From Table 6, the coefficients of both risk factors are small and insignificant. - (a) Lockup period is used to mitigate risk. Its coefficient is positive - (b) From the return model, I found that IPO issue prices in AIM market are more underpriced compared to the Main market.
Hypothesis 4. Underwriter charge more in the AIM because of rent-seeking. (Reject)	- Log Lockup Period coefficient is positive both in Table 6 for the Univariate and in the regression model.
Hypothesis 5. Underwriters charge less to attract future business. (Reject)	- The model of future IPOs shows the coefficient of the spread charged the previous year to be positive.

The third hypothesis was to examine the risk factor and the usage of underpricing to mitigate risk. I tested idiosyncratic risk and the Standard Deviation for the share price. They both proved to be insignificant. Then, I tested the underpricing; this proved significant for the smaller IPOs and the AIM market. However, I found a positive relationship with the log of the lockup period which indicates a level of risk. The fourth hypothesis was to examine the possibility of rent-seeking contributing to the overall charged fees. I used the lockup period as an indication of rent-seeking. When examining rent-seeking, I found that the coefficient has a positive value which means that it is not used for rent-seeking.

The fifth hypothesis was the usage of lower charges by the underwriters to attract future business. I have examined 2 models in terms of the number of IPOs and value of proceeds. Both showed to be insignificant. However, I have noticed a positive relation between the average proceeds from last year to the number of IPOs the following year. This could be a result of the bookrunners are emerging to be more reputable. A summary of the hypothesis findings can be found in Table 4.19.

This chapter is one of the first work which tries to shed light on the underwriter's role in bringing a company public in the Alternative Investment Market. Though the Alternative Investment market is one of the most popular markets for smaller companies (Vismara et al., 2013) there are no studies, to the best of our knowledge, which tries to examine the cost of raising money in such a market. I try to fill the gap in the literature.

I tested a number of hypotheses. I find that economies of scale are a strong determinant of gross spread. The higher the money raised, the lower the cost. I test risk hypotheses. However, none of the proxies for risk is significant. I cannot conclude

that by using our measures that higher the risk higher the gross spread. Our measures for risk could be problematic as all uses the share price data after the IPO. I find some evidence that Indirect cost of IPO, underpricing, is negatively related to the gross spread. I did not find evidence that gross spread charged works as a marketing tool. I also found some evidence that investment bankers charge higher spread at the same time they impose higher lockup on the company, which might be consistent with the rent-seeking behaviour. However, further research is needed to shed more lights on the rent-seeking behaviour and underwriters spread.

Chapter 5

5. Short-run IPO dynamics and value of textual analysis

5.1 Abstract

In this chapter, I examine whether IPO texts are related to the short-run IPO dynamics like IPO underpricing, spread, lockup length, volatility and idiosyncratic risk. I took advantage of hand-collected data from the London Stock Exchange. I do so because no previous study has examined the IPO tones and its impact on short-run IPO dynamics in the UK, where the market structure in the UK is different in comparison to the US from a number of ways. I find that IPO tone is related to underpricing, spread and lockup length. However, I did not find any relationship between IPO tone and volatility or idiosyncratic risk. This may be due to the fact that these are not good measures of ex-ante uncertainty.

Keywords: London Stock Exchange, Underpricing, spread, lockup length, volatility, idiosyncratic risk, Textual analysis.

JEL: G14, G24, G30, G32

5.2 Introduction

In this chapter, I have started with an abstract of the chapter. The second section is the introduction of the chapter. The third section is the literature review which is followed by the hypotheses of this chapter. The fifth section is the data and methodology where we discuss the textual analysis and the tone measures. The empirical results are discussed in the sixth section where I started with the descriptive data then went through the hypotheses for the short-run dynamics. On the seventh section, I concluded.

There has been a large literature on issues in IPOs in the last 40 years. Yet, there are some puzzles remain with respect to underpricing and design of lockups. Though there has been a number of theories to explain underpricing including winners curse (Benvenist and Spindt, 1989), asymmetric information (Rock, 1986), IPOs as a marketing tool (Demers and Lewellen, 2003), information cascades (Welch, 1992), lawsuit avoidance (Hughes and Thakor, 1992; Tinic, 1988), behavioural biases (Welch, 1992, Loughran and Ritter, 2002), still it is not clear the reasons for underpricing. Also, lockup length is another puzzling issue in the IPO literature. Though there have been theories of asymmetric information and moral hazard (Brav and Gompers, 2003 and Hoque and Lasfer, 2016) to explain the existence and length of lockup, the evidence is not conclusive. In this chapter, I am examining some of the puzzling issues in the IPO literature using textual tones.

Textual analysis has been used in a growing number of papers and research contributing to the literature in Accounting and Finance to measure the tone and sentiment of corporate news releases, Management discussions and analysis (MD&A), annual reports (10K fillings), articles in the newspaper, and message boards for investors. Examples include Engelberg (2008), Li (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008), Tetlock (2007) Antweiler and Frank (2004). The

results of these studies could be summarised as negative words could be an effective way of measuring tone as negative words show a significant correlation with financial variables.

Other papers examine the tone of various documents with the financial variables. For instance, Kothari, Li, and Short (2008) relate the tone of newspaper articles on the cost of capital, return volatility, and analyst forecasts. Demers and Vega (2008), Engelberg (2008), and Henry (2008) relates news releases with net income of firms, drift in earnings, or share price returns. Some papers relate the information content of IPO prospectuses with the share price returns, share price volatility and trading volume (e.g., Feldman et al. (2008), Hanley and Hoberg (2010), and Li (2008, 2009)).

In this chapter, I am examining the effect of the text used in the IPO prospectuses on the short-run IPO dynamics of the newly listed companies on the London Stock Exchange. In particular, I use the spread, lockup period, underpricing, idiosyncratic risk and share price volatility to test for the short-run effect. I do so to shed light on the puzzles that remain in the IPO literature.

All the studies that relate IPO underpricing, volatility and spread to the tone of the IPO are based on the US data. To the best of our knowledge, there is no study which use the UK data to examine the information content of an IPO prospectus. I fill this gap in the literature. The UK and US IPO markets are very different. For example, UK market uses book building method and the UK uses open offers and placing method. In the book building procedure, information is gathered from the investors and reliance on the underwriters is less. On the other hand, open offers and placing method rely heavily on the underwriters' due diligence. Hence, methods of IPO (e.g.,

book building versus open offer and placing) has direct implications on the underpricing and IPO gross spread.

Lockup length in the US is very standardised, and the median lockup length is 180 days, and most of the IPOs has 180 days of lockups (Field and Hanka, 2001; Brav and Gompers, 2003). In the UK, the lockup length is very diverse and heterogeneous (Hoque, 2011). The average lockup length in the UK is 365 days, and the highest is 1080 days. Also, there are few different types of lockups, absolute versus relative date, one versus gradual release of the shares (Hoque, 2011). Since the US lockup length is shorter, not much information is produced during the lockup length. However, in the UK more information is produced and disclosed as the lockup lengths are longer. Thus the longer length of the lockups in the UK might have implications on the IPO prospectuses. Also, there is a quiet period in the US which is 40 days after the IPO (before 2002, it was 25 days). There is no quiet period in the UK. Since more information is produced and disclosed after the IPO, asymmetric information is less of a problem in the UK. Hence, the tone measure at the time of IPO may be not related to the lockup length in the UK.

Our results show a significant relationship with the dictionary words and IPO underpricing, lockup length and spread. I find that superfluous words and underpricing are significantly related. I also find that positive words are significantly related to underpricing. These results are in sharp contrast to the US market, where Loughrun and McDonald (2015) show a positive relationship between uncertain, weak modal and negative words and IPO underpricing. It may show different IPO market structure present in the US as compared to the UK. While in the US, book building is the IPO method normally followed, while in the UK private placements and open offer are practised most of the times. Rather our results are more consistent with Hanley and

Hoberg (2010) where they find standard content is positively related, and informative content is negatively related to underpricing.

Our results show a positive relationship between most of the tone measures and IPO spread. This is somehow consistent with Hanley and Hoberg (2010) who find a positive relationship between standard content and informative content with the Spread in the US market. Higher information disclosure in the UK increases the spread, which shows the underwriters can possibly measure the riskiness of the company and charge an appropriate fee for the services they provide. I need to bear in mind that the IPO spread in the US is almost fixed at 7%, while in the UK, it shows a large variation ranging from 4-11%. Our sample is dominated by the small AIM companies which might be riskier. Thus bookrunners charge according to the risk of the companies as reflected by the IPO tones in the prospectuses.

I also find a significant negative relationship between negative, litigious and uncertainty words and lockup length. The results show higher the risk related words the lower the lockup length. These results are consistent with the asymmetric information explanation of lockup length. Since these words mitigate the information asymmetry – to some extent - it reduces the lockup length.

I do not find any relationship between the tone measure of IPO and total volatility and idiosyncratic risk. This might be due to several reasons. The first one is that volatility and idiosyncratic risk is not an appropriate measure of ex-ante uncertainty of the IPO. Both volatility and idiosyncratic risk are measured after the IPO, hence might not reflect the true ex-ante risk of an IPO. Secondly, the effect of tone measures might die out as the time progresses. Since total volatility and

idiosyncratic risk is measured after some time of the IPO, it is unable to capture the impact the tone as measured by words mentioned in IPO.

I contribute to the IPO literature in various ways. There has been a longstanding puzzle in the IPO literature regarding underpricing and lockups. This chapter tries to explain the underpricing and lockups with the tones from prospectuses. This is related to the literature on tones of IPOs and 10-K using the US data (see Loughrun and McDonald, 2011, Hanley and Hoberg, 2010). I did an out of sample test using the data from UK where the IPO market setup is different. I also tested whether the lockup length is related to the asymmetric information problem using the IPO tone. Previous papers find that lockups serve to mitigate the information asymmetry and they reduce moral hazard problems. Using the tone information from IPO prospectuses, I find that IPO lockups mitigate the information asymmetry. This paper is related to a number of lockup papers in the US (Field and Hanka, 2001, Brav and Gompers, 2003, Brau et al., 2005) and the UK (Hoque, 2011, Hoque, 2015, Hoque and Lasfer, 2016).

Our work expands on earlier literature that analyses disclosure with regards to IPOs. For instance, Beatty and Ritter (1986) provide evidence that more detail data in the prospectuses regarding the use of proceeds found to increase underpricing. Conversely, Leone, Rock, and Willenborg (2007) and Ljungqvist and Wilhelm (2003) find that organisations that disclose more (less) information regarding the utilizations of the funds are associated with lower (higher) first day returns. Different researchers analyse whether there is a connection between the length and level of disclosure in the risk factors and valuation of IPOs. Beatty and Welch (1996) and Arnold, Fishe, and North (2010) analyse the prospectus in terms of length and level of disclosure of risks and find that more information disclosure in risk related section is connected with

higher IPO first day returns. Guo, Lev, and Zhou (2004) concentrate on an industry which is characterised by high risk as information is valuable—the biotechnology firms. They find a negative connection between the degree of risk disclosure and the bid–ask spread, however, don't analyse if there is a connection to IPO underpricing.

The rest of the paper is organised as follows. In Section 2 I present a literature review. Section 3 develops hypotheses. Section four describes the data and methodology. Section 5 presents the empirical results. Finally, section six concludes.

5.3 Literature review

Recently, there has been a number of studies that uses textual analysis to analyse the financial information contents provided by the companies and firms working in the public market. When checking the term "text" in Oxford dictionary, the first definition is "A book or other written or printed work, regarded in terms of its content rather than its physical form" (2015). However, the text in its essence is a conversion of voice and sounds people use to communicate into a written format. Although that the tone of speech is not captured in this process of conversion, yet the words used would partially reflect the sentiment. Humans have been using the concept of the Textual analysis very early in the history as part to understand the conveyed message better regardless of its format (McKee, 2003). A more structured format of textual analysis emerged more than a thousand years ago as people were studying scriptures to understand them and to criticise their authenticity (Karcic, 2006). In the modern ages, text mining – a term that was more frequently used – was first noticed in mid-1980's. However, this process used to be expensive and cumbersome due to the fact that it could only be done by real human reading throughout the text and indexing all words (Bhattacharyya, Das, Mitra, Ganguly, Das, Bandyopadhyay and Kim, 2009).

However, as the computer started to have more power and to be affordable, textual analysis advanced noticeably since then.

Earlier use of textual analysis was mainly focusing on the analysing the political context and not specifically designed to analyse the financial documents. An important study by (Loughran and McDonald, 2011) made a major contribution by regenerating the dictionaries used to evaluate the sentiment of the documents. In this study, authors had examined the dictionaries created by Harvard University. Harvard dictionaries categorised words in a way to show their tones. However, the authors argue that some of the words are taken out of context when I look into financial text. They have reported that almost three-fourths of negative words used in 10 – K based on Harvard dictionary are typically not negative financial context. They worked on creating word lists containing words that reflect different tones. They have examined all the words used in 10-K documents and captured all the words that have been used for at least 5%. They have categorised these words into negative and positive lists with the financial meaning in mind. They came with a new set of dictionaries that is more related to the financial literature.

To be listed in the US market, a company has to fill and S1 form. S1 form contains initial data about the company. This data will be used in the preparation of prospectus. In this study by (Loughran and McDonald, 2013), they are measured changes in offer price and in S1 language to the final prospectus. They find that higher use of certain words in S1 results in upward offer price revisions. Loughran and McDonald (2015) have compared the results of the analysing financial documents using their word list (hereafter LM) with Diction's. Diction platform and its word lists are widely used in analysing financial documents in literature. However, (Loughran and McDonald, 2015) find that about 83% of optimistic words and about 70% of

pessimistic words used in the lists are misleading. They have analysed the pre-filing K-10 documents using both LM's and Diction's word lists. They have reported a negative correlation between the optimism word frequency and LM positive words. This means that the more optimistic tone of the document using Diction's, the less positive tone using LM positive word list. They ran regressions measuring the effect of the four list analysis (Diction's optimism and pessimism word lists and LM positive and negative word lists) on the post-filing return volatility. Using Diction optimism word frequency, its coefficient was significant and positive. They argue that this is somehow surprising as per previous empirical evidence investors do not value positive language in the business text. LM negative word list frequency coefficient was insignificant. This is what they have expected as it is consistent with the aforementioned previous empirical evidence. Both Diction's pessimism and LM negative frequency coefficients were significant and positive. They refer the ability of Diction pessimism word frequency to reasonably explain the pre-filing return volatility to the correlation of 0.688 with LM negative word list.

Kearney and Liu (2014) have surveyed about thirty-eight papers that have contributed to the financial textual analysis. The have studied them from three aspects. They examine the information sources, content analysis methods and financial models. The authors categorised the sources into three main categories. The first category uses sources to measure the corporation-expressed sentiment. This mainly includes the documents that are released by the corporates such as annual reports, earnings press releases and conference calls. I notice that among all the surveyed papers, (Ferris, Hao and Liao, 2013) was the only paper that used the IPO prospectuses as an information source. The second category was for the information sources used to major the media-express sentiment. Information sources that fall under this category include news

stories, in-depth commentary and analyst reports. The third category was the internet-expressed sentiment. The information sources included in this category are the messages on the Internet such as the messages posted on Yahoo!Finance and Raging Bull.

In terms of the content analysis methods, Kearney and Liu (2014) report that the two main methods used for textual analysis are the dictionary-based approach and machine learning. However, the dictionary-based approach is the dominated one among the surveyed studies. In earlier studies, they have reported the use of DICTION and GI/Harvard dictionaries. However, in the most recent studies, they have reported a tendency to use a finance specific words dictionaries such as the LM word list.

Finally, Kearney and Liu (2014) have surveyed the financial models used and reported that the most common used is the linear regression model on time series data. They have also reported four other less frequently used models such as vector autoregression model (VAR), logistic and probit regression, volatility model and textual sentiment-based trading model.

Ferris et al., (2013) studied the level of conservatism in the IPO prospectuses. Their sample consists of 1100 IPO prospectuses from the US market over the period from 1999 to 2005. In their study, they have used LM word list, Diction dictionary and Harvard dictionary. They measured the level of conservatism in an IPO prospectus using the negative word list and calculating its percentage to the total words in the prospectus. They have documented that a greater conservatism in a prospectus is related to an increased and replacing. In addition, they have reported that conservatism shows a significant inverse relationship to the industry return on assets for three years following the IPO. However, the above applies to the non-technology companies and

less or no significance on technology companies. They use a separate analysis for technology companies due to the high uncertainty in them. They conclude that lack of power of productivity concerns the idea of those companies being fundamentally hard to evaluate.

The US Securities and Exchange Commission (SEC) introduced a rule effective from the first of October 1998 for the companies willing to be listed to use plain English in their prospectus filings. The mandate states "*To enhance the readability of the prospectus, you must use plain English principles in the organization, language, and design of the front and back cover pages, the summary, and the risk factors section*" (1998). Loughran and McDonald (2014) have measured the change in the effect of the language used in the 10-K, 424 forms and issued prospectuses for IPOs that took place for the period between 1994 and 2009. They captured six components from the documents to measure the plain English in a document. These components are the average sentence length, the average word length, use of passive voice, legalese, personal pronouns and other low frequent measures such as negative phrases and superfluous. In their study, they have normalised all the measures for all the documents (mean of 0 and standard deviation of 1). As the measure increases, it implies more compliance with the plain English mandate. They have reported that the all the three types of documents have shown a reaction to the mandate towards compliance. The measure for the 424 Forms and the IPO prospectuses increased more than 3 standard deviations over the following 2 years. Moreover, the effect on IPO prospectuses was almost immediate as it was reported from 1999.

Demers and Joos (2007) did a study developing an IPO failure prediction model that includes accounting information, firm and IPO characteristics. They have

documented that their estimated probabilities of failure driven from their textual analysis model is significantly and negatively associated with abnormal returns after the IPO. Elliott, Rennekamp and White, (2015) measured the effect of using concrete language and clear disclosure on the willingness of the investors to invest. They reported that investors are more comfortable when more information is provided.

5.4 Hypothesis Development

In this chapter, I am examining the effect of the sentiment and tone used in the prospectuses writing on the short-run post-IPO performance. I use the spread, IPO underpricing, lockup period, idiosyncratic volatility and share price volatility.

I can categorise the dictionaries into 3 main groups. The first group contains the dictionaries that can be used in a conservative context. These dictionaries are Negative, Harvard IV, Constraining and Litigious. The second group contains the dictionaries that don't add much information to the context. However, using them at a higher rate can indicate a level of hesitance and uncertainty. The group of this category contains Superfluous, Modal and Uncertainty. The third group contains the dictionaries that can be used in marketing materials. These dictionaries are Interesting and Positive.

When the text used in the prospectus is more of a conservative tone, I would expect that the spread charged by the underwriters, underpricing and lockup period would be higher due to the fact that the underwriter is presenting more conservative information in terms of negative, constraints and litigious information. Ferris et al., (2013) reported a positive relationship between the conservatism of the prospectuses and underpricing. In terms of performance, I expect the shares to have lower

performance and volatility as risk factors are already presented in the prospectus and hence, less price correction and volatility are expected.

For prospectuses with higher usage of words from the second group, would be expected to have lower spread. The Higher rate of superfluous words can be an indication of a lack of the necessary information which would be a result of less effort in the preparation of the IPO and hence, fewer fees charged by the underwriters. However, to compensate for the lack of information, underwriters may tend to underprice more and apply longer lockup periods. In addition, shares are expected to perform less than the market during a higher volatility period.

The third group is the prospectuses that use more positive tone in their text. They show more interesting words in their prospectuses. These kinds of words are used to market the IPO and to make it more appealing. Presenting more positive words can be an indicative of more effort paid by the underwriters to show the strengths of the company. This is expected to be reflected in a higher spread. In addition, underpricing and lockup period are expected to be higher to mitigate the risk of share prices drop.

5.5 IPO Spread and tone

Hanley and Hoberg (2010) assert that a good amount of effort and resources spent by the underwriters in the pre-IPO stage. In their words: “Substantial resources are expended on due diligence by the underwriter, the issuing firm, and their legal counsel to gather information about the firm. While some of this expenditure is due to regulatory or liability concerns, it is plausible that greater effort expended in the premarket to acquire information about both the issuing company and its competitors. If the issuer and underwriter choose to have more accurate pricing in the premarket,

they will expend greater effort in acquiring information through enhanced due diligence about the issuing firm and its competitors." (Hanley and Hoberg, 2010, p. 2826-27). In that case underpricing will be lower and underwriter spread will be higher.

The advantage of expanded data collection in the US before the IPO is that the underlying offer price will be a more precise appraisal of the last offer price and aftermarket trading price (Hanley and Hobarg, 2010). Since fewer data will be assembled from investors during book building, underwriters who conduct a higher level of premarket due diligence gain from both lower value changes and lower compensation to educated investors through underpricing (Benveniste and Spindt, 1989). This advantage is balanced by the possible costs of uncovering exclusive data to rivals. Further, improved due diligence may turn out to be excessively costly exercise, making it impossible to higher lawful expenses and underwriters pay. This is especially valid if the underwriter's remuneration can't be raised sufficiently high to make up for the extra effort because authors find that underwriters make exactly 7% on IPOs (see Chen and Ritter 2000, Abrahamson et al., 2011). But this may not be the case in the UK as there is no fee clustering in the London Stock Exchange. I report while the average fee in the UK is 6.67%, there is a huge variation in the level of fees. For example, underwriters charge from 4.00% to 11.00% in the UK. So it is possible that the underwriters who pursue a higher level of diligence charge higher level of fees.

Hanley and Hobarg (2010) assert that the result of more prominent data generation in the premarket ought to be more informative substance in the initial prospectus. On the other hand, when book building is used the preliminary prospectus is likely to contain more standard content following the recent trends of the industry.

In the absence of book building in the UK, I expect that UK underwriters will try and collect more information regarding the company before IPO. Hence, by decomposing the information into positive and negative tones, I will be able to identify how information affects the spread and underpricing.

Issuers and underwriters who are willing to put more resources to get higher production of information before the IPO are expected to have better information content in the preliminary prospectus. Underpricing will be lower as higher information content is presented since the company or the investment banks do not need to reward investors who provide information in the book building phase of the IPO (Benveniste and Spindt, 1989). Since book building is not used in the UK, the tone of the IPO will be reflected in the IPO underpricing and underwriter spread. I conjecture that positive tone will reduce the underpricing and increase the spread.

Conservative and Uncertainty tones reflect a riskier IPO. Underwriters are expected to charge higher spread in this case to cover for the anticipated risk. On the other hand, when enough data is presented with a positive tone, this indicates more due diligence and extra effort paid in data collection and hence, higher spread as well.

H1: Underwriters charge higher spread for all the tones whether it is Conservative, Uncertainty or Positive tone.

5.6 IPO Underpricing and tone

A number of theories in IPO pricing state that uncertainty of the company is related to IPO underpricing (see Ritter, 1984; Rock, 1986). However, since the IPOs are not traded in the market, and there is not much information regarding them it is very hard to measure the ex-ante uncertainty and risks surrounding these companies. Though inverse issue price, firm age, pre-IPO sales, IPO gross proceeds are all used as a proxy

of uncertainty; these measures are subject to limitations, and they could be measuring many other things at the same time. Previous literature has used alternative measures like standard deviation of aftermarket returns. However, they are not without critic as standard deviation is measured in the aftermarket.

Beatty and Ritter (1986) study have demonstrated a positive association between ex-ante uncertainty surrounding the IPO and the expected underpricing. According to the previous studies, a higher ex-ante uncertainty of IPO firms should result in higher first-day returns. Consistent with the previous studies, Loughran and McDonald (2015) report that certain words which represent ex-ante risks of IPOs in the Form S-1 are significantly related to IPO underpricing after they control for other variables which are found as significant in the previous literature. Since IPO prospectus is the document which presents relevant financial and business information in the UK, I conjecture that certain risk related words like uncertain, weak modal and negative words are positively related to underpricing.

Underpricing is related to the riskiness of the IPO. Conservative and Uncertainty tones reflect a risker IPO and hence a underprice more. While a positive tone would indicate a less riskier IPO and hence will underprice less.

H2: Underwriters will underprice more when underwriters use a Conservative or Uncertainty tones while it will underprice less when using a Positive tone.

5.7 Lockup length and tone

Lockup is an agreement between the underwriters and insiders of the firm not to sell the shares for an agreed period after the IPO. Since there are no rules regarding lockups, it is a voluntary agreement. However, almost all the IPOs have lockups. Literature shows that lockups reduce the asymmetric information problem and moral

hazard problem (Brav and Gompers, 2003, Brau and Fawcett, 2005, Hoque and Lasfer, 2016). Literature has also proposed that lockup is a mechanism for rent-seeking, however, the evidence for this is not very clear (Brav and Gompers, 2003 and Hoque and Lasfer, 2016).

Most of the literature on Lockups on the US where lockup length is standardised and found to be 180 days. However, the lockup length in the UK is much more heterogeneous, and the length varies. The median lockup length is 365 days, with lowest of 90 days and highest of 1080 days. Also, there are few types of Lockups. According to Hoque (2011), there are fixed lockup date and relative lockup date. Also, there is the gradual release of shares. The varied nature of lockup in the UK in comparison to the US makes it more interesting to relate the IPO prospectuses tone to the lockup length. It is also interesting to look at it because since lockup length is larger in the UK, so more information is produced and disclosed during this time. It would be an interesting to examine the relationship between IPO tone and lockup length.

If lockups reduce the asymmetric information problem and provide investors with the confidence, the same objectives could be achieved by the tone of the IPO prospectuses. Thus I conjecture that that uncertain, weak modal and negative words are positively related to lockup length. However, if the lockups reduce the moral hazard in the UK, I would expect no relationship between the lockup and IPO tone.

Underwrites impose a lockup period to show the commitment of the internal management and hence mitigate the risk of information asymmetry. As a result, Conservative and Uncertainty tones is expected to be related with a higher lockup period. While a positive tone would indicate a less lockup period.

H3: Lockup period will be more when underwriters use a Conservative or Uncertainty tone and will be less when they use a Positive tone in the prospectuses.

5.8 Volatility and tone

There is a huge literature on the connection between disclosure and volatility (see the literature review by Healy and Palepu (2001)). Disclosure reduces the volatility because more disclosure is likely to reduce the degree of information asymmetry between the insiders and the outsiders (e.g., Amihud and Mendelson 1986; Diamond and Verrecchia 1991; Easley and O'Hara 2004). Information asymmetry widens the adverse selection component of the bid-ask spread required by the market makers and thus increases the volatility of security (Kothari, Li and Short, 2009).

The covariance of the firm's cash flows with other firms' cash flow decreases with a higher level of disclosures (Lambert, Leuz and Verrecchia, 2007). This reduces the volatility of the firms which is disclosing more information. Kothari, Li and Short (2009) assert that the higher quality of information disclosure improves the investment decisions made by managers. The improvement is caused by the reduced uncertainty as a result of better information.

Loughrun and McDonald (2015) assert that certain words such as uncertain, weak modal, and negative word frequencies proxy for ex-ante uncertainty about the IPO. Beatty and Ritter (1986), in an earlier paper, use two measures of ex-ante risk of an IPO: (1) log (1+ the number of different uses of proceeds mentioned in the IPO prospectus) and (2) the inverse of IPO gross proceeds. Beatty and Ritter (1986) rationalise the use of a number of different uses of IPO proceeds is motivated by the SEC regulations. During their sample period, there was a requirement by the SEC for the more speculative offerings to provide with the detailed uses of funds that are going

to be raised during the IPO. The less risky IPOs are not required to provide such information.

Beatty and Ritter (1986) uncertainty proxies are out of context in our study as this paper is using UK data and there are no such requirements in the UK. Firstly, in the UK there is no difference in requirements between the established firms and risky offering to report the intended use of proceeds. Secondly, as suggested by Loughrun and McDonald (2015), amid the Internet bubble, a few extremely risky offerings had huge net proceeds from the IPO. Due to differences in regulations imposed by the United Kingdom Listing Authorities (UKLA) as compared to the SEC, Beatty and Ritter's measures are not appropriate for our sample. Rather, I use the Loughrun and McDonald (2015) tone measures to relate to the share price volatility.

Since the Conservative and the Uncertainty tones are related with a higher risk, they are expected to result in more underpricing and higher volatility and hence higher first-day return. While the Positive tone is less received by the investors and is not expected to show any impact on the underpricing or the volatitlity on the first-day of trading.

H4: Share price volatility will be higher when underwriters use a Conservative or Uncertainty tones and no effect when they use a positive tone in the prospectuses.

5.9 Idiosyncratic risk and tone

One of the ways to measure business risk is an idiosyncratic risk. Idiosyncratic risk is the residual risk from the market model regression which is not explained by the market and specific to a business sector. Lambert et al. , (2007) hypothesise that the covariance of the firm's cash flows with other firms' cash flow decreases with a higher level of disclosures, which reduces the volatility of the firms with more disclosers. If

the covariance of the cash flows between two firms is zero, we are left with the idiosyncratic risk because the idiosyncratic risk is uncorrelated across firms. In that circumstance, firms with more disclosure have a lower idiosyncratic risk. Kothari, Li and Short (2009) assert that the higher quality of information disclosure improves the investment decisions made by managers. The improvement is a result of the reduced uncertainty as a result of better information. Again, this might be uncorrelated across the business or the industry. Since information about the business is reflected in the tone of the IPO, I try to relate the information content of the IPO with the idiosyncratic risk. To the best of my knowledge, no previous paper related the idiosyncratic risk of an IPO with the information content of IPO prospectus. Thus, I conjecture that:

With hihger idiosyncratic risk, underwriters are expected to be more conservative. On the other hand, a positive tone would indicate less idiosyncratic risk. However, Uncertainty tone is not expected to have any effect.

H5: Idiosyncratic risk is higher when underwriters use a Conservative tone, lower when using a Postive tone and will not be affected by the Uncertainty tone in the prospectuses.

Table 5.1 summarise the five hypothesis

Table 5.1: Hypothesis summary

Sentiment		H1 Spread	H2: Underpricing	H3: Lockup Period	H4: Volatility	H5: Idiosyncratic
Conservative Dictionaries	Negative	Higher	Higher	Higher	Higher	Higher
	Harvard_IV	Higher	Higher	Higher	Higher	Higher
	Constraining	Higher	Higher	Higher	Higher	Higher
	Litigious	Higher	Higher	Higher	Higher	Higher
Uncertainty Dictionaries	Uncertainty	Higher	Higher	Higher	Higher	Neutral
	Superfluous	Higher	Higher	Higher	Higher	Neutral
	Modal	Higher	Higher	Higher	Higher	Neutral
Marketing Dictionaries	Interesting	Higher	Lower	Lower	Neutral	Lower
	Positive	Higher	Lower	Lower	Neutral	Lower

This table shows a summary of the hypotheses I have in this chapter. It shows the expected effect of the usage of each of the dictionaries in the prospectuses on short terms variables.

5.10 Data and methodology

I use the same IPO prospectuses that I used in chapter 1. As per Table 5.3, The median of the number of words per prospectus in the main market is 66,662 words while for the AIM market it is 27,792 words. I have only included the IPOs that have more than 2000 words that is less than 10 percent of the median of the AIM market prospectuses and hence would not have a comparable data. The total number of prospectuses in our sample is 946 where 26 of them were photocopies and required optical character reader (OCR) to convert them to text. I used the tool provided by Google Drive to convert photocopies in PDF format into text.

I used the dictionaries developed by (Loughran and McDonald, 2011) in addition to Harvard IV dictionary for negative words.

To analyse each prospectus, I use the AntWordProfiler . For each prospectus, I have gathered the following information:

- Total number of words in the prospectus
- The frequency of appearance of each word in the prospectus
- The number of words appeared in the prospectus that belongs to each of the dictionaries.
- Numerical characters have been excluded from the counting

The pre-IPO financial data have been collected manually from the prospectuses. This includes the Pre-IPO Net Income and the Pre-IPO Sales. I also calculate the Overhang using the financial data provided in the prospectuses. The post-IPO financial data have been collected using DataStream. Market information and Indices have been collected using DataStream too.

Table 5.2: Dictionaries Statistics

Group	Dictionary	Number of words
	All words	84,669
	Uncategorised words	77,602
	Categorised words including Harvard IV	7,067
	Categorised words by (Loughran and McDonald, 2011)	4,118
Conservative	Negative	2,329
	Harvard_IV	4,187
	Constraining	187
	Litigious	886
Uncertainty	Superfluous	56
	Modal	60
	Uncertainty	297
Marketing	Interesting	68
	Positive	354
	Irr_Verb	158

This table shows the total number of words in the main dictionary and in each categorised dictionary.

Loughran and McDonald (2011) have introduced 9 different dictionaries in addition to the Harvard IV negative words dictionary. There is some overlap between those dictionaries were some words appear in more than one of them. The dictionaries are as follows (Negative, Positive, Uncertainty, Litigious, Constraining, Superfluous,

Interesting, Modal, Irregular Verb in addition to Harvard IV Negative). Table 5.2 shows some statistical information about the dictionaries.

I can notice from Table 5.2 that the sizes of the dictionaries vary a lot. The largest dictionary by far is Harvard IV. LM Negative dictionary is almost half the size of Harvard IV. The other LM dictionaries sizes range between few hundred words to less than a hundred.

Following the same methodology used by Loughran and McDonald (2011), I have calculated the proportional weight of each dictionary as follows:

$$Proportional\ Weigh\ of\ a\ Dictionary = \sum \frac{Word_i\ Frequency}{Total\ Nubner\ of\ Words} \tag{5.1}$$

For all the words from this specific dictionary.

This model simply calculate the weight of the accumulated frequency of words from a certain dictionary to the total number of words in the prospectus. This gives an independent weight of the tone used in a prospectus without taking in consideration the usage of the same words or terms on the other prospectuses.

I have also calculated the Time Frequency – Inverse Document Frequency (tf.idf) so that the weight of a certain word is in accordance with its importance across all the prospectuses under our study. The calculation of the weight of each word is as follows:

$$W_{i,j} = \begin{cases} \frac{(1+\log(tf_{i,j}))}{(1+\log(a))} log\frac{N}{df_i} & if\ tf_{i,j} \geq 1 \\ 0 & Otherwise \end{cases} \tag{5.2}$$

Where:

$W_{i,j}$: the weight of the i^{th} word in the j^{th} document

$tf_{i,j}$: the row word count of the i^{th} word in the j^{th} document

a : is the average word count in the document

N: Number of documents

df_i: Number of documents containing at least one occurrence of the i^{th} word

From this model, it is noticeable from the first part of the equation that the weight of the word in a certain prospectus increases as the number of appearances in the prospectus increases. However, it is multiplied by the log of the total number of documents divided by the number of documents that has at least one appearance of the word. This will give more weight to the words that do not appear much as this part will exponentially decrease as the number of documents with the same word increases.

I used the same regressions I used in Chapter 1 and introduced the dictionary weight as dependent variables. I introduce the variables one at a time to avoid the multicollinearity.

Table 5.3: Textual Analysis Statistics

		All	Standard Deviation	Main	Standard Deviation	AIM	Standard Deviation	Main to AIM In %	Weighted % to Average Words: Main	Weighted % to Average Words: AIM
Number of prospectuses		946		159		787				
Number of words										
Average		38,818	31,209	74,023	46,763	31,706	20,740	233%		
Median		30,183		66,662		27,792		240%		
Highest		288,796		288,796		157,637		183%		
Lowest		2,198		3,169		2,198		144%		
Average Number of words from dictionaries		465	192	627	222	433	167	145%		
Dictionaries										
Conservative	Negative	376	347	627	537	433	238	145%	0.85%	1.37%
	Harvard_IV	1356	1147	599	1787	269	750	223%	0.81%	0.85%
	Constraining	318	275	693	401	354	202	196%	0.94%	1.12%
	Litigious	411	299	786	443	342	220	230%	1.06%	1.08%
Uncertainty	Superfluous	10	11	586	19	264	7	222%	0.79%	0.83%
	Modal	524	415	184	597	62	297	297%	0.25%	0.19%
	Uncertainty	417	363	410	539	167	260	246%	0.55%	0.53%
Marketing	Positive	207	172	727	245	305	116	238%	0.98%	0.96%
	Interesting	82	99	17	181	9	53	189%	0.02%	0.03%
	Irr_Verb	325	248	965	380	435	163	222%	1.30%	1.37%

This table shows statistics of the usage of the dictionaries in prospectuses in our sample. It shows the average number of words used in All the sample, Main market and AIM market. Main to AIM in % shows ratio of the average number of the words used in the Main market prospectuses compared to AIM market prospectuses in percentage. Weighted % to Average words show the ratio of the average number of the words from each dictionary for a certain market to the average number of words of the dictionary of the whole sample.

5.11 Empirical Results

5.12 Descriptive Statistics

Most of the prospectuses under our study are for IPOs that took place in the AIM market. There are 787 prospectuses for AIM market IPOs compared to 159 prospectuses for the Main market IPOs. A brief statistical information about the textual analysis of the prospectus can be found in Table 5.3

Table 5.3 shows that the average word count of the Main IPO prospectuses is more than twice those of the AIM IPO prospectuses. In addition to that, the average word count for the categorised words is about 45% more in Main prospectuses compared to those in AIM. This implies that text used in Main prospectuses provide more information about the companies.

On the lower part of Table 5.3, I can still notice that the words used from each dictionary are higher in Main Market than it is in AIM market. This can be attributed to the fact that the word count is higher in the Main Market prospectuses. At the right most of the lower part of Table 5.3, the weighted percentage of the average word count from each dictionary to the average number of words of the prospectuses from each market is shown. It shows that on average, 1.37% of the words used on AIM IPO prospectuses are coming from LM negative dictionary compared to only 0.85% from the Main market prospectuses. This shows that prospectuses issued in the AIM market are showing more conservatism compared to the Main market. Harvard IV dictionary is also showing a slightly higher percentage in AIM than Main prospectuses. However, Harvard IV dictionary shows a higher Standard Deviation than the Mean. This indicates that the range of numbers is large. In other words, it implies that there are only a few words from this dictionary compared to a significantly larger usage of the

words in some other prospectuses. The third dictionary in the Conservative group is LM constraining dictionary. It shows 1.12% in AIM prospectuses compared to 0.94% in Main prospectuses. It is expected for the smaller companies to have more constraints than larger companies with more facilities. For LM Litigious, AIM prospectuses are showing a slightly higher percentage of 1.08 compared to 1.06 in Main. This shows that all companies in both markets are putting the same weight on litigious words. The percentage use of the other dictionaries is close between the two markets except for LM modal where it is slightly higher in Main prospectuses compared to AIM prospectuses. Yet, with a higher Standard Deviation than the Mean.

In this chapter, I have included a few variables in addition to those that I used in chapter 1. I have included Overhang, Pre-IPO Net Income and Pre-IPO Sales. First, I examined the correlation between them and the Dictionaries Proportional Weights and tf.idf Weights as shown in Table 5.4.

By comparing Panel A to Panel B in Table 5.4, I notice that the tf.idf weights have better correlation with the other variables in terms of significance. This is due to the fact that the tf.idf puts more weight on the terms that are less frequently used. Hence, the usage of such a word will attract more attention from the audience to the sentiment of the context. When looking at Panel A, I measure the occurrence of the words from a certain dictionary as a proportion of all the words used in a document in isolation from the other documents. On the other hand, Time Frequency – Inverse Document Frequency put different weights to different words within a certain dictionary based on the usage of that word in all the words used in all the documents under study.

From Panel A, showing the correlation with the proportional weight of the different dictionaries, I can notice that there is a positive correlation between Log(Proceeds) and Negative, Positive, Uncertainty and Harvard Negative. This shows that larger IPOs provide more information in the aforementioned domains. A negative correlation can be seen for Litigious and Superfluous in addition to the Irregular Verbs. This indicates that larger IPOs show less litigious information that the smaller IPOs as a proportional of the whole prospectus.

Table 5.4: Correlation table for the main variables with the different dictionaries weights

	Proceeds	Log (Proceeds)	Spread	Idiosyncratic	STDev	Age in Days	Potential Growth	D1 Underpricing	Overhang	Pre-IPO Net Income	Pre-IPO Log(Sales)
PrNegative	0.07**	0.152***	0.068**	0.001	-0.012	-0.035	-0.031	-0.072**	0.006	0.018	0.021
PrHarvard_IV	0.052	0.118***	-0.013	0.036	-0.005	-0.026	-0.101***	-0.06*	0.01	-0.002	0.127***
PrConstraining	0.026	0.109***	-0.051	0.011	-0.006	-0.085**	0.002	-0.028	-0.054*	-0.032	0.066
PrLitigious	-0.136***	-0.224***	0.086***	0.02	0.044	-0.074*	-0.007	0.191***	-0.02	-0.01	-0.051
PrSuperfluous	-0.125***	-0.173***	0.077**	-0.014	0.077**	-0.087**	0.006	0.012	-0.027	-0.02	-0.144***
PrModal	0.033	-0.004	0.014	-0.03	0.01	-0.121***	-0.022	0.041	-0.065**	0.046	-0.144***
PrUncertainty	0.043	0.125***	-0.011	0.031	0.029	-0.05	0.002	0.008	-0.054*	-0.007	-0.006
PrPositive	0.03	0.087***	-0.09***	-0.001	-0.001	0.081**	-0.002	-0.075**	-0.022	0.018	0.037
PrInteresting	0.026	0.016	0.011	-0.02	-0.015	0.097**	0.018	-0.005	0.032	-0.025	0.059
PrIrr_Verb	-0.1***	-0.242***	0.112***	-0.013	-0.005	0.024	-0.079**	0.094***	0.036	0.021	0.045
Panel A: Correlation of the main variables with the proportional weight of each dictionary											
TFNegative	0.411***	0.561***	-0.129***	0.004	-0.025	0.07*	0.018	-0.168***	-0.018	0.064*	0.23***
TFHarvard_IV	0.371***	0.515***	-0.111***	0.004	-0.028	0.044	0.016	-0.167***	-0.014	0.052	0.177***
TFConstraining	0.453***	0.605***	-0.134***	0.016	-0.027	0.077**	0.027	-0.171***	-0.003	0.075**	0.306***
TFLitigious	0.458***	0.595***	-0.137***	0.015	-0.027	0.094**	0.037	-0.159***	0.014	0.096**	0.275***
TFSuperfluous	0.195***	0.289***	-0.04	-0.022	-0.002	0.033	-0.001	-0.129***	-0.018	0.008	0.108***
TFModal	0.239***	0.361***	-0.061*	0.021	-0.033	0.022	0.03	-0.15***	0.001	0.012	0.075*
TFUncertainty	0.314***	0.452***	-0.087***	0.021	-0.027	0.072*	0.011	-0.137***	-0.007	0.034	0.166***
TFPositive	0.369***	0.539***	-0.128***	0.027	-0.036	0.091**	0.025	-0.19***	-0.002	0.068*	0.211***
TFInteresting	0.288***	0.418***	-0.088***	-0.003	-0.013	0.025	0.019	-0.118***	-0.001	0.033	0.09**
TFIrr_Verb	0.334***	0.472***	-0.096***	0.011	-0.039	0.061	0.008	-0.177***	-0.001	0.06	0.237***
Panel B: Correlation of the main variables with the time frequency – Inverse document frequency of each dictionary											

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation of the different sets of dictionaries weight with the main variables. Proceeds are the total amount raised on IPO. Log proceeds is the log of the proceeds raised from the IPO. The spread is the amount charged by the underwriter as a percentage of the IPO proceeds. Idiosyncratic risk shows the firm specific risk. The STDev is the standard deviation of share price during the first year of trading. Age in Days is the age of the company in days on the day of IPO. Potential growth is a factor showing if the company increased the size of its capital at the years following the IPO. D1 Underpricing is the return on the first day of trading if purchased at the IPO. Overhang is the percentage of shares that are not traded. Pro-IPO Net Income is the reported Net Income in the year prior to the IPO. Pre-IPO Log(Sales) is the log of the reported sales in the year prior to the IPO.

The spread is showing significant correlation with five of the dictionaries in Panel A. A positive correlation with Negative, Litigious, Superfluous and Irregular Verbs. The highest correlation after the Irregular Verbs is Litigious and Negative with a value of 0.086 and 0.068 in the same order. This indicates that underwriters tend to charge more as they disclose more litigious and negative information as a proportion of the whole prospectus. It also shows a significant positive correlation with the Superfluous dictionary. Underwriters use slightly more superfluous words when they charge higher fees.

Age in days before the IPO shows a significant positive correlation with Positive and Interesting. This indicates that prospectuses of older companies show more positive and interesting words than younger companies. It indicates that older companies present more marketing material in their IPO prospectuses compared to the younger companies. There is a significant negative correlation with Litigious and Constraining. This indicates that older companies have fewer constraints and litigious matters to address than other younger companies. Superfluous and Modal are showing a negative correlation as well. That indicate that older companies use fewer words from them.

Underpricing is showing a negative correlation with Negative, Positive and Harvard IV dictionaries' proportional weights. This indicates that more positive or negative information disclosure in a prospectus will result in less underpricing. Moreover, a positive correlation is present with the litigious dictionary. That means using more litigious words will result in more underpricing.

Overhang variable shows a negative correlation with Uncertainty, Constraining and Modal. This indicates that as a percentage of offered shares

decreases, the proportion of Uncertainty, Constraining and Modal words decreases as well. This is due to the fact that the owners are still holding more of the shares which show more trust from their side on the firm.

The Pre-IPO sales variable shows a positive correlation with Positive, Interesting, and Harvard Negative words. For the companies that have reported higher Pre-IPO sales, their prospectuses show slightly more positive and interesting words. In addition, it shows negative correlations with Litigious, Superfluous and Modal.

When looking to Panel B, I notice that Log(Proceeds) shows high positive significant correlation with all the dictionaries. The positive correlation indicates that more information is provided for the larger IPOs. The spread is also showing a high negative significant correlation with all the dictionaries except for the Superfluous and less significance for Modal. This indicates that Underwriters charge more fees when disclosing less information. Age in Days is showing a positive correlation with Negative, Positive Uncertainty, Litigious and Constraining. This shows that older companies show more information than younger ones.

Underpricing is significantly negatively correlated with all the dictionaries. This supports the literature that reported that underwriters compensate for the lack of information by underpricing more. Pre-IPO Sales is significantly and positively correlated to all dictionaries. This shows that a firm that reports pre-IPO sales are disclosing more information than other companies who did not report in per-IPO sales. To further examine the effect of the sentiment used in the prospectus on the spread charged by the underwriters, I ran the same regression model used in Chapter 1 with different dictionaries at a time.

From Table 5.5 I can notice that the coefficients of all the dictionaries are positive. This shows different results compared to what is reported in the correlation in Table 5.4. This can be resulted due to the fact that different segments have different correlation across the sample. To investigate this further, I ran all the regressions for the different markets, and different IPO proceed sizes as shown in Table 5.6.

5.13 Spread and tone

From Table 5.6, when looking at All Market, I can notice that the coefficients of all the dictionaries are positive. Only the Superfluous dictionary shows to be nonsignificant. The highest coefficient is those of the Interesting and Constraining dictionaries. However, I can notice this is significant and positive only on AIM market and Less than 16.6M and slightly less significant for the 16.6M – 66.6M. From Table 5.7, though the correlation is 99% significant and negative for the whole market with a value of -0.134, Constraining is showing to be 90% significant for IPOs with proceeds less than 16.6M with a value of 0.07. As a result, this indicates that the negative value with All Market is mainly coming from the insignificant negative correlation reported for the Main Market and larger IPOs. With respect to the Interesting, Table 5.7 shows that the correlation is only significant for the segment of the sample of the IPOs size between 16.6M – 66.M with a level of significant of 95% and a value of 0.165. From Table 5.6, we can notice that all the significant coefficients have a positive value which support our first hypothesis (H1). We can notice that the significant is more in AIM Market and the less than 16.6M which compromise most of the IPOs in our sample and hence All Market shows some significance as well.

I have also shown the same for the correlation in Table 5.7.

Table 5.5: Spread Regression for all Market showing the effect of the tf.idf

		TFNegative	TFHarvard_IV	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		0.03**									
		(2.46)									
TFHarvard_IV			0.02**								
			(2.21)								
TFConstraining				0.27***							
				(3.43)							
TFLitigious					0.15***						
					(3.49)						
TFSuperfluous						0.45					
						(1.41)					
TFModal							0.15***				
							(3.49)				
TFUncertainty								0.08*			
								(1.77)			
TFPositive									0.12**		
									(2.48)		
TFInteresting										0.38**	
										(2.43)	
TFIrr_Verb											0.22**
											(2.01)
Log(proceeds)	-3.23***	-3.38***	-3.34***	-3.49***	-3.47***	-3.25***	-3.47***	-3.31***	-3.35***	-3.31***	-3.33***
	(-17.69)	(-17.54)	(-17.63)	(-17.73)	(-17.87)	(-17.75)	(-17.87)	(-17.59)	(-17.76)	(-17.87)	(-17.6)
Market Dummy	-1.72***	-1.57***	-1.6***	-1.51***	-1.49***	-1.64***	-1.49***	-1.65***	-1.52***	-1.6***	-1.64***
	(-4.68)	(-4.23)	(-4.3)	(-4.07)	(-4.01)	(-4.42)	(-4.01)	(-4.47)	(-4.03)	(-4.34)	(-4.43)
Multi-bookrunner	0.63	0.4	0.46	0.26	0.3	0.56	0.3	0.49	0.54	0.56	0.51
	(1.41)	(0.89)	(1.01)	(0.57)	(0.65)	(1.25)	(0.65)	(1.08)	(1.22)	(1.25)	(1.14)
Prestigious	0.79***	0.78***	0.79***	0.78***	0.78***	0.81***	0.78***	0.8***	0.76***	0.76***	0.8***
	(3.51)	(3.49)	(3.51)	(3.46)	(3.48)	(3.59)	(3.48)	(3.57)	(3.36)	(3.39)	(3.56)
Day 1 Return	-0.01**	-0.01*	-0.01*	-0.01*	-0.01*	-0.01**	-0.01*	-0.01**	-0.01*	-0.01**	-0.01*
	(-2.09)	(-1.96)	(-1.95)	(-1.92)	(-1.94)	(-1.99)	(-1.94)	(-2.05)	(-1.81)	(-1.98)	(-1.93)
Log(Lockup)	0.73	0.83	0.8	0.87	0.9	0.72	0.9	0.8	0.8	0.82	0.8
	(1.12)	(1.28)	(1.23)	(1.33)	(1.39)	(1.11)	(1.39)	(1.22)	(1.22)	(1.26)	(1.22)
Constant	5.67***	5.19***	5.27***	5***	4.76***	5.59***	4.76***	5.38***	5.11***	5.23***	5.25***
	(3.15)	(2.88)	(2.92)	(2.78)	(2.64)	(3.11)	(2.64)	(2.98)	(2.83)	(2.9)	(2.9)
Observations	945	945	945	945	945	945	945	945	945	945	945
Adjusted R Square	0.3372	0.3332	0.3331	0.3323	0.3375	0.3311	0.3331	0.3311	0.3302	0.3317	0.3375
Prob > F	0	0	0	0	0	0	0	0	0	0	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the spread where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 5.6: Summary of dictionary tf.idf weight coefficient for 70 Spread regressions

	All Market	Main Market	AIM Market	Less than 16.6M	16.6M – 66.6M	66.6M – 333.4M	More than 333.4M
TFNegative	0.03**	0.01	0.05***	0.04**	0.03**	-0.01	-0.06
	(2.46)	(0.51)	(2.88)	(2.22)	(2.02)	(-0.25)	(-1.15)
TFHarvard_IV	0.02**	0.01	0.03**	0.03*	0.03**	0.01	-0.04
	(2.21)	(1)	(2.32)	(1.74)	(2.12)	(0.11)	(-1.1)
TFConstraining	0.27***	-0.06	0.39***	0.41***	0.11	-0.11	-0.37
	(3.43)	(-0.58)	(4.06)	(3.19)	(1.09)	(-1.04)	(-0.73)
TFLitigious	0.15***	0.03	0.2***	0.21***	0.02	0.04	-0.13
	(3.49)	(0.54)	(3.73)	(2.98)	(0.34)	(0.68)	(-0.57)
TFSuperfluous	0.45	0.61*	0.51	0.62	0.3	0.82	-3.36
	(1.41)	(1.78)	(1.28)	(1.34)	(0.76)	(1.49)	(-0.64)
TFModal	0.15***	0.03	0.2***	0.21***	0.02	0.04	-0.13
	(3.49)	(0.54)	(3.73)	(2.98)	(0.34)	(0.68)	(-0.57)
TFUncertainty	0.08*	-0.01	0.1**	0.09	0.14***	-0.11	-0.35
	(1.77)	(-0.09)	(2.04)	(1.37)	(2.65)	(-1.61)	(-0.96)
TFPositive	0.12**	0.06	0.14**	0.16**	0.06	0.07	-0.53
	(2.48)	(1.19)	(2.46)	(2.47)	(1.01)	(0.91)	(-0.6)
TFInteresting	0.38**	0.24	0.53**	0.52**	0.31*	0.14	-0.81
	(2.43)	(1.63)	(2.57)	(2.23)	(1.78)	(0.47)	(-0.44)
TFIrr_Verb	0.22**	-0.06	0.31**	0.31**	0.21	-0.15	-0.63
	(2.01)	(-0.36)	(2.42)	(2.06)	(1.37)	(-0.8)	(-0.92)
Observations	945	158	787	675	162	90	18

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 70 OLS Regressions of the spread. Only the coefficient of the dictionary weight and its level of significance is reported here in this table. The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of lockup period, Market dummy is 0 for the Main market and 1 for AIM market and year dummy variable.

Table 5.7: Spread Correlation with dictionaries for All Sample, Main, AIM and different IPO Proceed sizes

spread	All Market	Main Market	AIM Market	Less than 16.6M	16.6M – 66.6M	66.6M – 333.4M	More than 333.4M
TFNegative	-0.129***	-0.052	-0.046	0.027	0.161**	0.124	-0.059
TFHarvard_iv	-0.111***	-0.01	-0.036	0.022	0.17**	0.148	0.074
TFConstraining	-0.134***	-0.121	-0.036	0.07*	0.106	0.068	-0.145
TFLitigious	-0.137***	-0.083	-0.04	0.077**	0.052	0.145	-0.191
TFSuperfluous	-0.04	0.152*	0.004	0.031	0.089	0.24**	0.33
TFModal	-0.061*	-0.06	-0.008	0.022	0.206***	0.046	0.055
TFUncertainity	-0.087***	-0.065	-0.019	0.034	0.198**	0.042	0.056
TFPositive	-0.128***	0.03	-0.038	0.012	0.107	0.158	0.299
TFInteresting	-0.088***	0.056	-0.024	0.019	0.165**	0.082	0.247
TFIrr_verb	-0.096***	-0.08	-0.024	0.034	0.123	-0.004	0.276

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation of the spread with different sets of the time frequency – inverse document frequency weighted dictionaries for different markets and different sizes in terms of IPO proceeds.

From both Table 5.6 and Table 5.7, I can notice that the usage of superfluous words in the prospectus is related to a higher spread charged by the underwriters in the Main Market. This indicates that underwriters who charge higher fees would use more superfluous words to justify their fees.

Although AIM Market does not show any significant correlation in Table 5.7, yet, the regression reported in Table 5.6 shows significant coefficients across a number of dictionaries. The highest four coefficients in value are those of Interesting, Constraining, Irregular Verbs, Litigious and Modal with values of 0.53, 0.41, 0.31, 0.21 and 0.21, respectively. This indicates that the underwriters in AIM Market put their utmost effort towards disclosing the legal and constraints that would face the company so that the prospectus would be more interesting. Regressions in Table 5.6 for the segment of IPOs with proceeds less than 16.6M are showing very close results to those of the AIM Market. This is due to the fact the most of the IPOs in the AIM Market fall in this segment. However, Table 5.7 shows significance in the correlation with both Litigious and Constraining.

Moving to the next segment where IPO proceed is between 16.6M – 66.6M, I can see that there are four dictionaries showing a level of significance. Interesting words is showing a significant coefficient of 0.31 and a significant correlation of 0.24 as shown in Table 5.7. In addition, Uncertainty is also showing a significant and positive coefficient of 0.14 and 0.198 of significant correlation in Table 5.7. This indicates that underwriters are charging higher fees for IPOs with a higher level of uncertainty. The other 2 dictionaries are the Negative and Harvard IV dictionaries. From Table 5.6 I can notice that the coefficient values for those 2 dictionaries are close. A coefficient of 0.03 for both in Table 5.6 and 0.161 and 0.17 of correlation in

Table 5.7 for Negative and Harvard IV in this order. None of the remaining regressions in Table 5.6 for 66M – 333.4M and More than 333.4M show any significance.

I have introduced the tf.idf weight variables for all the dictionaries on the non-linear spread regression model:

$$Spread = \frac{1}{b_1 + b_2 \log(proceeds)} + b_3\, tfdictionary\, weight + b_4\, Underpricing + b_4\, Return\, Stdev + \varepsilon_i \quad (5.3)$$

From the Table 5.8, I can notice that for All IPOs, AIM and Less than 16.6M of proceeds that the coefficient of all the dictionary is significant and positive. This indicates that underwriters charge higher fees when they report more information. However, I can notice that Superfluous dictionary coefficient has the highest value. This indicates that underwriters will use more superfluous words in the prospectuses to justify the higher spread they are charging. The next higher coefficient is coming from the Interesting dictionary. This shows that prospectuses with marketing tone will cost the companies more in terms of spread. This also supports our first hypothesis (H1)

I also ran the regression for all the segments of our sample. From Table 5.9 I can see that for All Market, AIM Market and Less than 16.6M all have significant and negative coefficients for Irregular Verbs and Positive dictionary. From Table 5.9 I can notice that Positive dictionary coefficient high level of significance is present across all the segment less than 66.6M. The coefficients' values are higher for the larger IPOs than the smaller ones. This indicates that positive words used in the prospectus have better reception from the audience in larger IPOs rather than smaller IPOs. Main

Market and 16.6M - 66.6M is showing significance for Constraining dictionaries. In addition, the Main Market is also showing significance for Litigious, Modal and Negative dictionaries. This indicates that mainly in the Main Market, underwriters would underprice more when providing less information about constraints, negative information and litigious data.

5.14 Underpricing and tone

In Table 5.10 I am examining the effect of the dictionaries weight over underpricing for prestigious and non-prestigious underwriters. Panel A – shows the proportional weight variables coefficients. For the IPOs with prestigious underwriters, I can see that the highest significance is for the Litigious dictionary. With a positive coefficient, I can say that IPO with higher Litigious words tend to be underpriced more. On the other hand, Negative and Positive are both showing significant negative coefficients.

For the non-prestigious underwriters, I can notice that, in addition to Litigious, their IPOs are more affected by the Modal dictionary.

Panel B – shows the tf.idf weight variables coefficients. For the IPOs with prestigious underwriters, I can notice that as the prospectus is showing more conservatism, underpricing gets less. This is explained by the significant negative coefficient of Negative, Harvard IV and Constraining. Positive and Interesting dictionaries weight coefficients are also significant and negative. This shows that for a prestigious underwriter, when they show interesting words of the IPO, then, they have priced it closer to the market value and hence it will be less underpriced.

Table 5.8: Summary of Nonlinear Spread Regression - Dictionary weight coefficient for 70 regressions

	All Market	Main Market	AIM Market	Less than 16.6M	16.6M – 66.6M	66.6M – 333.4M	More than 333.4M
TFNegative	0.021***	0.004	0.035***	0.049***	0.02**	0.024*	-0.003
	(2.96)	(0.45)	(3.4)	(3.51)	(2.06)	(1.7)	(-0.14)
TFHarvard_iv	0.016***	0.006	0.024***	0.032***	0.016**	0.022*	0.005
	(3.02)	(0.95)	(3.21)	(3.14)	(2.18)	(1.98)	(0.38)
TFConstraining	0.139***	-0.028	0.246***	0.422***	0.097	0.085	-0.053
	(3.05)	(-0.42)	(3.73)	(4.41)	(1.31)	(0.98)	(-0.31)
TFLitigious	0.066***	0.005	0.116***	0.215***	0.023	0.078*	-0.048
	(2.66)	(0.14)	(3.14)	(3.97)	(0.58)	(1.67)	(-0.58)
TFSuperfluous	0.691**	0.83**	0.749**	0.901**	0.344	1.397**	1.355*
	(2.47)	(2.55)	(2.04)	(2.02)	(1.03)	(2.53)	(1.93)
TFModal	0.322***	0.023	0.396***	0.412**	0.426***	0.219	0.12
	(2.76)	(0.12)	(2.75)	(2.3)	(2.72)	(0.77)	(0.26)
TFUncertainty	0.091***	0.009	0.125***	0.155***	0.1**	0.056	0.028
	(3.14)	(0.21)	(3.3)	(3.1)	(2.59)	(0.91)	(0.31)
TFPositive	0.082***	0.061	0.125***	0.153***	0.06	0.135**	0.114
	(2.69)	(1.36)	(2.86)	(2.8)	(1.29)	(2.03)	(0.92)
TFInteresting	0.333***	0.194	0.493***	0.553**	0.305**	0.285	0.429
	(2.62)	(1.4)	(2.7)	(2.49)	(2.05)	(1.01)	(1.11)
TFIrr_Verb	0.211***	-0.014	0.305***	0.384***	0.17	0.055	0.258
	(2.82)	(-0.13)	(3.08)	(3.05)	(1.51)	(0.3)	(1.1)
Sample size	946	159	787	675	162	90	19

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 70 Nonlinear Regressions of the spread on different IPO sizes in terms of IPO proceeds. Only the coefficient of the dictionary weight and its level of significance is reported here in this table. The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of lockup period, Market dummy is 0 for the Main market and 1 for AIM market and year dummy variable.

Table 5.9: Summary of dictionary tf.idf weight coefficient for 70 Underpricing regressions

	All Market	Main Market	AIM Market	Less than 16.6M	16.6M – 66.6M	66.6M – 333.4M	More than 333.4M
TFNegative	-0.33	-1.21**	-0.09	-0.39	-1.03*	-0.07	0.01
	(-1.59)	(-1.99)	(-0.38)	(-1.36)	(-1.67)	(-0.18)	(1.13)
TFHarvard_IV	-0.28*	-0.84*	-0.16	-0.34*	-0.72	-0.07	0.01
	(-1.87)	(-1.81)	(-0.99)	(-1.69)	(-1.59)	(-0.21)	(0.43)
TFConstraining	-2.19	-11.4**	0.09	-1.36	-9.17**	-2.54	0.01
	(-1.49)	(-2.49)	(0.06)	(-0.66)	(-2.05)	(-1.09)	(1.19)
TFLitigious	-0.97	-4*	-0.1	-1.04	-3.9	-0.6	0.01
	(-1.27)	(-1.72)	(-0.12)	(-0.95)	(-1.64)	(-0.51)	(0.97)
TFSuperfluous	-13.03**	-29.94	-10.19*	-12.65*	-25.85	-19.01	-0.04
	(-2.19)	(-1.64)	(-1.66)	(-1.73)	(-1.45)	(-1.44)	(-0.23)
TFModal	-0.97	-4*	-0.1	-1.04	-3.9	-0.6	0.01
	(-1.27)	(-1.72)	(-0.12)	(-0.95)	(-1.64)	(-0.51)	(0.97)
TFUncertainty	-0.57	-4.55	-0.08	-0.96	-1.97	1.49	0.01
	(-0.73)	(-1.4)	(-0.11)	(-1.01)	(-0.79)	(0.92)	(0.75)
TFPositive	-2.8***	-5.33**	-2.12**	-3.08***	-6.44**	0.38	0.01
	(-3.32)	(-2.1)	(-2.46)	(-3.02)	(-2.46)	(0.23)	(0.26)
TFInteresting	-3.79	-12.14	-0.59	-1.47	-11.79	-9.23	0.01
	(-1.3)	(-1.58)	(-0.18)	(-0.4)	(-1.49)	(-1.43)	(0.22)
TFIrr_Verb	-4.65**	-10.05	-3.98**	-5.52**	-9.78	-0.56	-0.01
	(-2.28)	(-1.31)	(-1.99)	(-2.32)	(-1.39)	(-0.12)	(-0.11)
Observations	945	158	787	675	162	90	18

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 70 OLS Regressions of the Underpricing on different IPO sizes in terms of IPO proceeds. Only the coefficient of the dictionary weight and its level of significance is reported here in this table. The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of lockup period, Market dummy is 0 for the Main market and 1 for AIM market and year dummy variable.

Table 5.10: Summary of Underpricing regressions showing dictionary coefficients for prestigious and non-prestigious underwriters

Prestigious	PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
Coef.	-2539.115**	-184.955	1012.192	4138.239***	-12414.24	-798.71	-150.335	-3637.12**	-128.869	1524.941
	(-2.01)	(-0.34)	(1.11)	(3.71)	(-0.72)	(-0.78)	(-0.11)	(-2.08)	(-0.07)	(0.69)
A.R2	0.034	0.0254	0.0278	0.0546	0.0263	0.0265	0.0252	0.0346	0.0251	0.0262
Prob F	0.0011	0.0058	0.0036	0	0.0049	0.0047	0.0061	0.001	0.0061	0.005
Non-Prestigious										
Coef.	-48.175	-101.512	-99.014	2915.94***	3349.19	1717.7**	1528.19	-27.849	-72.51	3152.388*
	(-0.05)	(-0.23)	(-0.13)	(3.87)	(0.25)	(2.35)	(1.48)	(-0.02)	(-0.05)	(1.75)
A.R2	0.2256	0.2257	0.2256	0.2485	0.2257	0.2342	0.2291	0.2256	0.2256	0.2304
Prob F	0	0	0	0	0	0	0	0	0	0

Panel A: Dictionary proportional weight variables coefficient for prestigious and non-prestigious underwriters

Prestigious	TFNegative	TFHarvard_iv	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
Coef.	-0.464*	-0.36*	-3.575**	-1.496	-13.796	-7.919*	-1.205	-2.725**	-6.399*	-4.794*
	(-1.9)	(-1.94)	(-2.06)	(-1.58)	(-1.56)	(-1.87)	(-1.16)	(-2.53)	(-1.77)	(-1.71)
A.R2	0.033	0.0334	0.0344	0.0306	0.0305	0.0328	0.0281	0.039	0.032	0.0316
Prob F	0.0013	0.0012	0.001	0.0021	0.0022	0.0014	0.0034	0.0004	0.0016	0.0018
Non-Prestigious										
Coef.	-0.601**	-0.439**	-3.594**	-1.882**	-13.425*	-6.692**	-1.558*	-3.299***	-2.558	-6.228***
	(-2.49)	(-2.55)	(-2.16)	(-2.09)	(-1.95)	(-2.31)	(-1.92)	(-3.19)	(-0.59)	(-2.89)
A.R2	0.2353	0.2357	0.2329	0.2325	0.2315	0.2339	0.2314	0.2413	0.2262	0.2385
Prob F	0	0	0	0	0	0	0	0	0	0

Panel B: Dictionary tf.idf weight variables coefficient for prestigious and non-prestigious underwriters

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 40 OLS Regressions of the spread. Only the coefficient of the dictionary weight and its level of significance is reported here in this table. The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of lockup period, Market dummy is 0 for the Main market and 1 for AIM market and year dummy variable.

For the non-prestigious underwriters, Table 5.10 also shows that conservatism yields in less underwriting. In comparison to the prestigious underwriters, I can see that litigious is significant for the non-prestigious underwriters' IPOs. All other variables are showing significant and negative coefficients. From the previous underpricing tables, it can be noticed that for most of the significant coefficients, the value is negative. This goes in line with our hypothesis (H2) for the Postive tone part. However, we can not accept the hypothesis (H2) as values of the Conservative and Uncertainty tones are negative as well. In this case, the negative relationship has to do with the level of exposure where more information availability would reduce the underpricing.

5.15 Lockup length and tone

Table 5.11 shows the results of the Log(Lockup) regression of All Market introducing the dictionaries tf.idf weight as follows:

$$\begin{aligned} LogLockupPeriod \\ &= \beta_1 tf dictionary\ weight + \beta_2 Log\ Proceeds_i \\ &+ \beta_3 Market_i + \beta_4 MultiBookru_i \\ &+ \beta_5 Prestigious_i + \beta_6 D1\ Underpricing_i \\ &+ \left(\sum_{j=1999}^{2012} \beta_j Year_j \right) + \varepsilon_i \end{aligned} \tag{5.4}$$

I can see the coefficient is low in value. However, it is showing a high significance for the Litigious dictionary with a negative coefficient. This means that IPOs with more litigious language in their IPO prospectus will get less lockup period. The same applies on Uncertainty, Interesting, Constraining and Negative dictionaries with lower coefficient values.

To further examine the effect of the prospectus language on lockup period, I ran 70 regressions for each of the dictionaries weight for the different sample segments. As I can see in Table 5.12, Litigious and Uncertainty disclosure show a significant negative coefficient of a low value for All Market and AIM Market. The Interesting dictionary is showing significance with a negative coefficient on All Market and 66.6M – 333.4M of proceeds segment.

Table 5.11: Lockup period regression showing the effect of the dictionary weight variable

		TFNegative	TFHarvard_IV	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		-0.001*									
		(-1.93)									
TFHarvard_IV			-0.001								
			(-1.52)								
TFConstraining				-0.007*							
				(-1.84)							
TFLitigious					-0.005**						
					(-2.30)						
TFSuperfluous						0.006					
						(0.36)					
TFModal							-0.006				
							(-0.77)				
TFUncertainty								-0.004*			
								(-1.72)			
TFPositive									-0.003		
									(-1.21)		
TFInteresting										-0.014*	
										(-1.76)	
TFIrr_Verb											-0.008
											(-1.52)
Log(proceeds)	-0.004	0.002	1.2E-4	0.003	0.004	-0.004	-0.003	0	-0.001	-0.001	0
	(-0.44)	(0.23)	(0.01)	(0.31)	(0.41)	(-0.47)	(-0.28)	(0.01)	(-0.1)	(-0.08)	(-0.01)
Market Dummy	0.063***	0.057***	0.058***	0.057***	0.055***	0.064***	0.062***	0.059***	0.058***	0.058***	0.059***
	(3.39)	(3.02)	(3.11)	(3.03)	(2.92)	(3.41)	(3.32)	(3.18)	(3.03)	(3.13)	(3.19)
Multi-bookrunner	-0.155***	-0.146***	-0.149***	-0.145***	-0.143***	-0.156***	-0.153***	-0.148***	-0.153***	-0.152***	-0.15***
	(-7.16)	(-6.56)	(-6.74)	(-6.46)	(-6.44)	(-7.15)	(-7.00)	(-6.69)	(-7.02)	(-7)	(-6.86)
prestigious	-0.006	-0.005	-0.006	-0.005	-0.005	-0.005	-0.006	-0.006	-0.005	-0.005	-0.006
	(-0.5)	(-0.47)	(-0.49)	(-0.46)	(-0.47)	(-0.48)	(-0.54)	(-0.55)	(-0.43)	(-0.41)	(-0.53)
Underpricing	-8.1E-05	-9E-05	-8.9E-05	-8.9E-05	-9E-05	-7.9E-05	-8.5E-05	-8.5E-05	-9.3E-05	-8.8E-05	-9.1E-05
	(-0.91)	(-1.02)	(-1.01)	(-1.01)	(-1.02)	(-0.89)	(-0.96)	(-0.96)	(-1.04)	(-0.99)	(-1.03)
Constant	2.593***	2.601***	2.6***	2.601***	2.608***	2.591***	2.596***	2.598***	2.602***	2.6***	2.602***
	(79.96)	(79.6)	(79.35)	(79.48)	(78.91)	(79.3)	(79.13)	(79.81)	(77.97)	(79.63)	(78.8)
Observations	945	945	945	945	945	945	945	945	945	945	945
Adjusted R Square	0.1184	0.121	0.1197	0.1207	0.1225	0.1176	0.118	0.1203	0.1189	0.1204	0.1197
Prob > F	0	0	0	0	0	0	0	0	0	0	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Lockup period where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 5.12: Summary of Lockup Period Regression - Dictionary weight coefficient for 70 regressions

	All Market	Main Market	AIM Market	Less than 16.6M	16.6M – 66.6M	66.6M – 333.4M	More than 333.4M
TFNegative	-0.001*	-0.001	-0.001*	-0.001	-0.001	-0.002	-0.002
	(-1.93)	(-0.6)	(-1.93)	(-1.16)	(-0.52)	(-0.97)	(-1.39)
TFHarvard_IV	-0.001	-0.001	-0.001	-3.9E-4	-4.5E-4	-0.001	-0.001
	(-1.52)	(-0.64)	(-1.51)	(-0.69)	(-0.57)	(-0.9)	(-1.61)
TFConstraining	-0.007*	-0.009	-0.006	0.002	-0.008	-0.018	-0.015
	(-1.84)	(-0.96)	(-1.32)	(0.42)	(-1.03)	(-1.66)	(-0.86)
TFLitigious	-0.005**	-0.001	-0.005**	-0.005	-0.004	-0.001	-0.006
	(-2.3)	(-0.3)	(-2.19)	(-1.45)	(-1.04)	(-0.23)	(-0.86)
TFSuperfluous	0.006	0.047	-0.004	0.002	0.025	-0.005	-0.094***
	(0.36)	(1.35)	(-0.23)	(0.1)	(0.79)	(-0.07)	(-6.37)
TFModal	-0.006	0.009	-0.008	-0.003	-0.006	-0.005	-0.067
	(-0.77)	(0.4)	(-1)	(-0.32)	(-0.38)	(-0.15)	(-1.75)
TFUncertainty	-0.004*	-0.002	-0.004*	-0.002	-0.002	-0.008	-0.014
	(-1.72)	(-0.29)	(-1.88)	(-0.81)	(-0.4)	(-1.05)	(-1.97)
TFPositive	-0.003	-0.001	-0.003	-0.002	0.002	-0.008	-0.019**
	(-1.21)	(-0.17)	(-1.3)	(-0.62)	(0.34)	(-1)	(-4.95)
TFInteresting	-0.014*	-0.011	-0.015	-0.005	-0.007	-0.073**	-0.059*
	(-1.76)	(-0.74)	(-1.59)	(-0.51)	(-0.51)	(-2.45)	(-2.42)
TFIrr_Verb	-0.008	-0.016	-0.009	-0.005	-0.009	-0.022	-0.026
	(-1.52)	(-1.08)	(-1.58)	(-0.78)	(-0.72)	(-1.05)	(-2.26)
Observations	945	158	787	675	162	90	18

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 70 OLS Regressions of the Lockup period on different IPO sizes in terms of IPO proceeds. Only the coefficient of the dictionary weight and its level of significance is reported here in this table. The control Variables are the log of the proceeds adjusted for CPI index, Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, Log(Lockup) is the log of lockup period, Market dummy is 0 for the Main market and 1 for AIM market and year dummy variable.

Table 5.13: Underpricing Regression for All Market showing the effect of the tf.idf

		TFNegative	TFHarvard_IV	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		-0.33									
		(-1.59)									
TFHarvard_IV			-0.28*								
			(-1.87)								
TFConstraining				-2.19							
				(-1.49)							
TFLitigious					-0.97						
					(-1.27)						
TFSuperfluous						-13.03**					
						(-2.19)					
TFModal							-0.97				
							(-1.27)				
TFUncertainty								-0.57			
								(-0.73)			
TFPositive									-2.8***		
									(-3.32)		
TFInteresting										-3.79	
										(-1.3)	
TFIrr_Verb											-4.65**
											(-2.28)
Log(proceeds)	-17.04***	-15.08***	-15.1***	-14.87***	-15.37***	-16.25***	-15.37***	-16.39***	-13.77***	-16.13***	-14.77***
	(-5.05)	(-4.2)	(-4.28)	(-4.04)	(-4.24)	(-4.79)	(-4.24)	(-4.69)	(-3.93)	(-4.67)	(-4.2)
Market Dummy	-19.73***	-21.57***	-21.71***	-21.5***	-21.35***	-21.83***	-21.35***	-20.29***	-24.7***	-20.92***	-21.49***
	(-2.89)	(-3.12)	(-3.15)	(-3.11)	(-3.08)	(-3.18)	(-3.08)	(-2.96)	(-3.56)	(-3.04)	(-3.14)
Multi-bookrunner	-6.17	-3.16	-3.2	-2.93	-3.64	-4.19	-3.64	-4.99	-3.66	-5.25	-3.39
	(-0.77)	(-0.38)	(-0.39)	(-0.35)	(-0.44)	(-0.52)	(-0.44)	(-0.61)	(-0.46)	(-0.65)	(-0.42)
Prestigious	-1.49	-1.38	-1.45	-1.36	-1.4	-2	-1.4	-1.57	-0.62	-1.19	-1.7
	(-0.35)	(-0.33)	(-0.34)	(-0.32)	(-0.33)	(-0.47)	(-0.33)	(-0.37)	(-0.15)	(-0.28)	(-0.4)
Constant	52.83***	55.29***	56***	55.32***	55.89***	55.7***	55.89***	53.68***	61.88***	54.71***	57.9***
	(4.44)	(4.61)	(4.67)	(4.61)	(4.6)	(4.66)	(4.6)	(4.49)	(5.1)	(4.57)	(4.79)
Observations	946	946	946	946	946	946	946	946	946	946	946
Adjusted R Square	0.0469	0.0485	0.0494	0.0481	0.0475	0.0507	0.0475	0.0464	0.0571	0.0476	0.0512
Prob > F	0	0	0	0	0	0	0	0	0	0	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of Underpricing where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 5.13 shows the Underpricing regression model I used in Chapter 2 introducing the dictionaries one at a time. The table shows that the coefficient is negative for all the dictionaries. This is in line with the unilateral correlation of the Underpricing variable showed in Table 5.4. This supports the literature that Underpricing is less observed as more information is provided. Positive dictionary coefficient shows a value of -2.8 with high significance. Superfluous dictionary coefficient is showing the highest value of 13.03 followed by the dictionary of the Irregular Verbs and then Harvard IV with a value of -0.28. This indicates that the prospectus tone is not representing the the risk as it is not being reflected in a higher lockup period. This is too an indication that the tone indicate higher level of information exposure and hence lower lockup period. We can not accept the third hypothesis (H3).

5.16 Volatility and tone

From Table 5.14 It is noticeable that non of the dictionaries coefficient is significant. So is the case for the subsegment of the sample. We can not find a relationship between the prospectus tone and the volatility after the IPO. We can not accept our fourth hypothesis (H4).

5.17 Idiosyncratic and tone

From Table 5.15 It is also noticeable that non of the dictionaries coefficient is significant. So is the case for the subsegment of the sample. We can not find a relatsionship between the prospectus tone and the company specifc risk. We can not accept our fifth hypothesis (H5).

Table 5.14: Share Volatility Regression for all Market showing the effect of the tf.idf

Share Volatility		TFNegative	TFHarvard_IV	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		0.723									
		(0.11)									
TFHarvard_IV			-0.134								
			(-0.03)								
TFConstraining				9.048							
				(0.2)							
TFLitigious					-3.114						
					(-0.13)						
TFSuperfluous						135.214					
						(0.72)					
TFModal							-26.733				
							(-0.3)				
TFUncertainty								2.568			
								(0.11)			
TFPositive									-9.062		
									(-0.34)		
TFInteresting										22.167	
										(0.24)	
TFIrr_Verb											-27.349
											(-0.42)
Log(proceeds)	-50.809	-54.985	-49.913	-59.637	-45.561	-58.212	-44.518	-53.722	-40.884	-55.963	-38.001
	(-0.47)	(-0.48)	(-0.45)	(-0.51)	(-0.4)	(-0.54)	(-0.41)	(-0.48)	(-0.37)	(-0.51)	(-0.34)
Market Dummy	150.207	154.11	149.271	157.342	145.238	173.289	144.942	152.639	133.731	156.907	139.737
	(0.7)	(0.7)	(0.68)	(0.72)	(0.66)	(0.79)	(0.67)	(0.7)	(0.6)	(0.72)	(0.64)
Multi-bookrunner	-70.889	-76.81	-69.592	-83.246	-63.697	-91.89	-61.332	-75.838	-63.964	-75.012	-56.046
	(-0.27)	(-0.29)	(-0.26)	(-0.31)	(-0.24)	(-0.35)	(-0.24)	(-0.29)	(-0.25)	(-0.29)	(-0.21)
Prestigious	178.517	178.322	178.529	178.027	178.753	183.891	176.217	178.929	181.223	176.853	177.186
	(1.35)	(1.35)	(1.35)	(1.35)	(1.35)	(1.39)	(1.33)	(1.35)	(1.37)	(1.34)	(1.34)
Day 1 Return	-0.225	-0.219	-0.227	-0.215	-0.232	-0.173	-0.246	-0.223	-0.264	-0.214	-0.259
	(-0.22)	(-0.21)	(-0.22)	(-0.21)	(-0.22)	(-0.17)	(-0.24)	(-0.22)	(-0.26)	(-0.21)	(-0.25)
Log(Lockup)	-226.244	-223.577	-226.787	-221.746	-229.943	-229.467	-229.176	-223.96	-231.419	-220.925	-234.36
	(-0.59)	(-0.58)	(-0.59)	(-0.58)	(-0.6)	(-0.6)	(-0.6)	(-0.59)	(-0.61)	(-0.58)	(-0.61)
Constant	2205.638**	2193.02**	2208.657**	2183.143**	2225.36**	2181.263**	2230.865**	2195.679**	2250.414**	2180.252**	2258.284**
	(2.08)	(2.06)	(2.07)	(2.05)	(2.08)	(2.06)	(2.1)	(2.06)	(2.11)	(2.05)	(2.12)
Observations	945	945	945	945	945	945	945	945	945	945	945
Adjusted R Square	0.0158	0.0147	0.0147	0.0148	0.0147	0.0153	0.0148	0.0147	0.0148	0.0148	0.0149
Prob > F	0.0193	0.0274	0.0275	0.0272	0.0274	0.024	0.0268	0.0274	0.0267	0.0271	0.0262

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Share volatility where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 5.15: Idiosyncratic Regression for all Market showing the effect of the tf.idf

Idiosyncratic		TFNegative	TFHarvard_IV	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		-0.195									
		(-0.79)									
TFHarvard_IV			-0.124								
			(-0.7)								
TFConstraining				-0.5							
				(-0.29)							
TFLitigious					-0.135						
					(-0.15)						
TFSuperfluous						-8.308					
						(-1.17)					
TFModal							0.315				
							(0.09)				
TFUncertainty								-0.033			
								(-0.04)			
TFPositive									0.563		
									(0.56)		
TFInteresting										-0.252	
										(-0.07)	
TFIrr_Verb											-0.982
											(-0.4)
Log(proceeds)	0.408	1.534	1.24	0.895	0.634	0.862	0.334	0.445	-0.21	0.466	0.867
	(0.1)	(0.36)	(0.29)	(0.2)	(0.15)	(0.21)	(0.08)	(0.11)	(-0.05)	(0.11)	(0.21)
Market Dummy	-0.076	-1.129	-0.945	-0.471	-0.291	-1.494	-0.014	-0.108	0.948	-0.152	-0.452
	(-0.01)	(-0.14)	(-0.11)	(-0.06)	(-0.04)	(-0.18)	(0)	(-0.01)	(0.11)	(-0.02)	(-0.05)
Multi-bookrunner	-3.036	-1.439	-1.83	-2.353	-2.725	-1.745	-3.148	-2.972	-3.466	-2.989	-2.503
	(-0.31)	(-0.14)	(-0.18)	(-0.23)	(-0.27)	(-0.18)	(-0.32)	(-0.3)	(-0.35)	(-0.3)	(-0.25)
Prestigious	-1.985	-1.932	-1.974	-1.958	-1.975	-2.315	-1.958	-1.99	-2.153	-1.966	-2.033
	(-0.4)	(-0.39)	(-0.39)	(-0.39)	(-0.39)	(-0.46)	(-0.39)	(-0.4)	(-0.43)	(-0.39)	(-0.41)
Day 1 Return	0.001	-0.001	-0.001	0	0.001	-0.002	0.001	0.001	0.003	0.001	0
	(0.02)	(-0.02)	(-0.02)	(0.01)	(0.02)	(-0.06)	(0.03)	(0.02)	(0.08)	(0.02)	(-0.01)
Log(Lockup)	2.983	2.263	2.479	2.734	2.823	3.181	3.017	2.953	3.304	2.922	2.691
	(0.21)	(0.16)	(0.17)	(0.19)	(0.19)	(0.22)	(0.21)	(0.2)	(0.23)	(0.2)	(0.19)
Constant	-5.698	-2.294	-2.894	-4.455	-4.845	-4.2	-5.995	-5.569	-8.482	-5.409	-3.808
	(-0.14)	(-0.06)	(-0.07)	(-0.11)	(-0.12)	(-0.1)	(-0.15)	(-0.14)	(-0.21)	(-0.13)	(-0.09)
Observations	945	945	945	945	945	945	945	945	945	945	945
Adjusted R Square	0.0164	0.016	0.0159	0.0155	0.0154	0.0168	0.0154	0.0154	0.0157	0.0154	0.0155
Prob > F	0.0163	0.0198	0.0206	0.0229	0.0232	0.0162	0.0233	0.0234	0.0215	0.0233	0.0224

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Idiosyncratic risk where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

5.18 Conclusion

In this chapter, I measure the IPO tone using textual analysis and relate it to a number of short-run IPO dynamics such as IPO underpricing, spread, lockup length, volatility and idiosyncratic risk using hand-collected data from London Stock Exchange. Though there have been a few studies on textual analysis and IPO short-run dynamics in the US, to the best of our knowledge, this is the first study to analyse IPO dynamics using textual analysis in the UK.

The UK IPO market structure is different from the US in a number of ways. In the US, book building is the preferred method while in the UK, placements and open offer are the methods in use. IPO gross spreads are clustered at 7% in the US. In comparison, there is a large variation in IPO spread in the UK, ranging from 4-11%. Lockup contracts are standardised at 180 days in the US, while heterogeneous in the UK. Also, price support is very common in the US, but not in the UK. Because of these differences, I believe, UK serves as an out of sample test on the issues tested in the US using textual data from IPO prospectuses.

I found a significant relationship with IPO tone and spread implying that underwriters charge more for more disclosures. I also find a significant relationship between IPO tone and underpricing and IPO tone and lockup length. These findings imply certain words in the IPO prospectuses reduce information asymmetry, particularly the negative words. These findings are in line with the earlier works done by using information asymmetry as a major reason for underpricing. To the best of our knowledge, there is no study which relates IPO tone with lockup length. I found that certain words significantly related to the lockup length. This is consistent with the earlier work done on lockups in the US and the UK (e.g. , Brav and Gompers, 2003,

in the US; Hoque, 2011, in the UK) However, I did not find any relationship between IPO tone and volatility or idiosyncratic risk. I propose few reasons for that. When shares start trading after the IPO, may be new information become more important than the information from prospectuses. Also, the volatility and idiosyncratic risks might not be good measures of ex-ante uncertainty of IPO risk. These issues are agendas for future research.

Chapter 6

6. Long-run IPO dynamics and value of textual analysis

6.1 Abstract

I examined some of the puzzles in the IPO literature with the help of information from prospectuses. I used Loughrun and McDonald (2011) modified word list to relate to IPO lockup expiration, long-run IPO return and survival of the IPOs. Though previous literature has use IPO tone measures to relate it to various short-run measures, I am the first to relate it to various long-run issues in the IPO literature. Our tone measures show a significant relationship with the survival of the firms and time till dead. I showed that the information disclosed at the time of IPO is still valid for predicting the survival of IPO firms. However, I do not find much correlation with the IPO tone measured and lockup expiration returns and long-run IPO performance. This might be the result of the methodological controversies of measuring the long-run performance of the IPO.

Keywords: London Stock Exchange, Textual analysis, Lockup expiration returns, long-run IPO performance, survival analysis.

JEL Code: G14, G24, G30, G32

6.2 Introduction

The first section of this chapter is the abstract of the work presented in this chapter. The second section is an introduction. Then on the third section the literature is discussed which include the hypotheses development. Data and methodology come on the fourth section then on the fifth chapter I discussed the results. I concluded in the sixth section.

In the past 40 years, lots of research has been done on various issues of IPOs. Yet, there are some puzzles remain with respect to long-run performance and design of lockups. There has been a number of theories to explain long-run underperformance. Loughran and Ritter (1995, 2000), Baker and Wurgler (2000) and Hirshleifer (2001) put forward behavioural explanations, Aggarwal and Rivoli (1990) put prospect theory as an explanation of long-run underperformance. Long-run underperformance may arise from a combination of a number of factors such as extreme differences of heterogeneous beliefs among investors, short selling is expensive for this small floats on many IPOs (see Ritter and Welch (2002) for a review). Still, it is not clear the reasons for underperformance. Also, the drop in price around the lockup expiration date is another puzzling issue in the IPO literature. Though there have been theories of asymmetric information and moral hazard (Brav and Gompers, 2003 and Hoque and Lasfer, 2016) to explain the existence and length of lockup, the evidence is not conclusive. In this chapter, I examine some of the puzzling issues over the long-run in the IPO literature using textual tones.

Survival of the IPOs is another important issue that has been examined in the literature. For example, Espenlaub et al. (2012) examine whether the IPOs that join the Alternative Investment Market with more prestigious underwriters survives longer and they find evidence in support of that. Joining the market with prestigious underwriter could mitigate the information asymmetry, or it might be possible that

prestigious underwriters bring good quality IPOs. Ahmad and Jelic (2014) analyse the impact of lockup agreements on the survival of UK Initial Public Offerings (IPOs) during the time 1990–2011. Hensler et al., (1997) examine the effects of several firm characteristics and IPO characteristics on the survival of IPO firms by using an accelerated failure time (AFT) model. They find that survival time decreases with the increases of the number of risk factors mentioned in the prospectus. Demers and Joos (2007) develop an IPO failure risk model by incorporating accounting and other deal specific characteristics.

Textual analysis has been utilised in an increasing number of papers and research in Accounting and Finance to quantify the tone and sentiment of corporate news releases, Management discussions and analysis (MD&A), yearly reports (10K reports), daily paper articles, and investor message sheets. Examples incorporate Engelberg (2008), Li (2008), and Tetlock, Saar-Tsechansky, and Macskassy (2008), Tetlock (2007) Antweiler and Frank (2004). The outcomes in these studies could be summarised as negative words could be a successful method for measuring tone as negative words demonstrate a critical relationship with financial variables.

Different papers look at the tone of different documents with the finance related variables. For example, Kothari, Li, and Short (2008) relate the tone of daily paper articles on the cost of capital, return variability, and forecasts from analysts. Henry (2008), Engelberg (2008), and Demers and Vega (2008) relates news discharges with firm income, earnings drift, or stock returns. Some papers relate the information content of IPO prospectus with the share returns, price variability and trading volume (e.g., Loughrun and McDonald (2013) Li (2008, 2009), Feldman et al. (2008), Hanley and Hoberg (2010)).

In this chapter, I examine the effect of the text used in the IPO prospectuses on the long-run IPO dynamics of the newly listed companies on the London Stock Exchange. In particular, I use the lockup expiration returns, long-run underperformance and survival of the IPOs to test for the long-run effect. I do so to shed light on the puzzles that remain in the IPO literature in the long-run.

All previous studies that relate textual analysis with IPO concepts are based on IPO short-run dynamics with the tone measure of IPO prospectus. To the best of our knowledge, there is no study which relates the tone measures of IPO prospectus with the long-run IPO dynamics such as lockup expiration returns, long-run underperformance and survival of the IPOs. I use the UK data to examine the information content of an IPO prospectus. I fill this gap in the literature.

The UK and US IPO markets are very different. For example, the US market uses book building method and the UK uses open offers and placing method. In the book building procedure, information is gathered from the investors and reliance on the underwriters is less. On the other hand, open offers and placing method rely heavily on the underwriters' due diligence. Hence, methods of IPO (e.g., book building versus open offer and placing) has direct implications on the underperformance and survival of IPOs. Lockup length in the US is very standardised, and the median lockup length is 180 days, and most of the IPOs has 180 days of lockups (Field and Hanka, 2001; Brav and Gompers, 2003). In the UK, the lockup length is very diverse and heterogeneous (Hoque, 2011). The average lockup length in the UK is 365 days, and the highest is 1080 days. Also, there are few different types of lockups, absolute versus relative date, one versus gradual release of the shares (Hoque, 2011). Since the US lockup length is shorter, not much information is produced during the lockup length.

However, in the UK more information is produced and disclosed as the lockup lengths are longer. Thus the longer length of the lockups in the UK might have implications on the IPO prospectuses. Also, there is a quiet period in the US which is 40 days after the IPO (before 2002, it was 25 days). There is no quiet period in the UK. During the longer lockup periods, since more information is produced and disclosed after the IPO, asymmetric information is less of a problem in the UK. Hence, the tone measure at the time of IPO may be not related to the lockup expiration returns in the UK.

I do not find any relationship between IPO tone measures and lockup expiration returns (with the exception that Fama-French model return is negatively related to the positive tone measures). This might be related to the unique market settings in the UK. The lockup length is longer in the UK as I find that the average lockup length is 388 days with a minimum of 85 days and maximum of 1260 days. This is consistent with the previous evidence in the UK (e.g., Hoque 2011). Since much information is produced and disclosed during this time period, may be the information disclosed in the IPO prospectus become stale. The prestigious underwriter is negatively related to the lockup expiration return. This implies that prestigious underwriters mitigate the information asymmetry more as compared to the IPO tone reflected at the time of IPO.

I find very little significance when I use the IPO tone measures to explain the long-run returns in case of univariate regressions. However, in the multivariate setting they disappear. However, spread and underpricing are both negatively related to the long-term returns. Higher the spread lower the long-term returns. This might be consistent with the notion that bookrunners charge higher fees for the highly risky companies and they underperform in the long-run. This is consistent our previous

findings. I find that the higher the underpricing, the lower the long-run returns. This is consistent with the previous literature on long-run returns (Hoque and Lasfer, 2015). This is also consistent with the fact that IPOs are underpriced initially and as the divergence of the opinions decrease valuations are corrected, and hence they show underperformance (Morris, 1996).

I find a significant relationship with the IPO prospectus tone measure and survival of the UK IPOs. I use time till dead and dead dummy to relate it to the IPO prospectus tones. Our Tobit regression on time till dead and Logit and Cox hazard model on dead dummy show that IPO prospectus tone measures are significantly related to the survival of the firms. I also find that time till dead is negatively related to spread and multi bookrunners. It shows that risky companies would have paid higher fees and hence they survive less time and also risky companies would have come to the market with multiple bookrunners. The higher the lockup period, the higher the time companies survive. This is consistent with the previous paper by Ahmad and Jelic (2014).

I contributed to the IPO literature in various ways. There has been a longstanding puzzle in the IPO literature regarding underperformance, survival and lockup expiration returns. This chapter tries to explain the underperformance, lockup expiration returns a survival of IPOs with the tones from prospectuses. This is related to the literature on tones of IPOs and 10-K using the US data (see Loughrun and McDonald, 2011, Hanley and Hoberg, 2010). However, no one tried to relate the IPO tone measures with the long-run puzzles in the literature. I also tested whether the lockup expiration return is related to the asymmetric information problem using the IPO tone. Previous papers find that lockups serve to mitigate the information

asymmetry and they reduce moral hazard problems. Using the tone information from IPO prospectuses, I did not find much evidence that IPO lockup expiration returns are driven by information asymmetry. This chapter is related to a number of lockup papers in the US (Field and Hanka, 2001, Brav and Gompers, 2003, Brau et al., 2005) and the UK (Hoque, 2011, Hoque, 2015, Hoque and Lasfer, 2016).

I marry two streams of literature in the IPO. The first one is the survival of the IPOs and the second one is IPO tone measures. Demers and Joos (2007) developed an IPO failure risk model by incorporating accounting and other deal specific characteristics. While they use several accounting variables at the time of IPOs to distinguish between delisted and non-delisted IPO firms I used the IPO tone measure from prospectus to do examine whether the IPO prospectus tone could be related to the survival of IPOs. Hensler et al., (1997) measure the variable risk from Jay Ritter's website which is mentioned as a risk. The risk variable is simply the count of risk factors reported in the prospectus. If there is no specific risk factor section, the variable risk is zero. However, the measure of the risk factor is problematic as it just counts the number of risk factors. I used the Loughrun and McDonald (2011) dictionary measures in the prospectus to relate it to the survival firms of the IPOs.

The rest of the paper is organised as follows. In Section 2 I present a literature review and develops hypotheses. Section 3 describes the data and methodology. Section 4 presents the empirical results. Finally, section five concludes.

6.3 Literature review and Hypotheses development

6.4 IPO lockup expiration and tone

Field and Hanka (2001) find a statistically significant -1.5% drop in share price around lockup expiration date using US data over the 1988-1997 period. This drop is much

larger for venture capital (VC) backed firms as compared to non-VC-backed firms. This constitutes a puzzle as the number of shares locked up and the lockup date are a well-known event then I do not expect any market reaction to that. However, since then a number of studies have examined the reasons for a significant price drop around the IPO lockup expiration (e.g., Field and Hanka (2001), Hoque (2011), Hoque (2014)).

The role of lockups in IPO procedure has been examined by several papers such as Brav and Gompers (2003) and Brau, Lambson and McQueen (2005). Brav and Gompers (2003) propose three contending hypotheses (i) lockups signal firms' quality, (ii) lockups are commitment device, and (iii) lockups are rent-seeking mechanisms by underwriters. Their results give backing that lockups are commitment device, however, dismiss that they serve as a signal of firm quality and they are rent-seeking mechanisms. They find that of firms' lockup insider holdings for more time if the potential for moral hazard is high. Hoque (2014) also provide support for this. Brau et al. (2005) revisited these findings and gave backing to that lockups are commitment device, and they are signals of firm quality. They demonstrate that Brav and Gompers (2003) confirmation of an opposite relationship amongst straightforwardness and lockup length underpins the signal of firm quality as well as the commitment hypothesis. Later on, Yung and Zender (2010) provide some clarification with respect to the opposing finding in Brav and Gompers (2003) and Brau et al. (2005). Yung and Zender (2010) classify the IPO firms as dominated by moral hazard and some dominated by information asymmetry. The firms who joins the market with prestigious underwriters mitigate the information asymmetry, hence, they have a high moral hazard. The firms who comes to the market with other underwriters are subject

to high information asymmetry. They provide empirical support that lockups mitigate asymmetric information for some firms and moral hazard for some firms.

In line with Yung and Zender (2010), Hoque (2014) analyses the role of asymmetric information and moral hazard on IPO underpricing and lockups using the data from the UK market where the lockup lengths are longer and more diverse. The study document that high information asymmetry is related to underpricing while the lockup length and lockup expiration return is related to moral hazard. Hoque (2014) compares high information asymmetry firms (small firms, IPOs underwritten by low-ranked underwriters and AIM firms) with high potential for moral hazard firms (high director ownership firms), he is able to determine if lockups are more closely associated with information asymmetry or moral hazard. Furthermore, the study documents that the relationship between director ownership and lockup expiration returns and lockup length in non-linear.

Information asymmetries are high at the time of IPO because little is known about these new companies. In order to reduce the information asymmetry amongst managers and shareholders, lockup contracts do not permit managers to sell their holdings during or before the lockup lapses. As such, lockup contracts exist so that the insiders can't exploit outside shareholders (Brav and Gompers, 2003). Lockup contracts secure the general investors as Ibbotson and Ritter (1995) contended that investors are ready to pay more for a firm with longer lockup as any negative data being withheld will be reflected in price before the lockup termination. Firms characterised with higher information asymmetry and with higher ex-ante uncertainty will have strict lockup contracts imposed on them (Goergen, Renneboog and Khurshed (2006)). Brau et al. (2005) build up a model which says that more transparent firms

have a higher probability that the mimickers will be identified, and hence they can afford a shorter lockup. Transparency and information asymmetry is inversely related so very transparent firms will have less strict lockups too. More transparent firms will disclose more information in the IPO prospectuses as well. Thus transparency is directly related to the IPO tone. Hence, I expect that longer lockups and lower lockup expiration returns are related to the IPO tone at the time of IPO.

Ex-ante, it is difficult for investors to differentiate between a good firm and a bad firm (the Lemon's problem) based on the limited information available at the time of IPO. Good quality firms need to do something which is difficult or very costly for a bad firm to replicate. In other words, good firms need to send some signals which are hard for bad firms to replicate. Leland and Pyle (1977) develop up such a model where insiders hold a large holding in their firm and remain undiversified which would send a signal about the quality of the firm. Insiders of good quality firms can hold on to their shares which are difficult to do for the insiders of a bad quality firm. In accordance with this, Courteau (1995) extends the Leland and Pyle (1977) model where lockups are used by insiders in the IPOs lockups to signal the quality of their firms. Additionally, Brau and Fawcett (2006) survey evidence demonstrate that more than 77% of the CFOs concurred or firmly concurred that "insiders commit to a long lockup" was a positive sign. If the IPO tone reflects the quality of the firm, it should be related to the lockup length and lockup expiration returns.

Riskier IPOs that are associated with Conservative or Uncertainty tones are expected to underperform on the long-run. While less resier IPOs are expected to deliver better return on the long-run. Hence, my first hypothesis is:

H1: Lockup expiration return is negatively related to the use of Conservative and Uncertainty tones and will be positively related when Positive tone are used in the prospectuses.

6.5 Long-run IPO performance and tone

In an early study, Brav and Gompers (1997) examine the long-run underperformance of 934 VC-backed IPOs from 1972-1992 and 3,407 non-VC-backed IPOs from 1975-1992. They observe that VC-backed IPOs performed better than non-VC-backed IPOs by utilising equally weighted returns. Krishnan et al. (2011) provide support that VC-backed firms perform better than other firms. Contrasting findings have been reported by Levis (2011) in the context of UK IPOs. Value weighting essentially lessens the difference and considerably diminishes underperformance for non-VC-supported IPOs. In tests utilising a few benchmarks and the Fama-French (1993) three-factor model, VC-backed organisations don't fundamentally fail to meet expectations, while the non-VC-sponsored firms do. Levis (2011) report that private equity backed firms performed better than the VC backed and other firms. The main theme across these papers is that venture capitalists know better than the general investors about the firm as they are often repeated investors. This is the reasons that they may mitigate information asymmetry to some extent.

There has been a methodology debate on the long-run performance of IPOs. Brav and Gompers (1997) conclude that underperformance, notwithstanding, is not an IPO impact. Comparative size and book-to-business sector firms that have not issued equity performs as badly as IPOs. While the studies based on event studies report underperformance, the studies based on Fam-French calendar time portfolios did not

report underperformance. This has been termed as pseudo market timing by Schultz (2003).

Espenlaub, Gregory and Tonks (2000) reconsider the long-term returns of IPOs in the UK utilising information of firms over the period 1985-92, in which they examine IPO underperformance by applying various techniques including a calendar time approach. They find that, while using an event study methodology produces significant underperformance after the 3-year post-IPO period irrespective of the models used (e.g., market adjusted, market model adjusted or raw returns). Hoque and Lasfer (2015) find that for the IPOs, where insiders are net buyers, they underperform whereas where insiders are net seller over-perform. This is true for any benchmark used and irrespective of event studies or calendar time approach. However, Espenlaub et al. (2000) report that underperformance is less severe in 5-year post-IPO period and depends largely on the model used. The underperformance would be less severe if they used the calendar time regressions.

Carter et al. (1998) use the underwriter's reputation to check whether the IPOs that join the market with more prestigious underwriters perform better that who join with the other underwriters. As expected, Carter et al. (1998) find that the underperformance of IPO stocks in respect to the market over a three-year holding period is less extreme for IPOs that join the market with more prestigious underwriters. Consistent with earlier studies (e.g., Carter and Manster, 1990; Megginson and Weiss, 1991), they likewise find that IPOs managed by more prestigious underwriters are connected with less short-run underpricing. Habib and Ljungqvist (2001) and Ritter (1991) report the similar patterns for IPOs with prestigious underwriters. Again, the main theme across these papers is that prestigious underwriters mitigate the

asymmetric information to a certain extent, and hence they perform better than the IPOs underwritten by other underwriters.

Boehmer, Boehmer and Fishe (2006) investigate distributions to institutional and retail investors in 441 IPOs. Notwithstanding the understood positive first-day returns, they demonstrate that institutional investors additionally get more assignments in IPOs with better long-term returns. This is in line with prior studies which find that in good IPOs institutional investors get better allocations than the retail investors (Hanley, 1993, Aggarwal et al., 2002). Boehmer et al. (2006) observe that underlying institutional flips anticipate future returns, proposing that in any event an institutional investor hold profitable private data about IPO firms. These papers assert that institutional investors are better informed than the general investors about the firm as they are often repeated investors. This is the reasons that when they buy shares in an IPO, the IPO is a good quality one and hence it performs better than other issues.

There have been many reasons proposed for the underperformance of IPOs in the long-run. For example, the arguments in Miller (1977) and Morris (1996) which state that the heterogeneous beliefs among the investors and short selling are very costly if not impossible to limit the arbitrage in the small public floats. This heterogeneous belief may be mitigated to some extent with the specific wording and tone measures in the prospectus. Also the use of ambiguous and uncertain words may push the heterogeneous beliefs further. Eckbo and Norli (2001) argue that IPOs underperform in the long-run because they are low risk. I used the risk words by using the Loughrun and McDonald (2011) modified dictionary to examine whether the prospectus tone is related to the long-run IPO performance.

Prospectuses with a Conservative or Uncertainty tones indicate a more risky IPO, hence, the return on the long-run is expected to be lower. On theo ther hand, when the tone of the prospectus is Positive, it would be expected that return over the long-run woud be better. The second hypothesis is:

H2: long-run return is negatively related to the use of Conservative or Uncertainty tones and is positively related when Positive tone are used in the prospectuses.

6.6 Survival of IPOs and tone

Espenlaub et al., (2012) examine IPO survival in the Alternative Investment Market (AIM), where the nominated adviser plays a pivotal role in looking after the IPOs. They find that Nomad reputation significantly affects IPO survival. Initial public offerings managed by reputable Nomads survive longer by around two years than those sponsored by other Nomads. However, they do not find any differences in survival rates of AIM IPOs and to those of North American IPOs. Underwriter reputation is used in the IPO literature in a number of studies to mitigate the information asymmetry of the newly listed firms (Beatty and Ritter, 1986, Carter and Manaster, 1990). While Espenlaub et al., (2012) use the nomad reputation as a measure of information asymmetry and relate it to the survival of IPO firms, I used the tone of the IPO prospectus. Audtrestsch et al. (2005) examine the impact of ownership for young and high-tech firms by utilising a unique dataset of 341 firms traded on the Neuer Markt, the German equivalent of the NASDAQ. Their finding is in contrast to the expectation and also with previous studies on more conventional firms. In particular, they find that possession by CEOs has no impact on firm survival. Their

finding is not consistent with the moral hazard behaviours on the part of the owners and managers.

Ahmad and Jelic (2014) analyse the impact of lockup agreements on the survival of IPOs using the data from London Stock exchange during 1990–2011. Their cox hazard model demonstrates a significant impact of lockup length on the post-IPO survival. Their finding suggests that an IPO survival time increases by 27% if the median lockup length increases by a year. This is consistent with the signalling story commitment hypotheses of Brav and Gompers (2003). Moreover, Ahmad and Jelic (2014) report that the delisting rates for IPOs with shorter lockups are significantly higher than the delisting rates for IPOs with longer lockups regardless of the reasons for delisting. Information asymmetry is reduced to some extent by imposing a longer lockup. Their results highlight the significance of lockup agreements on the resulting survival of recently joined firms and bring confidence among the investors to invest in the newly listed firms. Also, this has an implication in terms of the IPO market development in terms of the listing and regulation standards of the UK IPO market.

Chadha (2003) analyse a small number of US IPOs that delist because of performance related reasons within 3 years after IPO. Her paper reveals insight into the intentions of the managers and owners taking their firm public. Managers are better informed about the firms and their trades reflect information in the market (Seyhun, 1986). Chadha (2003) look at the insider trade from the beginning till the end of IPOs to examine the hypotheses that Insiders offer overvalued shares intentionally. Lee (1997) examine insider trading around seasoned equity offering to examine whether they issue overvalued shares. Chadha (2003) examines the conduct and the trading of insiders at the offer stage like Hoque and Lasfer (2009), after the lockup expiration

(Field and Hanka, 2001), and preceding delisting (Seyhun and Bradely, 1997). Chadha (2003) did not find any evidence that proposes that insiders purposely issue overvalued shares. Insiders trading behaviour at the offer and after the lockup termination does not offer any confirmation of insiders methodically abusing their private information to pick up to the detriment of outside investors like Hoque and Lasfer (2015). Also, her findings demonstrate that insiders of delisted IPOs do not sell shares before the firm delists. This is consistent with the notion that managers are over optimistic and overconfident. On the other hand, it is likewise predictable with the conduct of managers being unique in relation to that of enlisted directors. Generally speaking, the outcomes provide support for either the "hubris" theory or the "entrepreneurial pride" clarification (Heaton, 2002); they are to a great extent conflicting with the "windows of opportunity" explanation (Loughrun and Ritter, 1995, 2002).

Hensler et al., (1997) examine the effects of several firm characteristics and IPO characteristics on the survival of IPO firms by using an accelerated failure time (AFT) model. The results show that large IPOs, old firms, firms which underprice more and firms with increased level of IPO activity in the aftermarket, firms with higher level of insider ownership survives longer. Large firms and old firms are firms with less information asymmetry (Ritter, 1984). They also find that survival time decreases with the increase of general market level at the time of offering and number of risk characteristics. There are differences in industry origin in terms of survival rates of the IPOs. The IPO survives shorter if the firm is in from certain industries such as computer and data, restaurant, wholesale industry, and airline and survives longer if the IPO is from optical or drug industries. Possibly, the degree of information asymmetry is different in different industries. High tech industry is characterised with high information asymmetry. Demers and Joos (2007) develop an IPO failure risk

model by incorporating accounting and other deal specific characteristics. They document statistically different failure rates between tech and non-tech IPOs, which is in line with Hensler et al., (1997) finding of industry differences in survival. While they use several accounting variables at the time of IPOs to distinguish between delisted and non-delisted IPO firms I use the IPO tone measure from prospectus to do that. Hensler et al., (1997) measure the variable risk from Jay Ritter's website which is mentioned as a risk. The risk variable is the count of risk factors reported in the prospectus. If there is no specific risk factor section, the variable risk is zero. However, the measure of the risk factor is problematic as it just counts the number of risk factors. I used the Loughrun and McDonald (2011) dictionary measures in the prospectus to relate it to the survival firms of the IPOs.

When the tone of thep prospectus is either Conservative or Uncertain, survival rate is expected to be lower. On the other hand, if the tone of the prospectus is Positive, that would indicated a less riskier IPO and hence higher survival rate. Thus I hypothesise that:

H3: Survival rate is negatively related to Conservative or Uncertainty tones and is positively related to Positive tones in the prospectuses.

6.7 Data and methodology

I use the same IPO prospectuses that I used in first two chapters. As per Table 5.3, The median of the number of words per prospectus in the main market is 66,662 words while for the AIM market it is 27,792 words. I have only included the IPOs that have more than 2000 words that is less than 10 percent of the median of the AIM market prospectuses and hence would not have a comparable data. I have a total number of 946 prospectuses in our sample. 26 of the prospectuses in our sample where

photocopies that required optical character reader (OCR) to convert them to text. I used the tool provided by Google Drive to convert them photocopies to text.

I used the dictionaries developed by (Loughran and McDonald, 2011) in addition to Harvard IV dictionary for negative words. I analysed each prospectus using AntWordProfiler software. More details about the dictionaries, calculating the prospectuses weights and the software can be found in section 5.5.

Lockup length and lockup expiration date have been collected manually from the prospectuses. The share prices, indices data and bankruptcy dates were collected from DataStream. I got Fama-French Daily Factors from the website of the Business School of the University of Exeter.

To test my hypothesis, I have used a number of methodologies. First, to study the return around the lockup expiry date, I used 2 methods to calculate the abnormal returns. I used both the market model and Fama-French 4-Factor.

For the market model, I run an OLS regression using the daily return for each company as the dependent variable and index return as the independent variable as follows:

$$R_s = \alpha + \beta R_i + \varepsilon_i \quad (6.1)$$

Where:

R_s : Daily return on share price

R_i : Daily return on index

α : Intercept

β : Slope

I ran the regression for the period starting from the date of the IPO until the starting date of the event window. After getting the coefficients of the intercept and the slope of each regression, I calculated the expected return during the event window. Then I calculated the daily abnormal return as the excess return over the expected return as follows:

$$AR = R_s - E(R) \tag{6.2}$$

Where:

AR : Abnormal Return

$E(R)$: Expected return using the regression coefficients

Then I calculated the cumulative abnormal return (CAR) for the period as the numerical sum of the daily return during the event study window.

I also examined the same using Fama-French 4-Factor Model.

$$R_s - R_f = \alpha + \beta_1(R_m - R_f) + \beta_2 SMB + \beta_3 HML + \beta_4 UMD + \varepsilon_i \tag{6.3}$$

Where:

R_s : Daily return on share price

R_f : Risk free rate of return

α : Intercept

β : Coefficient Multiplier

SMB : Small minus Big – The difference in return between small cap companies to large cap companies.

HML: High minus Low – The difference in return between growth companies to value companies.

UMD: Up minus Down – The difference in return between winner companies and loser companies.

First, I run an OLS regression for each company using the four factors to find the intercept and the coefficients multipliers. Then I do the same as what I did for the market model by calculating the expected return for every day during the event study window. Then I calculate the abnormal return as the difference between the actual return and the expected return.

For studying the bankruptcy, I used a number of methods. First, I ran an OLS regression.

$$\begin{aligned} Log(TimeTillDead) &= \alpha + \beta_1 Spread_i + \beta_2 Log\ Proceeds_i \\ &+ \beta_3 Market_i + \beta_4 MultiBookrunner_i \\ &+ \beta_5 Prestigious_i + \beta_6 D1\ Underpricing_i \\ &+ \beta_7 LogLockupPeriod_i \left(\sum_{j=1999}^{2012} \beta_j Year_j \right) + \varepsilon_i \end{aligned} \tag{6.4}$$

Then I ran a logit regression using Dead Dummy. Finally, I used Cox Proportional Hazard Model using Dead Dummy and Days Till Dead in days as the period.

For long period return, I have calculated the Buy-And-Hold Abnormal Return (BHAR). First, I calculated the monthly return on both the share prices and the index. Then I calculated the difference between the share return and index return. Finally, I calculated the compounded rate of return over the holding period.

6.8 Results:

6.9 Descriptive Statistics

Table 6.1, panel A shows the word count statistics. I can notice that After Harvard Negative list, Modal dictionary has highest frequency use per prospectus with a mean of 524 words. Uncertainty and Litigious come after with a mean of 411 and 416 words. On the other hand, the Superfluous dictionary is the most rarely used dictionary among the others by far with a mean of 10 words per prospectus. Interesting dictionary comes after with 82 words. This will give Superfluous and Interesting more weight when used. Their effect will be even more when using the tf.idf weight.

Panel B shows the proportional weight of each dictionary. Again, with Harvard Negative List aside, Modal dictionary shows the highest mean. Litigious and Uncertainty dictionaries come after with Litigious showing a slightly more weight that Uncertainty. With respect to the dictionaries with the least proportional weight, Superfluous and Interesting are again the least ones.

Panel C shows the Time-Frequency Inverse-Time-Frequency weight statistics. After Harvard Negative list, the mean of Negative dictionary comes way ahead of the other dictionaries with 15.24. Next to it comes Litigious, Positive and Uncertainty with

4.80, 4.65 and 3.55 in the same order. From the other end, Superfluous and Interesting are keeping their position as the lowest weight.

Table 6.1: Dictionaries Statistical Data

Variable	Mean	Std. Dev.	Min	Max
Negative	375.826	346.519	3	2823
Harvard_iv	1356.205	1147.381	45	10052
Constraining	318.081	274.685	8	1979
Litigious	411.262	298.681	16	2304
Superfluous	10.124	10.576	0	184
Modal	524.104	415.045	18	2554
Uncertainty	416.911	363.423	12	2619
Positive	207.449	171.990	9	1018
Interesting	82.171	99.425	0	1083
Irr_verb	324.720	247.997	12	2269
Panel A: Words count in prospectuses				
PrNegative	0.009	0.002	0.001	0.019
PrHarvard_iv	0.034	0.005	0.014	0.060
PrConstraining	0.008	0.003	0.002	0.040
PrLitigious	0.011	0.003	0.003	0.029
PrSuperfluous	2.7E-4	1.7E-4	0.000	0.001
PrModal	0.014	0.003	0.004	0.023
PrUncertainty	0.010	0.002	0.003	0.018
PrPositive	0.005	0.002	0.001	0.016
PrInteresting	0.002	0.002	0.000	0.013
PrIrr_Verb	0.009	0.001	0.003	0.018
Panel B: Words proportional weight for each dictionary				
TFNegative	15.244	13.329	0.063	102
TFHarvard_iv	20.865	17.231	0.542	129
TFConstraining	2.551	1.993	0.003	12
TFLitigious	4.803	3.623	0.042	25.1
TFSuperfluous	0.275	0.347	0.000	2.75
TFModal	1.005	0.803	0.016	4.61
TFUncertainty	3.555	3.241	0.020	20.9
TFPositive	4.657	2.916	0.067	17.2
TFInteresting	0.797	0.743	0.000	7.6
TFIrr_Verb	1.848	1.203	0.040	7.68
Panel C: Words tf.idf weight for each dictionary				
Observations	946			

This table has 3 sections. Panel A shows statistics of the number of words from each dictionary in the prospectuses in our sample. Panel B shows statistics of the proportional weight of each dictionary. Panel C shows statistics of the Time Frequency – Inverse Document Frequency weight for each dictionary.

Table 6.2 shows the lockup period and the return on CAR at the lockup expiry date. First, I can notice that for the companies that imposed a lockup period from our sample, their lockup periods vary from 85 days to about three and a half years. With a

mean of 391 days. This is driven by the high frequency of companies using 365 days as the lockup period. The second and third parts of the table show the CAR over 3 different periods using Market Model and Fama-French 4-Factor model. I can notice that the results are close for both models.

Table 6.2: Lockup Period and Cumulative Abnormal Return at Lockup period expiration date

Variable	Obs	Mean	Std. Dev.	Min	Max
Lockup period in days	945	388.268	135.556	85	1260
Market Model CAR0_2	776	0.003	0.080	-0.605	0.690
Market Model CAR0_5	776	0.005	0.103	-0.640	0.688
Market Model CAR0_10	775	0.005	0.139	-1.496	0.805
Fama-French CAR0_2	776	0.003	0.081	-0.609	0.689
Fama-French CAR0_5	776	0.006	0.105	-0.629	0.687
Fama-French CAR0_10	775	0.006	0.145	-1.569	0.793

This table shows statistics of Cumulative Abnormal Return (CAR) over 2, 5 and 10 days from the lockup period expiration date. There are three sections in this table. The first table shows statistics of the lockup period for the companies in our sample. The second section shows CAR calculated using Market Model. The third section shows CAR calculated using Fama-French 4-Factor model.

Table 6.3 shows the Buy and Hold Abnormal Return (BHAR) over 4 holding periods. I can notice that the BHAR over 1 year shows an increase in return and standard deviation. It shows the highest mean of returns. For BHAR for 2 and 3 years holding periods, the mean of return shows a drop from 23.6% to 7.7% and 3.5% in the same order. However, the standard deviation is still high with 1.8 and 2.5.

Table 6.3: Buy and Hold Abnormal Return (BHAR) for 6-Month, 1-Year, 2-Year and 3-Year

Variable	Obs	Mean	Median	Std. Dev.	Min	Max
BHAR 6 Months	925	0.022	-0.0411	0.508	-0.908	5.913
BHAR 1 Year	924	0.030	-0.1035	0.821	-1.034	9.001
BHAR 2 Years	883	0.008	-0.254	1.155	-0.993	20.983
BHAR 3 Years	835	0.025	-0.3176	1.529	-0.995	29.343
Panel A: Abnormal Return using Index						
BHAR 6 Months	833	0.057	-0.0152	0.589	-0.912	8.569
BHAR 1 Year	833	0.087	-0.0561	0.984	-0.983	12.112
BHAR 2 Years	833	0.095	-0.2329	1.355	-0.982	22.623
BHAR 3 Years	833	0.161	-0.3306	1.844	-0.997	29.807
Panel B: Abnormal Return Using Expected Return						

This table shows statistics of the Buy and Hold Abnormal Return (BHAR) over 4 different periods (6 months, 1 year, 2 years and 3 years). BHAR is calculated starting from the IPO date. In Panel A, BHAR is calculated as the difference between the share price return and the index return. In Panel B, BHAR is calculated using the difference between share price return and the expected share price return.

6.10 IPO lockup expiration return analysis

From Table 6.4, Panel A, I can notice that none of the correlations between the reported abnormal returns with the proportional dictionaries weights is significant. This applies to both the Market Model variables and the Fama-French Model. However, Panel B shows a significant correlation between Time Frequency – Inverse Time-Frequency Superfluous dictionary weight and the Cumulative Abnormal Return for the periods (0 – 5) and (0 – 10) for both models. It shows to be more significant for the period (0 – 10). The negative sign indicates that showing a tone of presenting extra data in the prospectuses is correlated with less cumulative abnormal return following the lockup expiration date.

Table 6.4: Correlation between (market model and Fama-French variables around the expiry date with different dictionaries' weights and variables)

	MAR0	MCAR0-2	MCAR0-5	MCAR0-10	FFAR0	FFCAR0-2	FFCAR0-5	FFCAR0-10
PrNegative	0.01	-0.036	-0.048	0.009	0.029	-0.026	-0.028	0.021
PrHarvard_iv	-0.003	-0.042	-0.047	-0.003	-0.027	-0.043	-0.047	-0.003
PrConstraining	0.039	0.009	0.003	0.025	-0.012	0.002	-0.004	0.018
PrLitigious	0.015	-0.039	-0.007	0.042	0.056	-0.038	-0.005	0.038
PrSuperfluous	0.004	-0.022	-0.02	-0.03	0.01	-0.019	-0.025	-0.028
PrModal	-0.018	-0.049	-0.031	0.008	0.007	-0.052	-0.024	0.018
PrUncertainty	-0.015	-0.031	-0.014	0.028	0.022	-0.032	-0.002	0.038
PrPositive	-0.023	-0.032	-0.049	-0.032	0.001	-0.034	-0.059	-0.043
PrInteresting	0	0.01	0.001	0.001	-0.006	0.014	0.004	0.001
PrIrr_Verb	-0.001	-0.034	-0.007	0.021	-0.046	-0.029	0.005	0.022

Panel A: Correlation with Proportional Weight Dictionaries

TFNegative	-0.018	-0.029	-0.051	-0.024	-0.013	-0.025	-0.044	-0.016
TFHarvard_iv	-0.012	-0.023	-0.043	-0.02	-0.024	-0.02	-0.037	-0.012
TFConstraining	-0.006	-0.005	-0.015	-0.009	0.01	0.001	-0.007	0.001
TFLitigious	0.003	-0.02	-0.019	-0.016	0.012	-0.012	-0.01	-0.008
TFSuperfluous	-0.038	-0.037	-0.061*	-0.106***	-0.031	-0.042	-0.064*	-0.107***
TFModal	-0.027	-0.029	-0.051	-0.033	-0.037	-0.028	-0.039	-0.022
TFUncertainty	-0.028	-0.024	-0.036	-0.022	-0.026	-0.024	-0.027	-0.013
TFPositive	-0.008	-0.042	-0.059*	-0.045	-0.037	-0.042	-0.058	-0.043
TFInteresting	-0.007	-0.005	-0.018	-0.025	-0.02	-0.004	-0.012	-0.019
TFIrr_Verb	-0.018	-0.047	-0.045	-0.028	-0.063*	-0.046	-0.037	-0.022

Panel B: Correlation with Time Frequency – Inverse Document Time-Frequency Weight Dictionaries

log(proceeds)	-0.049	-0.051	-0.051	-0.061*	-0.047	-0.051	-0.051	-0.059
Spread	0.064*	0.062*	0.053	0.069*	0.031	0.068*	0.066*	0.076**
Prestigious	0.004	0.062*	0.08**	0.051	-0.014	0.072**	0.085**	0.058
STDev	0.016	0.016	0.01	-0.014	0.023	0.033	0.035	0.005
Age in days	0.02	0.014	-0.01	-0.008	-0.031	0.006	-0.019	-0.017
Potential Growth	-0.006	0.001	-0.009	-0.016	-0.001	0	-0.011	-0.015
Log(Lockup)	0.22***	-0.012	-0.02	-0.023	-0.011	-0.014	-0.025	-0.029
Underpricing	-0.018	-0.033	0.012	0.023	0.074**	-0.033	0.009	0.025
Sales	-0.001	0.007	0.006	-0.012	-0.001	0.004	-0.002	-0.018
Overhang	0.047	0.044	-0.035	-0.007	-0.001	0.049	-0.032	-0.005
Net Income	-0.003	-0.123***	-0.104**	-0.042	-0.144***	-0.124***	-0.095**	-0.029
Post-IPO EPS	-0.025	-0.042	0.011	-0.014	-0.023	-0.038	0.007	-0.026

Panel C: Correlation with Other Variables

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation of the abnormal return. MAR0 is the Abnormal Return calculated using the Market Model calculated at the day of the lockup period expiration date. MCAR is the Cumulative Abnormal Return calculated using Market Model for periods of (2, 5 and 10 days). FFAR0 is the Abnormal Return calculated using Fama-French 4-Factor model. FFAR is the Cumulative Abnormal Return calculated using Fama-French 4-Factor Model. This table has three sections. In Panel A, it shows the correlation with the dictionaries proportional weights. In Panel B, it shows the correlations with the time-frequency – inverse document frequency dictionaries' weights. In Panel C it shows the correlations with other variables. Proceeds is the money raised in the IPO. The spread is the amount charged by the underwriter as a percentage of the IPO proceeds. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise. STDev is the standard deviation of share price during the first year of trading. Age in Days is the age of the company on the day of the IPO in days. Potential growth is a factor showing if the company increased the size of its capital at the years following the IPO. Lockup is the lockup period in days. the underpricing factor is calculated as the return on trading. If IPO was underpriced, the underpricing factor is positive, if it is overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0. Sales is the reported sales the year prior to the IPO. The Overhang is the percentage of shares that are not traded. Net Income is the reported net income the year prior to the IPO. Post IPO EPS is the reported Earnings Per Share the year prior to the IPO.

On the other hand, from Panel C, I can notice a positive correlation with the underwriters' prestigious dummy variable. This indicates that the IPOs undertaken by prestigious underwriters shows a slightly higher abnormal return after the lockup expiration date. In addition, for the IPOs where the underwriters charged higher spread fees, show slightly higher abnormal returns for both models.

The length of the lockup period is showing a positive correlation in terms of Log(Lockup) and the Market Model Abnormal Average Return on the expiration date. This indicates that the longer the lockup period, the higher the abnormal average return on the lockup period expiration date. As the lockup period is getting higher, the shareholders will take more action changing their positions and gaining more return compared to the shorter lockup periods.

The Underpricing variable is showing a significant positive correlation with Fama-French Abnormal Average Return on the expiration return. This indicates that the IPOs that are more underpriced tend to be slightly generating a more abnormal return on the lockup expiration date.

Finally, from the table, I can notice that the companies that have reported higher net income for the year prior to the IPO have reported less abnormal returns. This indicates that the companies that have been reporting positive net income before going public are more stable and hence show less abnormal returns during the event of the lockup expiration.

I ran regressions for all the variables from both models with different dictionaries. i have noticed that the tone of the prospectus does not show a significant relationship with the event study variables around the lockup period expiration date. This indicates that the effect of the prospectus tone decays over the period. Table 6.5

shows the regressions model coefficients of the Market Model CAR (0 – 5). I can notice that prestigious dummy variable is showing significance with a positive coefficient. This goes in line with the results from the correlation in Table 6.4. IPOs undertaken by prestigious underwriters show higher abnormal return during the lockup expiration date.

Table 6.5: Regression of Market Model CAR 0-5 with Proportional Weight Dictionaries

MM CAR 0-5		PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
PrNegative		-2.418									
		(-1.33)									
PrHarvard_iv			-0.815								
			(-1.1)								
PrConstraining				0.399							
				(0.33)							
PrLitigious					-0.575						
					(-0.41)						
PrSuperfluous						-21.314					
						(-0.97)					
PrModal							-0.93				
							(-0.71)				
PrUncertainty								-0.087			
								(-0.05)			
PrPositive									-3.089		
									(-1.33)		
PrInteresting										-0.546	
										(-0.22)	
PrIrr_Verb											-1.717
											(-0.56)
Spread	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(0.76)	(0.87)	(0.74)	(0.77)	(0.74)	(0.74)	(0.75)	(0.76)	(0.72)	(0.76)	(0.75)
Log(proceeds)	-0.012	-0.011	-0.011	-0.012	-0.012	-0.013	-0.012	-0.012	-0.012	-0.012	-0.013
	(-1.51)	(-1.36)	(-1.39)	(-1.54)	(-1.54)	(-1.59)	(-1.5)	(-1.51)	(-1.48)	(-1.51)	(-1.58)
Market Dummy	-0.002	-0.003	-0.001	-0.003	-0.002	-0.003	-0.002	-0.002	-0.002	-0.002	-0.002
	(-0.16)	(-0.22)	(-0.09)	(-0.2)	(-0.12)	(-0.18)	(-0.12)	(-0.16)	(-0.15)	(-0.16)	(-0.16)
Multi-bookrunner	0.022	0.021	0.022	0.021	0.022	0.021	0.022	0.022	0.02	0.021	0.021
	(1.1)	(1.07)	(1.1)	(1.07)	(1.11)	(1.06)	(1.14)	(1.1)	(1)	(1.09)	(1.08)
Prestigious	0.02**	0.021**	0.02**	0.02**	0.02**	0.02**	0.02**	0.02**	0.02**	0.02**	0.02**
	(2.45)	(2.53)	(2.51)	(2.44)	(2.45)	(2.48)	(2.45)	(2.45)	(2.49)	(2.46)	(2.45)
Underpricing	-1.5E-5	-1.1E-5	-1.8E-5	-1.6E-5	-9.3E-6	-1.5E-5	-8.3E--6	-1.5E-5	-2.3E-5	-1.7E-5	-1.3E-5
	(-0.2)	(-0.14)	(-0.23)	(-0.2)	(-0.12)	(-0.2)	(-0.1)	(-0.19)	(-0.3)	(-0.21)	(-0.16)
Log(Lockup)	-0.027	-0.031	-0.026	-0.028	-0.027	-0.028	-0.027	-0.027	-0.026	-0.027	-0.027
	(-0.72)	(-0.84)	(-0.7)	(-0.74)	(-0.71)	(-0.74)	(-0.71)	(-0.72)	(-0.71)	(-0.72)	(-0.73)
Constant	0.089	0.118	0.113	0.09	0.094	0.098	0.099	0.09	0.106	0.091	0.106
	(0.87)	(1.13)	(1.08)	(0.88)	(0.91)	(0.95)	(0.96)	(0.86)	(1.03)	(0.89)	(0.99)
Observations	776	776	776	776	776	776	776	776	776	776	776
Prob > F	0.4578	0.4132	0.4461	0.5151	0.5112	0.4624	0.4898	0.5218	0.4124	0.5189	0.5023
Adj R-squared	0	0.001	0.0003	-0.0011	-0.0011	0	-0.0006	-0.0013	0.0011	-0.0012	-0.0009

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Cumulative Abnormal Return for 5 days calculated using Market Model where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 6.6: Regression of Fama-French Model CAR 0-5 with Proportional Wight Dictionaries

FF CAR 0-5		PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
PrNegative		-1.363									
		(-0.73)									
PrHarvard_iv			-0.815								
			(-1.08)								
PrConstraining				0.233							
				(0.19)							
PrLitigious					-0.44						
					(-0.31)						
PrSuperfluous						-26.411					
						(-1.18)					
PrModal							-0.509				
							(-0.38)				
PrUncertainty								0.854			
								(0.44)			
PrPositive									-3.972*		
									(-1.68)		
PrInteresting										-0.338	
										(-0.13)	
PrIrr_Verb											-0.692
											(-0.22)
Spread	0.002	0.002	0.001	0.002	0.001	0.001	0.001	0.002	0.001	0.002	0.001
	(1.22)	(1.28)	(1.21)	(1.23)	(1.21)	(1.2)	(1.22)	(1.24)	(1.17)	(1.23)	(1.22)
Log(proceeds)	-0.011	-0.01	-0.01	-0.011	-0.011	-0.012	-0.011	-0.011	-0.011	-0.011	-0.011
	(-1.35)	(-1.27)	(-1.24)	(-1.37)	(-1.37)	(-1.46)	(-1.35)	(-1.37)	(-1.32)	(-1.36)	(-1.37)
Market Dummy	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002	-0.001	-0.001	-0.001	-0.002	-0.002
	(-0.1)	(-0.14)	(-0.04)	(-0.13)	(-0.07)	(-0.13)	(-0.09)	(-0.09)	(-0.1)	(-0.11)	(-0.1)
Multi-bookrunners	0.016	0.016	0.016	0.016	0.016	0.015	0.016	0.016	0.013	0.016	0.016
	(0.81)	(0.78)	(0.81)	(0.78)	(0.81)	(0.75)	(0.82)	(0.79)	(0.67)	(0.8)	(0.8)
Prestigious	0.021**	0.022***	0.022***	0.021**	0.021**	0.022***	0.021**	0.021**	0.022***	0.021**	0.021**
	(2.57)	(2.61)	(2.63)	(2.56)	(2.57)	(2.61)	(2.57)	(2.56)	(2.62)	(2.57)	(2.57)
Underpricing	-2.7E-5	-2.4E-5	-2.9E-5	-2.7E-5	-2.2E-5	-2.7E-5	-2.3E-5	-3.1E-5	-3.7E-5	-2.8E-5	-2.6E-5
	(-0.34)	(-0.3)	(-0.37)	(-0.34)	(-0.27)	(-0.33)	(-0.28)	(-0.38)	(-0.46)	(-0.34)	(-0.32)
Log(Lockup)	-0.035	-0.038	-0.035	-0.036	-0.035	-0.036	-0.035	-0.035	-0.035	-0.035	-0.035
	(-0.93)	(-0.99)	(-0.91)	(-0.94)	(-0.92)	(-0.96)	(-0.92)	(-0.91)	(-0.92)	(-0.93)	(-0.93)
Constant	0.125	0.141	0.148	0.125	0.128	0.136	0.13	0.116	0.146	0.126	0.131
	(1.19)	(1.32)	(1.39)	(1.19)	(1.22)	(1.29)	(1.23)	(1.08)	(1.39)	(1.2)	(1.21)
Observations	776	776	776	776	776	776	776	776	776	776	776
Prob > F	0.1213	0.139	0.1221	0.1539	0.152	0.1162	0.1505	0.1491	0.0852	0.1544	0.1535
Adj R-squared	0.0098	0.0092	0.01	0.0086	0.0086	0.0103	0.0087	0.0088	0.0122	0.0085	0.0086

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Cumulative Abnormal Return for 5 days calculated using Fama-French 4-Factor Model where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

The same applies to Table 6.6. With Fama-French Model, the regression showed significance with a negative coefficient of the Proportional weight of the positive dictionary. The more positive tone of the prospectus, the less abnormal return during the lockup expiration date period. From the above tables, I cannot accept our first hypothesis (H1). Positive, Conservative and Uncertainty do not show a significant relationship with lockup expiration return. Only Superfluous tone shows a negative coefficient and hence less lockup expiration return as the prospectus uses superfluous tone.

6.11 Long-run IPO performance analysis

From Table 6.7 I can see that none of the regressions of the Buy and Hold Abnormal Return (BHAR) in panel A with the proportional weight dictionaries shows any significance. However, in panel B, for 6 months holding period, the Interesting dictionary is showing a level of significance with a negative coefficient. This shows that the use of this dictionary in the prospectuses will reduce the return over the holding period. As I calculate the return from the IPO dates, this indicates that more interesting words will result in less underpricing and hence less return. The Superfluous dictionary is showing significance with 1 year holding period with a positive coefficient. This shows that prospectuses that use more superfluous words result in higher abnormal holding return over 1 year period.

Table 6.7: Correlation of different Buy and Hold Abnormal Return BHAR with 4 different holding periods with different dictionaries' weight and variables

	BHAR 6 Months	BHAR 1 Year	BHAR 2 Years	BHAR 3 Years
PrNegative	0.002	0.01	-0.008	-0.018
PrHarvard_iv	0.002	0.01	0.009	0.022
PrConstraining	0.018	-0.001	0.002	0.001
PrLitigious	-0.004	-0.006	-0.03	-0.05
PrSuperfluous	-0.004	0.012	-0.016	-0.03
PrModal	0.046	0.028	0.018	0.01
PrUncertainty	0.015	0.022	0.015	0.014
PrPositive	0.023	0.008	0.023	-0.012
PrInteresting	0.005	-0.018	0.013	0.012
PrIrr_Verb	-0.018	0.011	0.005	-0.001
Panel A: Correlation with Proportional Weight Dictionaries				
TFNegative	-0.032	0.029	0.016	0.01
TFHarvard_iv	-0.037	0.026	0.008	0.006
TFConstraining	-0.026	0.028	0.019	0.013
TFLitigious	-0.036	0.031	0.023	0.002
TFSuperfluous	0.014	0.059*	0.049	0.031
TFModal	-0.045	0.012	-0.003	0.008
TFUncertainty	-0.031	0.024	0.013	0.024
TFPositive	-0.04	0.004	0.001	-0.003
TFInteresting	-0.071**	-0.03	-0.039	-0.032
TFIrr_Verb	-0.039	0.017	0.014	0.034
Panel B: Correlation with Time Frequency – Inverse Document Frequency Weight Dictionaries				
Log(proceeds)	0.039	0.022	0.037	0.021
Spread	-0.145***	-0.095***	-0.096***	-0.075**
Prestigious	-0.028	-0.045	-0.001	0.025
STDev	-0.017	0.004	-0.01	-0.001
Age in days	0.03	0.044	0.052	0.042
Potential Growth	0.009	0.031	0.057	0.052
Log(Lockup)	0.043	0.038	0.012	-0.014
Underpricing	0.022	-0.048	-0.075**	-0.045
Sales	-0.005	0.013	0.070*	0.109***
Overhang	-0.07**	-0.062*	-0.069**	-0.059*
Net Income	0	-0.005	-0.003	-0.009
Post-IPO EPS	0.016	0.047	0.036	0.071*
Panel C: Correlation with Different Variables				

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation of Buy and Hold Abnormal Return (BHAR) over 4 different periods (6 months, 1 year, 2 years and 3 years). BHAR is calculated starting from the IPO date. This table has three sections. In Panel A, it shows the correlation with the dictionaries proportional weights. In Panel B, it shows the correlations with the time-frequency – inverse document frequency dictionaries' weights. In Panel C it shows the correlations with other variables. Proceeds is the money raised in the IPO. The spread is the amount charged by the underwriter as a percentage of the IPO proceeds. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise. STDev is the standard deviation of share price during the first year of trading. Age in Days is the age of the company on the day of the IPO in days. Potential growth is a factor showing if the company increased the size of its capital at the years following the IPO. Lockup is the lockup period in days. the underpricing factor is calculated as the return on trading. If IPO was underpriced, the underpricing factor is positive, if it is overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0. Sales is the reported sales the year prior to the IPO. The Overhang is the percentage of shares that are not traded. Net Income is the reported net income the year prior to the IPO. Post IPO EPS is the reported Earnings Per Share the year prior to the IPO.

From Panel C, I can notice that Spread is showing significance across all holding periods. The coefficient is negative which means that prospectuses with underwriters charging higher fees will result in less abnormal return. As I was calculating the return from the IPO issuing price, this indicates that they are more accurately priced and hence less underpriced. The absolute coefficient value is decreasing as the holding period increases. Though it is expected for an underpriced share to generate a higher return, yet, Underpricing is showing negative coefficients and only significant of 2 years holding period. The same applies to Overhang. The higher the overhang shares, the less the abnormal holding period return.

Table 6.8 shows univariate regressions of BHAR with different holding periods and different dictionaries' weights. It shows significance with a negative correlation between 6 months holding period and Interesting tf.idf dictionary. It means that more interesting words used in the prospectus will result in less return over a holding period of 6 months. Over the 1-year holding period, the superfluous dictionary is showing some significance with a positive coefficient. More superfluous words result in higher abnormal return over a holding period of 1 year. Results shown in Table 6.7 and Table 6.8 do not show high significance for most of the tested variables. However, 2 variables showing significance (for 6-month and 1-year holding periods), yet, they do not give support to our second hypothesis (H2). For 6-month I notice that a positive tone represented by the Interesting dictionary has a negative relationship with the return. For the 1-year holding period, superfluous tone shows a positive relationship with the return.

Table 6.8: Univariate regressions of Buy and Hold Abnormal Returns (BHAR) with different holding periods with different dictionaries' weights

	BHAR 6 Months		BHAR 1 Year		BHAR 2 Years		BHAR 3 Years	
	Proportional	tf.idf	Proportional	tf.idf	Proportional	tf.idf	Proportional	tf.idf
Negative	0.417	-0.001	3.501	0.002	-4.091	0.001	-12.295	0.001
	(0.06)	(-0.97)	(0.29)	(0.89)	(-0.24)	(0.46)	(-0.52)	(0.28)
Harvard_iv	0.209	-0.001	1.583	0.001	1.918	0.001	6.334	0.001
	(0.07)	(-1.13)	(0.31)	(0.78)	(0.26)	(0.24)	(0.64)	(0.17)
Constraining	2.946	-0.007	-0.377	0.012	0.645	0.011	0.575	0.01
	(0.54)	(-0.8)	(-0.04)	(0.86)	(0.05)	(0.55)	(0.03)	(0.36)
Litigious	-0.642	-0.005	-1.701	0.007	-12.143	0.007	-26.867	0.001
	(-0.11)	(-1.11)	(-0.18)	(0.93)	(-0.9)	(0.68)	(-1.45)	(0.05)
Superfluous	-11.542	0.02	55.124	0.14*	-107.169	0.162	-263.121	0.138
	(-0.12)	(0.42)	(0.35)	(1.8)	(-0.48)	(1.44)	(-0.86)	(0.89)
Modal	7.996	-0.029	7.818	0.013	6.972	-0.004	5.309	0.016
	(1.41)	(-1.38)	(0.85)	(0.37)	(0.53)	(-0.07)	(0.29)	(0.24)
Uncertainty	3.444	-0.005	8.288	0.006	7.727	0.005	10.179	0.012
	(0.45)	(-0.95)	(0.67)	(0.74)	(0.43)	(0.38)	(0.41)	(0.7)
Positive	7.152	-0.007	4.22	0.001	16.448	0	-11.318	-0.002
	(0.69)	(-1.21)	(0.25)	(0.13)	(0.69)	(0.02)	(-0.35)	(-0.1)
Interesting	1.565	-0.048**	-9.306	-0.033	9.558	-0.06	11.939	-0.067
	(0.15)	(-2.15)	(-0.54)	(-0.92)	(0.39)	(-1.16)	(0.36)	(-0.93)
Irr_verb	-6.899	-0.016	6.717	0.012	4.247	0.014	-1.102	0.046
	(-0.54)	(-1.17)	(0.33)	(0.51)	(0.14)	(0.42)	(-0.03)	(0.98)

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows a summary of the univariate regressions of the Buy and Hold Abnormal Return (BHAR) over 4 different periods (6 months, 1 year, 2 years and 3 years) and both the proportional weight and time-frequency – inverse document frequency of all the dictionaries.

Table 6.9 Shows regressions of BHAR with different holding periods returns with different variables. I can notice that spread is showing high significance. The negative coefficients indicate that IPOs with higher spread yield lower returns. This is the case over all the holding periods reported in the table. Underpricing is also showing some significance, however, with a very small coefficient value.

Table 6.9: Regression of Buy and Hold Abnormal Return (BHAR) with different holding periods

	6 Month Holding Period	1 Year Holding Period	2 Year Holding Period	3 Year Holding Period
Spread	-0.02***	-0.027***	-0.037***	-0.052***
	(-3.72)	(-3.02)	(-2.76)	(-2.75)
Log(proceeds)	0	-0.038	-0.069	-0.147
	(-0.01)	(-0.67)	(-0.83)	(-1.28)
Market Dummy	0.003	-0.007	-0.107	-0.184
	(0.05)	(-0.07)	(-0.74)	(-0.94)
Multi-bookrunner	0.031	0.164	0.185	0.04
	(0.42)	(1.39)	(1.09)	(0.17)
Prestigious	-0.047	-0.074	-0.041	0.076
	(-1.25)	(-1.21)	(-0.47)	(0.64)
Underpricing	0	-0.001	-0.001**	-0.001
	(0.55)	(-1.34)	(-2.15)	(-1.39)
Log(Lockup)	0.192	0.27	0.183	-0.087
	(1.47)	(1.27)	(0.6)	(-0.21)
Constant	-0.155	-0.096	0.325	0.841
	(-0.43)	(-0.17)	(0.39)	(0.75)
Observations	924	923	883	835
Prob > F	0.0289	0.0714	0.1332	0.7001
Adj R-squared	0.0148	0.0108	0.0081	-0.0045

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows 4 OLS Regressions of the Buy and Hold Abnormal Return (BHAR) over 4 different periods (6 months, 1 year, 2 years and 3 years). The control Variables are the spread is the amount charged by the underwriter as a percentage of the IPO proceeds, the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

In Table 6.10 I have introduced the tf.idf dictionaries' weights to the 2 years holding period return regression. None of the deaconries' weights coefficients showed any significance. I did the same for the different holding periods and got the same results.

Table 6.10: Introducing different dictionaries weights variables to (BHAR) over 2 years holding period regression

		TFNegative	TFHarvard_iv	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		0.001									
		(0.34)									
TFHarvard_iv			0								
			(0.06)								
TFConstraining				0.014							
				(0.45)							
TFLitigious					0.009						
					(0.54)						
TFSuperfluous						0.172					
						(1.39)					
TFModal							-0.011				
							(-0.18)				
TFUncertainty								0.006			
								(0.36)			
TFPositive									-0.006		
									(-0.34)		
TFInteresting										-0.092	
										(-1.53)	
TFIrr_Verb											0.014
											(0.32)
Spread	-0.037***	-0.038***	-0.037***	-0.038***	-0.038***	-0.039***	-0.037***	-0.038***	-0.037***	-0.035***	-0.038***
	(-2.76)	(-2.78)	(-2.75)	(-2.79)	(-2.8)	(-2.85)	(-2.73)	(-2.78)	(-2.72)	(-2.6)	(-2.77)
Log(proceeds)	-0.069	-0.079	-0.071	-0.085	-0.087	-0.082	-0.066	-0.077	-0.061	-0.042	-0.077
	(-0.83)	(-0.9)	(-0.81)	(-0.94)	(-0.97)	(-0.98)	(-0.78)	(-0.9)	(-0.71)	(-0.49)	(-0.89)
Market Dummy	-0.107	-0.1	-0.106	-0.098	-0.095	-0.082	-0.109	-0.102	-0.117	-0.131	-0.102
	(-0.74)	(-0.68)	(-0.73)	(-0.67)	(-0.65)	(-0.56)	(-0.75)	(-0.71)	(-0.79)	(-0.9)	(-0.71)
Multi-bookrunner	0.185	0.172	0.183	0.165	0.163	0.155	0.189	0.172	0.189	0.203	0.177
	(1.09)	(0.99)	(1.06)	(0.94)	(0.94)	(0.91)	(1.1)	(1)	(1.11)	(1.2)	(1.03)
Prestigious	-0.041	-0.042	-0.041	-0.042	-0.042	-0.032	-0.043	-0.04	-0.04	-0.036	-0.04
	(-0.47)	(-0.48)	(-0.47)	(-0.48)	(-0.48)	(-0.37)	(-0.49)	(-0.46)	(-0.46)	(-0.41)	(-0.46)
Underpricing	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	(-2.15)	(-2.13)	(-2.14)	(-2.12)	(-2.13)	(-2.05)	(-2.16)	(-2.14)	(-2.17)	(-2.22)	(-2.12)
Log(Lockup)	0.183	0.195	0.184	0.202	0.204	0.192	0.179	0.193	0.176	0.142	0.19
	(0.6)	(0.63)	(0.6)	(0.66)	(0.66)	(0.63)	(0.59)	(0.63)	(0.58)	(0.46)	(0.62)
Constant	0.325	0.285	0.319	0.262	0.247	0.269	0.339	0.291	0.36	0.464	0.293
	(0.39)	(0.34)	(0.38)	(0.31)	(0.29)	(0.32)	(0.41)	(0.35)	(0.43)	(0.56)	(0.35)
Observations	883	883	883	883	883	883	883	883	883	883	883
Prob > F	0.1332	0.165	0.1688	0.1622	0.1593	0.1141	0.1678	0.1646	0.1651	0.1043	0.1656
Adj R-squared	0.0081	0.0071	0.007	0.0072	0.0073	0.0092	0.007	0.0071	0.0071	0.0097	0.0071

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the Buy and Hold Abnormal Return (BHAR) over 2 years calculated using Index Return where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

6.12 Survival of IPOs analysis

From Table 6.11, I have the correlation between the dead dummy and the log(Days Till Dead) and the dictionaries' weights. Proportional Negative and Uncertainty are showing significant negative correlations. This shows that the exposure of negative and uncertainty information in the prospectus reduces the chances for the company to go dead. However, as the Log(Days Till Dead) variable in this table is only calculated for the companies that went dead, the negative correlation indicates that more negative tones in the prospectus will result in a shorter life for the company. For the dead dummy variable, litigious and Interesting dictionaries are showing a positive correlation. This means that exposure of more litigious and using more interesting language is correlated with the survival of the companies. On the other hand, Litigious is showing a negative correlation with Log(Days Till Dead). This indicates that more litigious language in the prospectus though it indicates a higher surviving chance, yet, it shows a shorter life among the dead companies. Significant negative correlation with Log(Days Till Dead) is also observed for Superfluous and Modal.

From Panel B, the Dead dummy is showing significant negative correlation with all the tf.idf weight dictionaries. This indicates higher survival probability in correlation with more exposure of information. However, among the bankrupted companies, more exposure results in shorter life till dead.

Table 6.11: Correlation of Bankruptcy variables with different dictionaries weights and variables

	Dead Dummy	Log(DaysTDead)
PrNegative	-0.123***	-0.204***
PrHarvard_iv	-0.002	-0.082
PrConstraining	0.024	-0.015
PrLitigious	0.084***	-0.146***
PrSuperfluous	0.038	-0.118**
PrModal	-0.014	-0.113**
PrUncertainty	-0.176***	-0.161***
PrPositive	0.001	0.074
PrInteresting	0.074**	0.02
PrIrr_Verb	0.179***	0.032
Panel A: Correlation with Proportional Weight Dictionaries		
TFNegative	-0.268***	-0.146***
TFHarvard_iv	-0.267***	-0.126**
TFConstraining	-0.235***	-0.169***
TFLitigious	-0.236***	-0.171***
TFSuperfluous	-0.173***	-0.01
TFModal	-0.251***	-0.085*
TFUncertainty	-0.283***	-0.137***
TFPositive	-0.197***	-0.1**
TFInteresting	-0.166***	-0.03
TFIrr_Verb	-0.256***	-0.058
Panel B: Correlation with tf.idf Weight Dictionaries		
Log(proceeds)	-0.219***	-0.004
Spread	0.073**	-0.09*
Prestigious	-0.008	0.062
STDev	0.042	0.021
Age in Days	-0.04	0.095*
Log(Lockup)		0.088*
Underpricing	0.03	-0.084*
Sales	-0.055	-0.202***
Overhang	0.061*	-0.009
Net Income	-0.038	-0.148**
Post-IPO EPS	0.054	0.084
Panel C: Correlation with Other Variables		

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table shows the correlation of the abnormal return. MAR0 is the Abnormal Return calculated using the Market Model calculated at the day of the lockup period expiration date. MCAR is the Cumulative Abnormal Return calculated using Market Model for periods of (2, 5 and 10 days). FFAR0 is the Abnormal Return calculated using Fama-French 4-Factor model. FFAR is the Cumulative Abnormal Return calculated using Fama-French 4-Factor Model. This table has three sections. In Panel A, it shows the correlation with the dictionaries proportional weights. In Panel B, it shows the correlations with the time-frequency – inverse document frequency dictionaries' weights. In Panel C it shows the correlations with other variables. Proceeds is the money raised in the IPO. The spread is the amount charged by the underwriter as a percentage of the IPO proceeds. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise. STDev is the standard deviation of share price during the first year of trading. Age in Days is the age of the company on the day of the IPO in days. Potential growth is a factor showing if the company increased the size of its capital at the years following the IPO. Lockup is the lockup period in days. the underpricing factor is calculated as the return on trading. If IPO was underpriced, the underpricing factor is positive, if it is overpriced, the underpricing factor is negative, and if fair price, the underpricing factor is 0. Sales is the reported sales the year prior to the IPO. Net Income is the reported net income the year prior to the IPO. Post IPO EPS is the reported Earnings Per Share the year prior to the IPO.

Panel C shows correlations with other variables. Log(Proceeds) is negatively correlated with the dead dummy. This shows that companies with larger IPOs have a higher probability to survive than those with smaller IPOs. Spread on the other hand is positively correlated with the dead dummy variable. This shows that underwriters charged slightly more fees to the companies that went dead. This could be an indication that the underwriters would charge more if a higher risk is anticipated. The Overhang variable is also showing a positive correlation with the Dead dummy variable. This indicates that the more shares kept by the shareholders, the higher the probability of going dead.

For Log(Days Till Dead) which in this table only measures the age in days for the companies that went dead, I can see a significant negative correlation with the Spread, Underpricing, Reported Sales before the IPO and the reported Net Income the year before the IPO. The negative correlation with the spread supports and goes in line with the discussion of the correlation between the dead dummy variable and the spread. In addition to that, the negative correlation with underpricing would also indicate that underwriters would underprice more if more risk is anticipated.

Among the companies that went dead, the companies that reported higher sales and higher net income for the year prior to the IPO had a shorter life. This will be an area to study their methods of reporting their sales and income and the level of discretionary measures in a different research.

Table 6.12: Regression of Log(TimeTillDead) with Proportional Weight Dictionaries

		PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
PrNegative		-30.815**									
		(-1.99)									
PrHarvard_iv			-2.816								
			(-0.43)								
PrConstraining				15.179							
				(1.21)							
PrLitigious					-12.254						
					(-1.09)						
PrSuperfluous						-346.526**					
						(-2.02)					
PrModal							-0.858				
							(-0.08)				
PrUncertainty								-2.927			
								(-0.19)			
PrPositive									5.019		
									(0.27)		
PrInteresting										-5.122	
										(-0.27)	
PrIrr_Verb											39.771*
											(1.85)
Spread	-0.024**	-0.023**	-0.024**	-0.024***	-0.024**	-0.024**	-0.024**	-0.024**	-0.024**	-0.023**	-0.023**
	(-2.58)	(-2.49)	(-2.58)	(-2.63)	(-2.6)	(-2.6)	(-2.57)	(-2.58)	(-2.57)	(-2.51)	(-2.5)
Log(proceeds)	-0.07	-0.057	-0.066	-0.079	-0.086	-0.082	-0.07	-0.07	-0.07	-0.07	-0.053
	(-1.11)	(-0.92)	(-1.03)	(-1.25)	(-1.34)	(-1.3)	(-1.11)	(-1.12)	(-1.12)	(-1.12)	(-0.84)
Market Dummy	-0.078	-0.081	-0.073	-0.094	-0.079	-0.073	-0.079	-0.08	-0.078	-0.08	-0.069
	(-0.7)	(-0.73)	(-0.65)	(-0.84)	(-0.71)	(-0.66)	(-0.7)	(-0.71)	(-0.7)	(-0.72)	(-0.62)
Multi-bookrunner	-0.501**	-0.474**	-0.505**	-0.499**	-0.516**	-0.514**	-0.501**	-0.5**	-0.5**	-0.497**	-0.514**
	(-2.36)	(-2.24)	(-2.38)	(-2.35)	(-2.43)	(-2.44)	(-2.36)	(-2.35)	(-2.35)	(-2.33)	(-2.43)
Prestigious	-0.022	-0.004	-0.02	-0.022	-0.022	-0.017	-0.022	-0.021	-0.022	-0.021	-0.011
	(-0.35)	(-0.07)	(-0.32)	(-0.36)	(-0.35)	(-0.27)	(-0.35)	(-0.34)	(-0.35)	(-0.34)	(-0.19)
Underpricing	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001*
	(-1.59)	(-1.44)	(-1.59)	(-1.6)	(-1.32)	(-1.53)	(-1.57)	(-1.55)	(-1.55)	(-1.61)	(-1.81)
Log(Lockup)	0.384	0.357	0.382	0.386*	0.365	0.393*	0.384	0.383	0.378	0.385*	0.44*
	(1.65)	(1.54)	(1.64)	(1.66)	(1.56)	(1.7)	(1.65)	(1.64)	(1.62)	(1.65)	(1.88)
Age in Days	1.34E-05	1.11E-05	1.33E-05	1.38E-05	1.33E-05	1.3E-5	1.33E-05	1.32E-05	1.3E-5	1.35E-05	1.12E-05
	(1.08)	(0.9)	(1.07)	(1.12)	(1.07)	(1.05)	(1.06)	(1.06)	(1.04)	(1.09)	(0.9)
Constant	6.485***	6.76***	6.578***	6.378***	6.671***	6.562***	6.497***	6.514***	6.47***	6.493***	5.962***
	(10.11)	(10.35)	(9.71)	(9.86)	(10.05)	(10.26)	(9.85)	(9.86)	(10.03)	(10.1)	(8.53)
Observations	314	314	314	314	314	314	314	314	314	314	314
Prob > F	0	0	0	0	0	0	0	0	0	0	0
Adj R-squared	0.1819	0.19	0.1797	0.1832	0.1824	0.1903	0.1792	0.1792	0.1793	0.1793	0.1885

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the log of time till dead in days where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 6.12 shows Logit Regression of Dead Dummy Variable with proportional weight dictionaries. I noticed from the table that Negative and Superfluous dictionaries have negative coefficients. This shows that these dictionaries have a negative relationship with the age of the company till it becomes dead. The Spread also shows a negative coefficient here which shows that underwriters charge slightly higher for the companies that have higher anticipated risk.

Multi-bookrunner dummy variable is showing a negative coefficient. This indicates that among the companies that went dead, the IPOs that has a multi-bookrunners have a shorter life. However, the prestigious dummy is showing insignificant here.

Table 6.13 shows similar data as Table 6.12 but with tf.idf weight dictionaries. None of the dictionaries shows significance in the reported regressions.

From Table 6.14 I noticed that Uncertainty has a significant negative coefficient. This means that the company is less likely to go dead as the tone on the prospectus is showing more uncertainty data. On the other hand, I can notice significant positive coefficients for litigious, constraining and less significant for Harvard Negative List. This shows that the likelihood for a company to go dead increases as the weight of the litigious or constraining tone of the prospectus goes higher. The same applies to the Harvard Negative List with less significance. This support our third hypothesis (H3) for the conservative tone part. I can notice that survival rate is less when the tone of the prospectus is more conservative. However, uncertainty is showing a negative coefficient which implies higher survival rate. On the other hand, positive tone does not show any significance here.

In addition to the tone of the prospectus, I also noticed that the coefficients of Log(proceeds) and Multi-bookrunners dummies are significant and negative. This shows that IPOs with higher proceeds and with multi-bookrunners are less likely to go dead that IPOs with smaller proceeds and one bookrunner.

Table 6.13: Regression of Log(TimeTillDead) with tf.idf Weight Dictionaries

		TFNegative	TFHarvard_iv	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		0.002									
		(0.45)									
TFHarvard_iv			0.001								
			(0.36)								
TFConstraining				0.014							
				(0.5)							
TFLitigious					-0.003						
					(-0.17)						
TFSuperfluous						0.113					
						(1.04)					
TFModal							0.01				
							(0.17)				
TFUncertainty								0.007			
								(0.37)			
TFPositive									-0.011		
									(-0.82)		
TFInteresting										0.046	
										(0.97)	
TFIrr_Verb											0.03
											(0.74)
Spread	-0.024**	-0.024**	-0.024**	-0.024***	-0.023**	-0.023**	-0.024**	-0.024**	-0.023**	-0.024***	-0.024***
	(-2.58)	(-2.61)	(-2.6)	(-2.62)	(-2.52)	(-2.55)	(-2.58)	(-2.6)	(-2.48)	(-2.64)	(-2.62)
Log(proceeds)	-0.07	-0.081	-0.078	-0.084	-0.066	-0.073	-0.072	-0.078	-0.054	-0.082	-0.084
	(-1.11)	(-1.2)	(-1.17)	(-1.22)	(-0.97)	(-1.16)	(-1.12)	(-1.17)	(-0.82)	(-1.28)	(-1.28)
Market Dummy	-0.078	-0.071	-0.074	-0.07	-0.081	-0.062	-0.076	-0.074	-0.094	-0.076	-0.067
	(-0.7)	(-0.63)	(-0.66)	(-0.62)	(-0.72)	(-0.55)	(-0.67)	(-0.66)	(-0.83)	(-0.68)	(-0.6)
Multi-bookrunner	-0.501**	-0.53**	-0.517**	-0.534**	-0.49**	-0.525**	-0.505**	-0.521**	-0.49**	-0.527**	-0.515**
	(-2.36)	(-2.39)	(-2.39)	(-2.4)	(-2.2)	(-2.46)	(-2.36)	(-2.38)	(-2.31)	(-2.46)	(-2.42)
Prestigious	-0.022	-0.023	-0.022	-0.021	-0.021	-0.024	-0.021	-0.021	-0.019	-0.024	-0.021
	(-0.35)	(-0.37)	(-0.36)	(-0.34)	(-0.34)	(-0.38)	(-0.34)	(-0.34)	(-0.31)	(-0.38)	(-0.34)
Underpricing	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001*	-0.001	-0.001
	(-1.59)	(-1.56)	(-1.55)	(-1.56)	(-1.59)	(-1.47)	(-1.56)	(-1.57)	(-1.68)	(-1.54)	(-1.51)
Log(Lockup)	0.384	0.4*	0.392*	0.406*	0.377	0.386*	0.386*	0.399*	0.367	0.395*	0.399*
	(1.65)	(1.7)	(1.68)	(1.71)	(1.59)	(1.66)	(1.65)	(1.69)	(1.57)	(1.7)	(1.71)
Age in Days	1.34E-05	1.37E-05	1.36E-05	1.35E-05	1.34E-05	1.39E-05	1.34E-05	1.33E-05	1.33E-05	1.38E-05	1.3E-5
	(1.08)	(1.1)	(1.1)	(1.09)	(1.08)	(1.12)	(1.08)	(1.07)	(1.07)	(1.11)	(1.05)
Constant	6.485***	6.434***	6.455***	6.419***	6.511***	6.459***	6.474***	6.44***	6.561***	6.448***	6.416***
	(10.11)	(9.86)	(9.97)	(9.78)	(9.87)	(10.06)	(10.01)	(9.85)	(10.12)	(10.03)	(9.89)
Observations	314	314	314	314	314	314	314	314	314	314	314
Prob > F	0	0	0	0	0	0	0	0	0	0	0
Adj R-squared	0.1819	0.1797	0.1795	0.1798	0.1792	0.1821	0.1792	0.1795	0.181	0.1817	0.1806

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 OLS Regressions of the log of time till dead in days where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 6.14: Logit Regression of Dead Dummy with proportional weight dictionaries

		PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
PrNegative		-10.004									
		(-0.29)									
PrHarvard_iv			24.733*								
			(1.74)								
PrConstraining				49.784**							
				(2.09)							
PrLitigious					61.193**						
					(2.25)						
PrSuperfluous						-0.551					
						(0)					
PrModal							17.227				
							(0.69)				
PrUncertainty								-74.803**			
								(-2.03)			
PrPositive									-3.367		
									(-0.08)		
PrInteresting										41.146	
										(0.9)	
PrIrr_Verb											191.26***
											(3.31)
Spread	0.005	0.006	0.005	0.005	0.007	0.005	0.006	0.003	0.005	0.004	0.007
	(0.22)	(0.24)	(0.23)	(0.22)	(0.31)	(0.22)	(0.24)	(0.13)	(0.22)	(0.18)	(0.28)
Log(proceeds)	-0.528***	-0.524***	-0.553***	-0.574***	-0.494***	-0.528***	-0.526***	-0.522***	-0.527***	-0.528***	-0.451***
	(-3.43)	(-3.4)	(-3.57)	(-3.69)	(-3.2)	(-3.41)	(-3.42)	(-3.38)	(-3.43)	(-3.43)	(-2.89)
Market Dummy	-0.068	-0.069	-0.104	-0.153	-0.13	-0.068	-0.071	-0.081	-0.068	-0.059	-0.057
	(-0.25)	(-0.26)	(-0.39)	(-0.56)	(-0.48)	(-0.25)	(-0.26)	(-0.3)	(-0.25)	(-0.22)	(-0.21)
Multi-bookrunner	-0.705*	-0.706*	-0.712*	-0.76*	-0.706*	-0.705*	-0.707*	-0.706*	-0.706*	-0.713*	-0.687*
	(-1.77)	(-1.78)	(-1.79)	(-1.89)	(-1.78)	(-1.77)	(-1.78)	(-1.77)	(-1.77)	(-1.79)	(-1.7)
Prestigious	0.166	0.169	0.144	0.149	0.174	0.166	0.17	0.165	0.167	0.159	0.174
	(1.05)	(1.06)	(0.9)	(0.93)	(1.09)	(1.05)	(1.07)	(1.04)	(1.05)	(1)	(1.09)
Underpricing	-0.002	-0.002	-0.001	-0.002	-0.002	-0.002	-0.002	-0.001	-0.002	-0.001	-0.002
	(-1.22)	(-1.21)	(-1.21)	(-1.26)	(-1.6)	(-1.22)	(-1.25)	(-1.06)	(-1.22)	(-1.2)	(-1.38)
Log(Lockup)	0.315	0.304	0.314	0.265	0.35	0.315	0.322	0.263	0.317	0.324	0.463
	(0.57)	(0.55)	(0.56)	(0.48)	(0.63)	(0.57)	(0.58)	(0.47)	(0.57)	(0.58)	(0.82)
Constant	-0.084	0.011	-0.819	-0.228	-0.767	-0.084	-0.309	0.703	-0.072	-0.192	-2.188
	(-0.06)	(0.01)	(-0.52)	(-0.15)	(-0.5)	(-0.06)	(-0.2)	(0.45)	(-0.05)	(-0.13)	(-1.33)
Observations	918	918	918	918	918	918	918	918	918	918	918
Prob > chi2	0	0	0	0	0	0	0	0	0	0	0
Pseudo R2	0.0983	0.0983	0.1007	0.1017	0.1023	0.0983	0.0986	0.1016	0.0983	0.0989	0.1073

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 Logit Regressions of the log of Dead Dummy where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 6.15 shows the same regressions using the tf.idf weights. Here I noticed that all the coefficients of the dictionaries weight that show significance are all negative. This includes Negative, Uncertainty, Modal, Harvard Negative List and less significance for Superfluous. For those dictionaries, the regressions show that the more usage of these dictionaries in the prospectuses the less likely it is for the company to go dead.

Table 6.16 Shows the Cox Proportional Hazard Model using Dead Dummy Variable and DaysTillDead variable. In this table, I have used a censored data which only uses companies that went dead during the time of the study. From the first column, I can see the regression without the inclusion of any of the dictionaries weights. From this regression, I can notice that the coefficient of Spread variable is in advantage of the company as its value is less than 1. This indicates that the higher the charged fees, the higher chance for the company to survive. However, Multi-bookrunner variable coefficient is showing a value that is larger than 1. This indicates that companies that have multi-bookrunners are more likely to go dead. In the other columns, I have introduced different proportional weights for different dictionaries. They all were insignificant except for Superfluous. The coefficient value is higher than 1. This shows that prospectuses with more superfluous content have a higher likelihood to go dead. I have also run the regression with the tf.idf where none of the coefficients of the dictionaries' weights was significant.

Table 6.15: Logit Regression of Dead Dummy with tf.idf weight dictionaries

		TFNegative	TFHarvard_iv	TFConstraining	TFLitigious	TFSuperfluous	TFModal	TFUncertainty	TFPositive	TFInteresting	TFIrr_Verb
TFNegative		-0.022**									
		(-2.4)									
TFHarvard_iv			-0.019***								
			(-2.81)								
TFConstraining				-0.009							
				(-0.14)							
TFLitigious					-0.033						
					(-1.02)						
TFSuperfluous						-0.466*					
						(-1.85)					
TFModal							-0.412***				
							(-3.33)				
TFUncertainty								-0.132***			
								(-3.58)			
TFPositive									-0.019		
									(-0.58)		
TFInteresting										-0.152	
										(-1.31)	
TFIrr_Verb											-0.225**
											(-2.55)
Spread	0.005	0.012	0.012	0.006	0.009	0.008	0.011	0.014	0.007	0.008	0.012
	(0.22)	(0.5)	(0.51)	(0.23)	(0.37)	(0.34)	(0.47)	(0.57)	(0.27)	(0.34)	(0.48)
Log(proceeds)	-0.528***	-0.393**	-0.394**	-0.519***	-0.465***	-0.502***	-0.421***	-0.369**	-0.503***	-0.485***	-0.412**
	(-3.43)	(-2.41)	(-2.44)	(-3.11)	(-2.81)	(-3.25)	(-2.67)	(-2.3)	(-3.15)	(-3.08)	(-2.57)
Market Dummy	-0.068	-0.158	-0.168	-0.074	-0.108	-0.146	-0.139	-0.166	-0.1	-0.105	-0.138
	(-0.25)	(-0.58)	(-0.61)	(-0.27)	(-0.4)	(-0.53)	(-0.51)	(-0.6)	(-0.36)	(-0.38)	(-0.51)
Multi-bookrunner	-0.705*	-0.553	-0.566	-0.694*	-0.631	-0.642	-0.588	-0.501	-0.695*	-0.684*	-0.616
	(-1.77)	(-1.37)	(-1.41)	(-1.71)	(-1.56)	(-1.6)	(-1.47)	(-1.24)	(-1.75)	(-1.72)	(-1.54)
Prestigious	0.166	0.171	0.17	0.167	0.167	0.149	0.141	0.151	0.172	0.173	0.16
	(1.05)	(1.08)	(1.07)	(1.05)	(1.05)	(0.94)	(0.88)	(0.94)	(1.08)	(1.09)	(1.01)
Underpricing	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-1.22)	(-1.35)	(-1.4)	(-1.22)	(-1.25)	(-1.35)	(-1.45)	(-1.32)	(-1.27)	(-1.26)	(-1.42)
Log(Lockup)	0.315	0.202	0.218	0.307	0.25	0.318	0.266	0.162	0.308	0.271	0.268
	(0.57)	(0.36)	(0.39)	(0.55)	(0.45)	(0.57)	(0.48)	(0.29)	(0.55)	(0.49)	(0.48)
Constant	-0.084	0.332	0.337	-0.056	0.162	0.007	0.273	0.463	-0.01	0.086	0.253
	(-0.06)	(0.22)	(0.22)	(-0.04)	(0.11)	(0)	(0.18)	(0.3)	(-0.01)	(0.06)	(0.17)
Observations	918	918	918	918	918	918	918	918	918	918	918
Prob > chi2	0	0	0	0	0	0	0	0	0	0	0
Pseudo R2	0.0983	0.1031	0.1049	0.0983	0.0991	0.1011	0.1078	0.1094	0.0985	0.0997	0.1036

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 Logit Regressions of the log of Dead Dummy where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

Table 6.16: Cox Proportional Hazard Model using Dead and DaysToDead variables

		PrNegative	PrHarvard_iv	PrConstraining	PrLitigious	PrSuperfluous	PrModal	PrUncertainty	PrPositive	PrInteresting	PrIrr_Verb
PrNegative		37.36									
		(1.19)									
PrHarvard_iv			-3.246								
			(-0.26)								
PrConstraining				-32.063							
				(-1.46)							
PrLitigious					17.002						
					(0.77)						
PrSuperfluous						726.619**					
						(2.22)					
PrModal							-10.118				
							(-0.52)				
PrUncertainty								-24.972			
								(-0.86)			
PrPositive									28.937		
									(0.81)		
PrInteresting										-0.506	
										(-0.02)	
PrIrr_Verb											-53.659
											(-1.35)
spread	0.041**	0.038**	0.041**	0.042**	0.04**	0.04**	0.041**	0.04**	0.042**	0.041**	0.042**
	(2.43)	(2.26)	(2.43)	(2.53)	(2.39)	(2.38)	(2.43)	(2.38)	(2.48)	(2.41)	(2.51)
log(proceeds)	0.02	-0.001	0.023	0.051	0.024	0.034	0.021	0.023	0.015	0.02	0
	(0.19)	(-0.01)	(0.22)	(0.47)	(0.23)	(0.32)	(0.2)	(0.21)	(0.14)	(0.19)	(0)
Market Dummy	0.111	0.101	0.114	0.148	0.076	0.141	0.115	0.114	0.108	0.111	0.079
	(0.55)	(0.5)	(0.56)	(0.72)	(0.36)	(0.69)	(0.56)	(0.56)	(0.53)	(0.55)	(0.39)
Multi-bookrunner	1.166***	1.191***	1.17***	1.229***	1.174***	1.226***	1.16***	1.159***	1.193***	1.166***	1.255***
	(3.05)	(3.11)	(3.06)	(3.19)	(3.08)	(3.2)	(3.02)	(3.02)	(3.11)	(3.04)	(3.28)
prestigious	0.065	0.052	0.065	0.048	0.071	0.061	0.061	0.065	0.059	0.065	0.058
	(0.57)	(0.46)	(0.58)	(0.43)	(0.63)	(0.54)	(0.54)	(0.57)	(0.52)	(0.57)	(0.51)
Underpricing	0.002*	0.002*	0.002*	0.002**	0.002*	0.002*	0.002*	0.002*	0.002*	0.002*	0.002**
	(1.84)	(1.7)	(1.85)	(1.99)	(1.65)	(1.69)	(1.88)	(1.92)	(1.92)	(1.84)	(2.01)
log(Lockup)	-0.624	-0.57	-0.613	-0.555	-0.597	-0.671	-0.625	-0.637	-0.635	-0.624	-0.698
	(-1.36)	(-1.24)	(-1.33)	(-1.2)	(-1.3)	(-1.45)	(-1.36)	(-1.39)	(-1.38)	(-1.36)	(-1.52)
Observations	388	388	388	388	388	388	388	388	388	388	388
Prob > chi2	0	0	0	0	0	0	0	0	0	0	0

*Significant at 10%, **Significant at 5%, ***Significant at 1%

This table summarises 11 Cox Hazard Models where each one is presented in one of the columns. The first column shows the regression without introducing any of the dictionaries weight variables. The following columns show regressions after introducing the dictionaries one at a time. The control Variables are the log of the proceeds adjusted for CPI index, Market dummy is 0 for the Main market and 1 for AIM market, Multi-bookrunner is a dummy with a value of 1 if the IPO has more than 1 bookrunner and 0 if only one bookrunner. Prestigious is a dummy with a value of 1 if the IPO has a prestigious underwriter and 0 otherwise, D1 Underpricing is the return on the first day of trading if purchased at the IPO. Log(Lockup) is the log of lockup period. I also controlled for years but are not reported here in this table.

6.13 Conclusion

To shed light on some of the puzzles in IPO literature, I examined the tone from IPO prospectuses and relate that to IPO lockup expiration, long-run IPO performance and survival of the IPOs. Though there have been a few studies on textual analysis and IPO short-run dynamics in the US, to the best of our knowledge, this is the first study to analyse long-run IPO dynamics using textual analysis.

The UK IPO market structure is different from the US in a number of ways. Lockup contracts are standardised at 180 days in the US, while heterogeneous in the UK. Also, their lengths are longer. I found a median lockup length of 388 days. Much information is produced and disclosed during the lockup period. Therefore, information asymmetry is of less important during the lockup period in the UK. I found evidence in line with this as our prospectus information does not explain the lockup expiration returns. I also find a huge variation in the long-run performance of IPOs. The information content of the prospectus to explain long-run IPO performance is rather limited.

While previous studies show that accounting information can predict the IPO survival (e.g., Demers and Joose, 2007), I show that IPO tone measures can predict the IPO survival. Hensler et al., (1997) use the variable risk from Jay Ritter's website which is mentioned as risk and relate it to IPO survival. The risk variable is just the count of the number of risk factors referred to in the prospectus. Recognising the limitation of previous studies, I use the Loughrun and McDonald (2011) dictionary measures in the prospectus to relate it to the survival firms of the IPOs. I find a significant relationship with IPO tone and survival of IPOs. This is important if I can distinguish ex-ante the quality of the companies that joins the London Stock

Exchange. However, to what extent our model is predictive is an empirical issue and could be agenda for further research.

Chapter 7

7. Conclusions and Directions for future research

This thesis is a fresh attempt to understand one of the most important IPO market in the world—the London Stock Exchange. The alternative Investment Market (AIM) is one of the successful markets for small, young and growth companies. Though there are few studies on different aspects of AIM (e.g., Doukas and Hoque, 2016 examine the reasons why companies prefer AIM as compared to the Main Market), there are some gaps in the literature that I examined in this theses. For example, no previous study examines the cost of raising money in the London Stock Exchange, particularly, for AIM. This thesis is an effort to understand different dimensions of the IPO market. I analysed the role of the underwriters in the IPO process from two main aspects. I studied underwriter fees in term of the spread of the total proceeds of the IPO, and tone they use in writing the IPO prospectuses. I also related the tone of the IPO prospectus with several IPO dynamics in the short-run and over the long-run. I used IPO spread, underpricing, lockup length, standard deviation and idiosyncratic risk as short-term IPO phenomenon and use lockup expiration returns, long-run returns and survival of IPOs as a long-run phenomenon to relate to IPO.

I constructed a unique database of IPOs from the London Stock Exchange. Since there is no readily available information on IPO spread and it is not reported in any databases, I collect the information from IPO prospectuses. I collected underwriter spread, lockup information and other IPO characteristics and create a unique database. I also machine read the tone of IPOs from prospectuses. My sample contains the IPOs that took place in both the Main Market and the AIM of London Stock Exchange for a period of 13 years starting from 1999 to 2012. I excluded the financial companies,

the companies that are not incorporated in the UK and those with missing information. Our final sample contains a total of 972 IPOs.

The Alternative Investment market is considered as one of the most popular markets for smaller companies (Vismara et al., 2013). Yet, to the best of our knowledge, there are no studies that try to examine the cost of raising money in such a market. I tried to fill the gap in the literature. The first empirical chapter is one of the first efforts which tries to shed light on the underwriter's role in bringing a company public in the Alternative Investment Market. When I calculate the fees charged (spread) as a percentage of the total IPO proceeds, our results show that the overall spread for the London Stock Exchange ranges between 4% and 6.43% with a median of 5%. Our findings show that fees are not clustered like the US market (Chen and Ritter, 2000), they vary across the IPOs. As the proceeds of the IPOs increase, the spread decreases. This result holds in the multivariate regression setting. Another factor that affects the fees is the year of the IPO.

But when I partition the markets into the Main market and AIM, I find that the spread is much higher in the AIM as compared to the Main Market. My results show that the spread charged in AIM market has a mean of 6.47% and a median of 5.08% compared to a mean of 4.04% and a median 4.00% of for the Main Market. Then I concentrated on the question why bookrunners in the AIM charge higher fees as compared to the Main Market. I examined a number of factors such as the lockup length, the age of the company at the time of the IPO, some risk factors such as idiosyncratic risk, and the potential growth.

To investigate the reason for the higher fees in the AIM market, I tested the dis-economies of scale as the main reason for charging higher fees in the AIM because

IPO size is larger in the Main Market. I examined this hypothesis from three different aspects. First, I examine the regression model for the spread. My findings show that log(proceeds) has a negative relationship with the spread. The second aspect is that I noticed a convergence to 5% that continues below it as the value of the proceeds increases as per the distribution of the fees charged. The third aspect is that underpricing is higher for smaller IPOs. This might be an indication that the implicit cost of raising money, that is, underpricing is higher for smaller IPOs. Our findings show that economies of scale are a strong determinant of gross spread. The higher the money raised, the lower the cost.

I tested the hypotheses that whether the risk is higher in the AIM and that's why they pay more. I calculated idiosyncratic risk and the standard deviation for the share price. However, none of the risk proxies is significant. I cannot conclude that by using our measures that higher the risk higher the gross spread. Our measures for risk could be problematic as all uses the share price data after the IPO. Previous research also raises questions about the use of aftermarket variables to measure the ex-ante risk, e.g., Yung and Zender, (2010). In the second chapter, I use the "tone" of the IPOs to measure the ex-ante risk of the IPOs. To the extent tone can measure ex-ante uncertainty is an empirical question, and I try to address it in chapter 5.

I tested whether the lockup period mitigates some of the risks in IPOs. Our findings show a significant positive relationship with smaller IPOs and in AIM market IPOs. Because bookrunners act as nominated advisors in the AIM, I also investigate the contribution of a possible rent-seeking to the overall fees charged. I used the lockup length as a proxy for rent-seeking. My findings show that the coefficient has a positive value. This implies that investment bankers charge higher spread and impose higher

lockup on the company at the same time, which might be consistent with the rent-seeking behaviour. However, further research is needed to shed more lights on the rent-seeking behaviour and underwriters spread.

I examined the use of under-pricing as an alternative mechanism to mitigate the risk of IPOs. The indirect cost of IPO, underpricing, is negatively related to the gross spread. Hence, underpricing works as a substitute of spread. I also investigate the pricing of the shares as an early indicator of the underwriters' proof of their due diligence and hence the post-IPO performance. Our findings suggest that under-pricing is less in the Main Market compared to the AIM. I calculated underpricing on both first and fifth trading days. IPOs issued in AIM show that they are underpriced when calculated on the first trading day. When I calculate it for the fifth day of trading, I noticed that the gap gets narrower. When looking at the median of the share return on the first and fifth trading days, my findings imply that the underwriters try to come up with a fair price.

I also examine the hypotheses that fees charged by the underwriters works as a marketing tool to attract future business. I examined two models in terms of the number of IPOs and value of proceeds. I did not find evidence that gross spread charged works as a marketing tool. Nevertheless, the results show a positive relationship between the average proceeds from one year to the number of IPOs of the following year. However, these tests are far from conclusive. More research is needed to discern whether the fees charged work as a marketing tool as part of the small boutique banks in the AIM and is an agenda for future research.

Motivated by the different IPO market structure in the UK as compared to the IPO market structure in the US in a number of ways, this thesis uses the tone

information from IPO prospectus to relate it to a number of IPO mechanisms. For example, in the US book building is the preferred method while in the UK placements and open offer are the methods in use. IPO gross spreads are clustered at 7% in the US. In comparison, there is a large variation in IPO spread in the UK, ranging from 4-11%. Lockup contracts are standardised at 180 days in the US, while heterogeneous in the UK. Also, price support is very common in the US, but not in the UK. Because of these differences, I believe, UK serves as an out of sample test on the issues tested in the US using textual data from IPO prospectuses.

In the second empirical chapter, I measured the IPO tone using textual analysis and relate it to a number of short-run IPO dynamics such as IPO underpricing, spread, lockup length, volatility and idiosyncratic risk using hand-collected data from London Stock Exchange. Though there have been a few studies on textual analysis and IPO short-run dynamics in the US, to the best of our knowledge, this is the first study to analyse IPO dynamics using textual analysis in the UK.

I analysed the prospectuses using 10 dictionaries as proposed by Loughrun and McDonald (2011). I categorise the dictionaries into three main groups— Negative, Positive and Uncertainty. For each dictionary, I calculated the dictionary weight in two different ways. The first way is by simply using a proportional weight of the words from the dictionary to the total word count in the document. The second way is the time-frequency – inverse document frequency weight. The second way considers the weight of each word across all the text used in all the prospectuses in our sample. Then it puts more weight on the words that are less commonly used.

In chapter two, I examined a number of hypotheses. First, I investigated if underwriters charge higher spread when the tone used in prospectuses fall under the conservative and the uncertainty categories groups and higher when it falls under the positive category group, and I find evidence in line with this. These results imply that if there are more uncertainty words the spread charged is higher. Thus the uncertainty words measure the ex-ante uncertainty of the IPOs. I find a significant relationship with IPO tone and spread implying that underwriters charge more for risky companies. I contribute to the existing literature by providing an alternative measure of risk which captures the ex-ante uncertainty of IPOs. Though previous literature has used a number of variables like size, age, offer price, standard deviation to measure the ex-ante uncertainty of an IPO, they could be measuring many other things. Our uncertainty tone could be a better measure than those proposed in the literature.

Then I tested if underwriters underprice less when they use a conservative tone and underprice more when they use uncertainty or positive tone in the prospectuses. Our findings show a significant relationship between uncertainty tone and underpricing. Again, the ex-ante uncertainty increases the underpricing which has been documented in a number of earlier studies. I provide evidence with an alternate measure of risk that ex-ante uncertainty increases underpricing from the UK market, which uses different methods to raise money as compared to the US. I examine whether the tone of the IPO prospectus is related to the lockup period. Do the underwriters impose shorter lockup period when underwriters use a conservative tone and longer lockup period when they use Uncertainty or Positive tones in the prospectuses? Our findings show a significant relationship with IPO tone and lockup length. These findings of under-pricing and the lockup period imply certain words in the IPO prospectuses reduce information asymmetry, particularly the negative words.

These findings are in line with the earlier works done by using information asymmetry as a major reason for underpricing. To the best of our knowledge, there is no study which relates IPO tone with lockup length. I found that certain words significantly related to the lockup length. This is consistent with the earlier work done on lockups in the US and the UK (e.g., Brav and Gompers, 2003, in the US; Hoque, 2011, in the UK).

I also related whether share price volatility is related to the tone of IPO. I examine whether standard deviation is lower when underwriters use a conservative tone, higher when they use uncertainty tone and no effect when they use a positive tone in the prospectuses. I did not find any relationship between IPO tone and volatility. I also tested the hypothesis that idiosyncratic risk would be lower when underwriters use a conservative or uncertainty tones and no affected by the positive tone in the prospectuses. For both the hypotheses, our results show no relationship between IPO tone and volatility or idiosyncratic risk. I propose few explanations. As new information is released after the IPO, the new information may be more important than the information reported in the prospectuses. In addition, the volatility and idiosyncratic risks might not be good measures of ex-ante uncertainty of IPO risk. These issues are agendas for future research.

In the third empirical chapter, I used the IPO tone measured in the same way following Loughrun and McDonald (2011) in previous empirical chapter and relate it to a number of long-run IPO dynamics such as cumulative abnormal return (CAR) after lockup expiration date, long-term buy-and-hold abnormal return (BHAR) and companies' survival rate. To the best of our knowledge, this is the first study to analyse IPO long-run dynamics using textual analysis in the UK.

In this chapter, I examined a number of issues. The first event I used is the market reaction around lockup expiration day. Though lockup expiration is a known event and reported in the IPO prospectus, share price drops around 2.00% around the lockup expiry day (Hoque, 2011). Previous research shows that lockup expiration returns in related to downward sloping demand curve and worse than expected insider selling (Field and Hanka, 2001). The first hypothesis in this chapter examines if the lockup expiration return will be less when underwriters use a conservative tone and will be more when they use Uncertainty or Positive tones in the prospectuses. I found significant relationship with IPO positive and superfluous tones and $CAR_{(0,+1\ )}$. Our finding imply that positive tone reduces $CAR_{(0,+10)}$. However, superfluous tone which implies uncertainty resulted increased $CAR_{(0,+10)}$.

As the second IPO anomaly, I use the long-run return. I examined whether the long-run return will be less when underwriters use a conservative tone and will be more when they use uncertainty or positive tones in the prospectuses. I calculated BHAR using Market Model and Fama-French Four Factor Model. Our findings show a significant relationship with IPO tone and BHAR. More usage of interesting words in prospectuses decreases BHAR over 6 months period. However, the use of superfluous words increases BHAR over a 1-year period. I did not find significance with the other dictionaries or the longer holding period. In sum, there is a lack of clear relationship between the tone of IPOs and long-run IPO returns. This indicated that the effect of the IPO tone decays in terms of return over longer periods. These results might be driven by the problems of measurement in the long-run IPO returns as reported in previous studies (Schultz, 2003). Our findings also show that using prestigious underwriters, charging higher spread, longer lockup period and more underpricing increases the IPO performance over the longer run.

The third phenomenon I test whether IPO survival rate is lower when underwriters use a conservative tone and is higher when they use uncertainty or positive tones in the prospectuses. I used the dead dummy and the log of the number of days until the company goes dead. I found a significant relationship with IPO tone and the probability of the company to go dead. When studying the proportional weight of dictionaries, our results suggest that exposure of negative and uncertainty information is related to less likelihood of the company to go dead. However, higher usage of a litigious tone is related to increase the chance of the company to go dead. Positive tone represented by interesting dictionary also increases the chances of going dead. Findings from testing tf.idf weighted dictionaries that show negative coefficients across all dictionaries with dead dummy suggest that more exposure decreases the probability for the company to go dead. Our findings also imply that larger IPOs with larger proceeds are less likely to go dead compared to smaller IPOs. Among the companies that went dead, I found a significant relationship with the IPO tone and how soon they go dead. Our findings suggest that more exposure will result in a shorter lifetime for those companies. In addition, companies that reported positive pre-IPO sales and net income have shorter time span among the companies that went dead. The methods of reporting pre-IPO sales and income and the level of discretionary measures in the IPOs could be agenda for future research.

There has been a growing body of literature on IPOs. Yet, literature has reported a number of anomalies like underpricing, lockup expiration returns and long-run underperformance of IPOs. By using the tone information from prospectuses, I examine a number of such IPO phenomenon like underpricing, IPO spread, lockup length, standard deviation, idiosyncratic risk, lockup expiration returns, long-run IPO performance and survival of IPOs. Our results show that “tone” of IPO prospectuses

could provide valuable information in explaining most of the IPO phenomenon that I use. These results are consistent with the information asymmetry that exists between the managers and the outside shareholders. Our results are consistent with the notion that tone information mitigates the information asymmetry to some extent. However, to what extent “tone information” can be used to design profitable trading strategy is not addressed in this theses and are avenues for future research.

Chapter 8

8. Appendix

Table 8.1 Key Terms Difinition

Term	Definition
Dictionary	A group of words that convey a certain tone such as (negative, positive, uncertainty)
Idiosyncratic	The risk factor that is related to the firm itself without the effect of the market
lockup	A contractual restriction that prevent the insiders whom had ownership of the company prior to the IPO from selling their shares
lockup length	The length of the period of the Lockup.
Multi-bookrunner	If an IPO is undertaken by more than one bookrunner then it is a multi-bookrunner IPO
OLS	Ordinary Least Square regression
Overhang	The percentage of share that are not traded
Potential growth	The increase of the size of a firm in the years following the IPO
Proceed	The money raised in the IPO
Spread	The amount charged by the underwriter as a percentage of the IPO proceeds
tf.idf	Time Frequency. Inverse Document Frequency. Used to weigh of a word in a document among a set of documents.
Tone	The sentiment of a document
Underpricing	Calculated as the return of the first-trading day after the IPO
Underwriter	The investment firm that undertake the process of taking a company the public market

9. Reference List

Abrahamson, M., Jenkinson, T., & Jones, H. (2011). Why Don't US Issuers Demand European Fees for IPOs? *The Journal of Finance, 66*(6), 2055-2082.

Aggarwal, R., Leal, R., & Hernandez, L. (1993). The aftermarket performance of initial public offerings in Latin America. *Financial Management*, 42-53.

Aggarwal, R., & Rivoli, P. (1990). Fads in the initial public offering market? *Financial Management*, 45-57.

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

Ahmad, W., & Jelic, R. (2014). Lockup agreements and survival of UK IPOs. *Journal of Business Finance & Accounting, 41*(5-6), 717-742.

Allen, F., & Faulhaber, G. R. (1988). Optimism invites deception. *The Quarterly Journal of Economics, 103*(2), 397-407.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*(2), 223-249.

Anthony, L. (2015). Retrieved from http://www.laurenceanthony.net/software/antwordprofiler/

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of Finance, 59*(3), 1259-1294.

Armitage, S. (2000). The direct costs of UK rights issues and open offers. *European Financial Management, 6*(1), 57-68.

Arnold, T., Fishe, R. P., & North, D. (2010). The effects of ambiguous information on initial and subsequent IPO returns. *Financial Management, 39*(4), 1497-1519.

Audretsch, D. B., & Lehmann, E. E. (2005). The effects of experience, ownership, and knowledge on IPO survival: Empirical evidence from Germany. *Review of Accounting and Finance, 4*(4), 13-33.

Barry, C. B., Muscarella, C. J., Peavy, J. W., & Vetsuypens, M. R. (1990). The role of venture capital in the creation of public companies: Evidence from the going-public process. *Journal of Financial Economics, 27*(2), 447-471.

Beatty, R. P., & Ritter, J. R. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics, 15*(1), 213-232.

Beatty, R. P., & Welch, I. (1996). Issuer expenses and legal liability in initial public offerings. *JL & Econ., 39*, 545.

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

Bhattacharyya, D., Das, P., Mitra, K., Ganguly, D., Das, P., Bandyopadhyay, S. K., & Kim, T.-h. (2009). In Search of Unstructured Documents Categorization *Advances in Communication and Networking* (pp. 116-125): Springer.

Boehmer, B., Boehmer, E., & Fishe, R. P. (2006). Do institutions receive favorable allocations in IPOs with better long-run returns? *Journal of Financial and Quantitative Analysis, 41*(04), 809-828.

Booth, J. R., & Smith, R. L. (1986). Capital raising, underwriting and the certification hypothesis. *Journal of Financial Economics, 15*(1), 261-281.

Bradley, D. J., Jordan, B. D., Yi, H. C., & Roten, I. C. (2001). Venture capital and IPO lockup expiration: An empirical analysis. *Journal of Financial Research, 24*(4), 465-493.

Brau, J. C., Carter, D. A., Christophe, S. E., & Key, K. G. (2004). Market reaction to the expiration of IPO lockup provisions. *Managerial Finance, 30*(1), 75-91.

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

Brau, J. C., Lambson, V. E., & McQueen, G. (2005). Lockups revisited. *Journal of Financial and Quantitative Analysis, 40*(03), 519-530.

Brav, A., & Gompers, P. A. (1997). Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. *The Journal of Finance, 52*(5), 1791-1821.

Brav, A., & Gompers, P. A. (2000). Insider trading subsequent to initial public offerings: Evidence from expirations of lock-up provisions. *working paper*.

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

Brennan, M. J., & Franks, J. (1997). Underpricing, ownership and control in initial public offerings of equity securities in the UK. *Journal of Financial Economics, 45*(3), 391-413.

Cai, J., & Wei, K. J. (1997). The investment and operating performance of Japanese initial public offerings. *Pacific-Basin Finance Journal, 5*(4), 389-417.

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

Carter, R. B. (1992). Underwriter reputation and repetitive public offerings. *Journal of Financial Research, 15*(4), 341-354.

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

Chadha, S. (2003). Do insiders knowingly issue overvalued equity? Evidence from IPOs that get delisted. *working paper*.

Chambers, D., & Dimson, E. (2009). IPO underpricing over the very long-run. *The Journal of Finance, 64*(3), 1407-1443.

Chemmanur, T. J. (1993). The pricing of initial public offerings: A dynamic model with information production. *The Journal of Finance, 48*(1), 285-304.

Chen, C. R., & Mohan, N. J. (2002). Underwriter spread, underwriter reputation, and IPO underpricing: A simultaneous equation analysis. *Journal of Business Finance & Accounting, 29*(3-4), 521-540.

Chen, H. C., & Ritter, J. R. (2000). The seven percent solution. *The Journal of Finance, 55*(3), 1105-1131.

Christie, W. G., & Schultz, P. H. (1994). Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes? *The Journal of Finance, 49*(5), 1813-1840.

Clarkson, P. M., & Merkley, J. (1994). Ex-ante uncertainty and the underpricing of initial public offerings: Further Canadian evidence. *Canadian Journal of Administrative Sciences, 11*(1), 54.

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

Courteau, L. (1995). Under-diversification and retention commitments in IPOs. *Journal of Financial and Quantitative Analysis, 30*(04), 487-517.

Demers, E., & Joos, P. (2007). IPO failure risk. *Journal of Accounting Research, 45*(2), 333-371.

Demers, E., & Lewellen, K. (2003). The marketing role of IPOs: evidence from internet stocks. *Journal of Financial Economics, 68*(3), 413-437.

Demers, E., & Vega, C. (2008). *Soft information in earnings announcements: News or noise?* : Federal Reserve Board.

Derrien, F., & Kecskes, A. (2007). The initial public offering of listed firms. *Journal of Finance, 62*(1), 447-479.

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance, 46*(4), 1325-1359.

Doukas, J. A., & Hoque, H. (2016). Why firms favour the AIM when they can list on main market? *Journal of International Money and Finance, 60*, 378-404.

DuCharme, L., Rajgopal, S., & Sefcik, S. (2001). Lowballing for'Pop': The case of internet IPO underpricing. *Available at SSRN 257725*.

Dunbar, C. G. (2000). Factors affecting investment bank initial public offering market share. *Journal of Financial Economics, 55*(1), 3-41.

Dutta, P. K., & Madhavan, A. (1997). Competition and collusion in dealer markets. *The Journal of Finance, 52*(1), 245-276.

Easley, D., & O'hara, M. (2004). Information and the cost of capital. *The Journal of Finance, 59*(4), 1553-1583.

Eckbo, B. E., & Norli, Ø. (2001). Risk and long-run IPO returns. *Unpublished Working Paper (Tuck School of Business, Dartmouth College, NH)*.

Elliott, W. B., Rennekamp, K. M., & White, B. J. (2015). Does concrete language in disclosures increase willingness to invest? *Review of Accounting Studies, 20*(2), 839-865.

Engelberg, J. (2008). *Costly information processing: Evidence from earnings announcements.* Paper presented at the AFA 2009 San Francisco Meetings Paper.

Espenlaub, S., Gregory, A., & Tonks, I. (2000). Re-assessing the long-term underperformance of UK Initial Public Offerings. *European Financial Management, 6*(3), 319-342.

Espenlaub, S., Khurshed, A., & Mohamed, A. (2012). IPO Survival in a Reputational Market. *Journal of Business Finance & Accounting, 39*(3-4), 427–463.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56.

Fama, E. F., & French, K. R. (2004). New lists: Fundamentals and survival rates. *Journal of Financial Economics, 73*(2), 229-269.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2008). The incremental information content of tone change in management discussion and analysis. *Working paper, Insead*.

Ferris, S. P., Hao, Q., & Liao, M.-Y. (2013). The Effect of Issuer Conservatism on IPO Pricing and Performance. *Review of Finance, 17*(3), 993-1027.

Field, L. C., & Hanka, G. (2001). The expiration of IPO share lockups. *The Journal of Finance, 56*(2), 471-500.

Fink, J., Fink, K., Grullon, G., & Weston, J. P. (2005). *IPO vintage and the rise of idiosyncratic risk*.

Finn, F. J., & Higham, R. (1988). The performance of unseasoned new equity issues-cum-stock exchange listings in Australia. *Journal of Banking & Finance, 12*(3), 333-351.

Goergen, M., Renneboog, L., & Khurshed, A. (2006). Explaining the diversity in shareholder lockup agreements. *Journal of Financial Intermediation, 15*(2), 254-280.

Gompers, P. A., & Lerner, J. (1997). Venture capital and the creation of public companies: do venture capitalists really bring more than money? *The Journal of Private Equity, 1*(1), 15-32.

Gregory, A., Tharyan, R. and Christidis, A., 2013. Constructing and testing alternative versions of the Fama–French and Carhart models in the UK. Journal of Business Finance & Accounting, 40(1-2), pp.172-214.

Grinblatt, M., & Hwang, C. Y. (1989). Signalling and the pricing of new issues. *The Journal of Finance, 44*(2), 393-420.

GUO, R. J., Lev, B., & Zhou, N. (2004). Competitive costs of disclosure by biotech IPOs. *Journal of Accounting Research, 42*(2), 319-355.

Habib, M. A., & Ljungqvist, A. P. (2001). Underpricing and entrepreneurial wealth losses in IPOs: Theory and evidence. *Review of Financial Studies, 14*(2), 433-458.

Hamao, Y., Packer, F., & Ritter, J. R. (2000). Institutional affiliation and the role of venture capital: Evidence from initial public offerings in Japan. *Pacific-Basin Finance Journal, 8*(5), 529-558.

Hanley, K. W. (1993). The underpricing of initial public offerings and the partial adjustment phenomenon. *Journal of Financial Economics, 34*(2), 231-250.

Hanley, K. W., & Hoberg, G. (2010). The information content of IPO prospectuses. *Review of Financial Studies, 23*(7), 2821-2864.

Hanley, K. W., & Hoberg, G. (2012). Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics, 103*(2), 235-254.

Hansen, R. S. (2001). Do investment banks compete in IPOs?: The advent of the "7% plus contract. *Journal of Financial Economics, 59*(3), 313-346.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of accounting and economics, 31*(1), 405-440.

Heaton, J. B. (2002). Managerial optimism and corporate finance. *Financial Management*, 33-45.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication, 45*(4), 363-407.

Hensler, D. A., Rutherford, R. C., & Springer, T. M. (1997). The survival of initial public offerings in the aftermarket. *Journal of Financial Research, 20*(1), 93-110.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance, 56*(4), 1533-1597.

Hoque, H. (2011). The choice and role of lockups in IPOs: evidence from heterogeneous lockup agreements. *Financial Markets, Institutions & Instruments, 20*(5), 191-220.

Hoque, H. (2014). Role of Asymmetric Information and Moral Hazard on IPO Underpricing and Lockup. *Journal of International Financial Markets, Institutions and Money*.

Hoque, H., & Lasfer, M. (2009). IPO lockup agreements and trading by insiders. *Unpublished paper, Cass Business School*.

Hoque, H., & Lasfer, M. (2010). Insider Trading and the Long-run Performance of IPOs. *Available at SSRN 1570762*.

Hoque, H., & Lasfer, M. (2015). Why do Directors change their holdings during lockup period? *working paper, Cass Business School*.

Hoque, H., & Lasfer, M. (2016). Are IPO Lockups Binding? *Working paper, Cass Business School*.

Hughes, P. J., & Thakor, A. V. (1992). Litigation risk, intermediation, and the underpricing of initial public offerings. *Review of Financial Studies, 5*(4), 709-742.

Ibbotson, R. G. (1975). Price performance of common stock new issues. *Journal of Financial Economics, 2*(3), 235-272.

Ibbotson, R. G., & Ritter, J. R. (1995). Initial public offerings. *Handbooks in operations research and management science, 9*, 993-1016.

Ibbotson, R. G., Sindelar, J. L., & Ritter, J. R. (1994). The market's problems with the pricing of initial public offerings. *Journal of applied corporate finance, 7*(1), 66-74.

James, C. (1992). Relationship-specific assets and the pricing of underwriter services. *The Journal of Finance, 47*(5), 1865-1885.

Jenkinson, T., & Ljungqvist, A. (2001). *Going public: The theory and evidence on how companies raise equity finance*: Oxford University Press on Demand.

Jenkinson, T., & Ramadorai, T. (2013). Does one size fit all? The consequences of switching markets with different regulatory standards. *European Financial Management, 19*(5), 852-886.

Johnson, J. M., & Miller, R. E. (1988). Investment banker prestige and the underpricing of initial public offerings. *Financial Management*, 19-29.

Karcic, F. (2006). Textual analysis in islamic studies: A short historical and comparative survey. *Islamic studies, 45*(2), 191-220.

Karpoff, J. M., Lee, G., & Masulis, R. W. (2013). Contracting under asymmetric information: Evidence from lockup agreements in seasoned equity offerings. *Journal of Financial Economics, 110*(3), 607-626.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis, 33*, 171-185.

Kothari, S., Li, X., & Short, J. E. (2008). The effect of disclosures by management, analysts, and financial press on cost of capital, return volatility, and analyst forecasts: A study using content analysis.

Kothari, S., Li, X., & Short, J. E. (2009). The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review, 84*(5), 1639-1670.

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

Krishnan, C., Ivanov, V. I., Masulis, R. W., & Singh, A. K. (2011). Venture capital reputation, post-IPO performance, and corporate governance. *Journal of Financial and Quantitative Analysis, 46*(05), 1295-1333.

Lambert, R., Leuz, C., & Verrecchia, R. E. (2007). Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research, 45*(2), 385-420.

Lee, I. (1997). Do firms knowingly sell overvalued equity? *The Journal of Finance, 52*(4), 1439-1466.

Leland, H., & Pyle, D. (1977). Information Asymmetries, Financial Structure and Firm Intermediation. *Journal of Finance, 50*, 301-318.

Leone, A. J., Rock, S., & Willenborg, M. (2007). Disclosure of intended use of proceeds and underpricing in initial public offerings. *Journal of Accounting Research, 45*(1), 111-153.

Levis, M. (1993). The long-run performance of initial public offerings: The UK experience 1980-1988. *Financial Management*, 28-41.

Levis, M. (2011). The performance of private equity-backed IPOs. *Financial Management, 40*(1), 253-277.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of accounting and economics, 45*(2), 221-247.

Li, F. (2009). The determinants and information content of the forward-looking statements in corporate filings-a Naïve Bayesian machine learning approach. *Working paper, University of Michigan*.

Lin, T. H., & Smith, R. L. (1998). Insider reputation and selling decisions: the unwinding of venture capital investments during equity IPOs. *Journal of Corporate Finance, 4*(3), 241-263.

Ljungqvist, A., Marston, F., & Wilhelm, W. J. (2009). Scaling the hierarchy: How and why investment banks compete for syndicate co-management appointments. *Review of Financial Studies, 22*(10), 3977-4007.

Ljungqvist, A. P., Jenkinson, T., & Wilhelm, W. J. (2003). Global integration in primary equity markets: The role of US banks and US investors. *Review of Financial Studies, 16*(1), 63-99.

Logue, D.E., 1973. On the pricing of unseasoned equity issues: 1965–1969. Journal of Financial and Quantitative Analysis, 8(01), pp.91-103.

London Stock Exchange History. (2013). Retrieved from http://www.londonstockexchange.com/about-the-exchange/company-overview/our-history/our-history.htm

Loughran, T., & McDonald, B. (2010). Measuring readability in financial text. *SSRN eLibrary*.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35-65.

Loughran, T., & McDonald, B. (2013). IPO first-day returns, offer price revisions, volatility, and form S-1 language. *Journal of Financial Economics, 109*(2), 307-326.

Loughran, T., & McDonald, B. (2014). Regulation and financial disclosure: The impact of plain English. *Journal of Regulatory Economics, 45*(1), 94-113.

Loughran, T., & McDonald, B. (2015). The use of word lists in textual analysis. *Journal of Behavioral Finance, 16*(1), 1-11.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *The Journal of Finance, 50*(1), 23-51.

Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of market efficiency. *Journal of Financial Economics, 55*(3), 361-389.

Loughran, T., & Ritter, J. R. (2002). Why don't issuers get upset about leaving money on the table in IPOs? *Review of Financial Studies, 15*(2), 413-444.

Loughran, T., & Ritter, J. R. (2003). Why has IPO underpricing changed over time?

McKee, A. (2003). *Textual analysis : a beginner's guide*. London: SAGE.

Megginson, W. L., & Weiss, K. A. (1991). Venture capitalist certification in initial public offerings. *The Journal of Finance, 46*(3), 879-903.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of Finance, 32*(4), 1151-1168.

Miller, R. E., & Reilly, F. K. (1987). An examination of mispricing, returns, and uncertainty for initial public offerings. *Financial Management*, 33-38.

Morris, S. (1996). Speculative investor behavior and learning. *The Quarterly Journal of Economics*, 1111-1133.

Ofek, E., & Richardson, M. (2000). DotCom mania: the rise and fall of Internet stock prices. *Journal of Finance, 58*.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*(1), 435-480.

A Plain English Handbook: How to create clear SEC disclosure documents. (1998). Retrieved from https://www.**sec**.gov/pdf/**handbook**.pdf.

Prabhala, N., & Puri, M. (1998). How does underwriter price support a! ect IPOs?: empirical evidence. *Unpublished Working Paper, University of Maryland, College Park*.

Ritter, J. (2016). IPO data. Retrieved from https://site.warrington.ufl.edu/ritter/ipo-data/

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

Ritter, J. R. (1987). The costs of going public. *Journal of Financial Economics, 19*(2), 269-281.

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance, 46*(1), 3-27.

Ritter, J. R. (2003). Differences between European and American IPO markets. *European Financial Management, 9*(4), 421-434.

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

Robins, N. (2012). *The Corporation That Changed the World: How the East India Company Shaped the Modern Multinational* (2nd Edition ed.): Pluto Press.

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

Schultz, P. (1993). Unit initial public offerings: A form of staged financing. *Journal of Financial Economics, 34*(2), 199-229.

Schultz, P. (2003). Pseudo market timing and the long-run underperformance of IPOs. *The Journal of Finance, 58*(2), 483-518.

Seguin, P. J., & Smoller, M. M. (1997). Share price and mortality: An empirical evaluation of newly listed Nasdaq stocks. *Journal of Financial Economics, 45*(3), 333-363.

Seyhun, H. N. (1986). Insiders' profits, costs of trading, and market efficiency. *Journal of Financial Economics, 16*(2), 189-212.

Seyhun, H. N., & Bradley, M. (1997). Corporate Bankruptcy and Insider Trading*. *The Journal of Business, 70*(2), 189-216.

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. *Journal of Accounting Research, 32*(1), 38-60.

Skinner, D. J. (1997). Earnings disclosures and stockholder lawsuits. *Journal of accounting and economics, 23*(3), 249-282.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance, 62*(3), 1139-1168.

Tetlock, P. C., SAAR-TSECHANSKY, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance, 63*(3), 1437-1467.

text - definition of text in English from the Oxford dictionary. (2015). Retrieved from http://www.oxforddictionaries.com/definition/english/text

Tinic, S. M. (1988). Anatomy of initial public offerings of common stock. *The Journal of Finance, 43*(4), 789-822.

Torstila, S. (2001). What determines IPO gross spreads in Europe? *European Financial Management, 7*(4), 523-541.

Torstila, S. (2003). The clustering of IPO gross spreads: International evidence. *Journal of Financial and Quantitative Analysis, 38*(03), 673-694.

Vismara, S., Paleari, S., & Ritter, J. R. (2012). Europe's second markets for small companies. *European Financial Management, 18*(3), 352-388.

Weber, J., & Willenborg, M. (2003). Do expert informational intermediaries add value? Evidence from auditors in microcap initial public offerings. *Journal of Accounting Research, 41*(4), 681-720.

Welch, I. (1992). Sequential sales, learning, and cascades. *The Journal of Finance, 47*(2), 695-732.

Welch, I. (1996). Equity offerings following the IPO theory and evidence. *Journal of Corporate Finance, 2*(3), 227-259.

Willenborg, M., & McKeown, J. C. (2000). Going-concern initial public offerings. *Journal of accounting and economics, 30*(3), 279-313.

World Fedration of Exchanges. (2012). *Monthly reports*. Retrieved from Paris:

Yung, C., & Zender, J. F. (2010). Moral hazard, asymmetric information and IPO lockups. *Journal of Corporate Finance, 16*(3), 320-332.