Which IPO Conditions Affect a Firm's Ownership Structure?

MSc Economics and Business: Financial Economics

Erasmus University Rotterdam

Supervisor: Mrs. J. Zhao

Floris van Rijn, 432566

October , 2018

Abstract

This paper examines the change in insider ownership as well as an alternative measure of ownership dispersion of IPO firms from 1999 to 2015 to understand which factors affect affect U.S. IPO firms' ownership structure the most. Further developing a new measure for a popularity variable through the use of Google Trends data. Robustness checks come in the form of a Durbin-Watson test, as well as exploring an alternate measure of the dependent variable. The findings from the % change in insider ownership regression show that VC backing, R&D spending and control variables for year and industry are significant in explaining the change in ownership during an IPO. The alternate measure of dispersion (holdings of largest 5 owners) finds market-to-book, free cash flow and VC dummy to be significant in explaining changes in dispersion. Lastly, this paper develops a python script to allow researchers to gather information in large quantities from Google Trends.

Table of Contents

1 Introduction

The Modern Corporation and Private Property, released in 1932 by Berle and Means was a book that called attention to the prevalence of widely held corporations in the United States. They showed that though capital was dispersed among small shareholders, control was concentrated in the hands of the managers. Their book formed the foundation of research for decades to come and introduced a focus on agency costs and governance within the research field. Agency costs are inherent in any situation that involves cooperative effort by two or more people. In the case of firm ownership agency costs arise due to the decision making capabilities of the agent that are entrusted to them by the principal. If both the agent and the principal are utility maximizing there is reason to believe that the agent will not always act in the best interest of the principal. This is especially relevant when control is concentrated in the managers whilst capital is dispersed. The principal aims for welfare maximization and the agent can direct the company towards a number of other goals such as revenue maximization. To combat this, costs are incurred by the principal for monitoring as well as bonding the agent. A further issue from dispersed ownership is shirking. Shirking refers to the ability of the agent to use his time and energy on other tasks where the benefit is accrued entirely by him. The cost of shirking is in the form of poorer performance of the firm and is bared by all the shareowners. Dispersed ownership is possible and beneficial when, as La Porta, Lopez-de-Silanes, and Shleifer (1999), (hereafter LLS) argue *"there is better legal protection for investors".* Jensen and Meckling (1976) predicted that ownership would be more concentrated when controlling insiders found it easier to take advantage of outside shareholders and stronger legal protection for outsiders lays the foundation for dispersed ownership. Berle and Means' model led to an intense debate within academia that focused on whether firms with managers that own more shares performed better and had higher valuations than those that did not.

In recent times, research has focused more on challenging the empirical validity of this model. Demsetz (1983) and Shleifer and Vishny (1986) show that even the largest of U.S. firms have a very modest concentration of ownership. Wide dispersion in ownership is a finding more prevalent for the U.S. and U.K. than other markets (due in part to its better protection of investors); more significant concentration of ownership is found in Germany, Japan, Italy and

seven other OECD countries than for the U.S. Recent research contradicts the Berle and Means image and shows that managers are in fact accountable. LLS largely do away with the Berle and Means picture of ownership structure for the modern corporation in countries other than the richest common law countries. They cast a shadow of doubt on the findings of Roe (1994) who attribute ownership dispersion in the United States to specific policies that discourage ownership concentration; instead their findings suggest that ownership structures are primarily an equilibrium response to the domestic legal environments that companies operate in. A differing opinion is found by Bolton and von Thadden (1998) who attribute the dispersion to the greater liquidity of markets in the United States and the United Kingdom compared to most countries. A large body of research has focused on understanding ownership structure over time, whereas no studies have taken into focus the factors that affect ownership structure at the time of the IPO.

The majority of firms begin their life with a high amount of insider ownership. This is defined by LLS hereafter as blockholders owning more than 20% of the total votes. LLS find on average that 75% of the companies that conduct their IPO are majority owned at the time. Privately held firms can be enticing due to their lack of disclosure requirements compared to publicly listed companies. Publicly listed firms are subject to detailed disclosure laws including: information about their financial condition, operating results and management compensation. In the U.S. the Securities and Exchange Commission (SEC) require publicly owned companies to disclose relevant information when new stocks are issued, as well as to disclose the financial and business data on a regular basis. Publicly owned companies prepare two annual reports that are subject to strict rules of disclosure and formatting. These act as a monitoring tool that promote greater dispersion of ownership. Privately held firms are able to take advantage of information asymmetries as they are not subject to strict disclosure requirements. Over a relatively short amount of time after the IPO a variety of factors including stock market performance and market liquidity lead to a significant drop in the insider ownership. To understand the dispersion in the U.S. market, it is important to study the change in ownership structure at the moment a company becomes listed.

Factors that lead to a wider ownership of a company years after the initial public offering have been researched in depth by Helwege, Pirinsky and Stulz (2007). Furthermore, research on "Why companies become more dispersed over time", "How do they become dispersed" and "What is the performance difference between dispersed and non-dispersed companies" have

been studied in depth by the likes of Coles, Lemmon, Meschke (2003), Demsetz, Villalonga (2001) and Himmelberg, Hubbard, and Palia (1999). These papers focus on understanding if ownership structure affects performance as well as why insiders own a larger fraction of shares in one firm but not in another. The reason behind the number of shares issued at IPO has not been subject to studies before and due to this, very little is known about the causes of ownership structure changes immediately following an IPO. This paper attempts to answer the question "Which IPO Conditions Affect a Firm's Ownership Structure?". Investigating the dynamics of ownership structure at the time of the firm's IPO in the United States would provide a deeper understanding of ownership structure patterns as well as provide practical information useful for the market.

This paper will conduct an empirical study that will find the factors that affect ownership changes at the time of the IPO and the extent each factor contributes. This paper's data follows all firms that have completed an IPO from 1999 onwards; this timeframe allows us to grasp a strong understanding of how firm ownership structures have changed as well as which factors affect ownership structure and to what extent. The start year of the data sample is chosen at 1999 to focus on the explanatory power of variables in more recent times. Taking data from 1999 through till 2013 gives us fourteen years of active data on which to base our regression.

This paper's methodology is similar to that of Helwege, Pirinsky, and Stulz (2007) and regresses a large number of factors on the dependent variable (insider ownership). The contribution of each factor can be easily measured as the dependent variable is measured as a change in percentage. This means the coefficient of the independent variables is equal to the % change in insider ownership if the value of the independent variable is 1. The robustness check comes in the form of a Durbin-Watson test.

An example of the benefit of the outcome of this research for investors is as follows: Brealey, Leland and Pyle (1977) model a retention of shares by insiders as a signal of the firm quality when information asymmetries are high. Dispersed ownership at IPO can signal that the informational advantage of insiders has become less important and for given industries/companies this information can be critical for the future of the company. An advantage of this research for future IPOs lies in the factors of the regression studied. In understanding what factors have caused the largest changes in ownership structure in peers, a company is able to determine what ownership structure change it can actively pursue. The

outcome of this research provides future IPOs and investors with key knowledge in understanding what has shifted the ownership structure. Further, the information gained through this study benefits the market by providing information on signals the company going through their IPO is sending out. If the factor "Google Trends" is deemed to have the greatest effect on change in ownership, then a company preparing for its own IPO can calculate from its own trending on Google how likely a proposed change in ownership structure is to be accepted by the market (i.e. its shares are bought). Taking the previous Brealey, Leland and Pyle (1977) example, an investor is able to study the factors that caused the largest changes in ownership structure and deduce whether the information advantage of insiders has become less important or that the IPO was taking advantage of its trending factor to raise the maximum amount of funds. Before Helwege, Pirinsky and Stulz (2007), factors that affected ownership had seen very little research despite being significant. This is important as the factors can signal information about the company for investors. In their research they argued that to differing extents the factors: moral hazard (controlling insiders find it easier to take advantage of outside/minority shareholders), adverse selection (insiders sell shares at an advantageous price and adverse selection makes it expensive to sell shares when information asymmetries are high) and timing (insiders time the sale of their own shares) played a key role in the change in ownership over time. Knowing the contribution of each of these factors provides information to the market that is of interest. This paper adds to this area of research by investigating which factors affect the change in ownership in the initial public offering as well as a brand new focus on a 'trending' factor. The research will provide key information that relates to the future performance of a firm after its IPO and allow investors to be more informed as well as furthering understanding of ownership structure for companies in the U.S. In relation to existing literature there has been no study that focuses on the change in ownership structure at the time of the IPO. Limited literature has looked into how stocks become dispersed over time (Helwege, Pirinsky and Stulz (2007)) and other research has focused on uncovering the ownership traits of companies around the world (La Porta, Lopez-de-Silanes and Shleifer (1999)). With the development of an empirical model, future IPOs can be scrutinized to a larger degree and more information contained in the level of ownership change can be understood.

The rest of this paper will proceed as follows. Section II focuses on the literature review outlining the hypotheses of this paper. Section III explains the methodology employed. Section IV focuses on the data and descriptive statistics. Section V explains the results and lastly section VI concludes.

2 Literature review

Ownership structure is affected through a variety of different variables. Research over the past eight decades has taken the understanding of these variables in different directions and often clashed over the significance of them. This section investigates the history of the research, the main debates in the literature as well as establishing the relevance of the variables included in the regression.

2.1 History of ownership structure

Ownership structure was first popularized within research in the 1932 publication from Adolph Berle and Gariner Means. Their book showed that widely held corporations in the United States, where ownership of capital was dispersed among shareholders, still had concentrated ownership of control in the hands of managers. They painted an image of managers that were held unaccountable for the decisions they made as the shareholders didn't have the power to act. This model was further adopted in important book publications such as that of Baumol (1959) and Marris (1964). In the same year Williamson published *"The economics of discretionary behavior: Managerial objectives in a theory of the firm"* which showed the manager of a firm not to be acting in the interest of profit maximizing but in utility maximizing as control of the firm was in the hands of the managers further perpetuating the view of Berle-Means. This spurred an intense controversy over whether firms in which managers own more shares perform better and have higher valuations. This view continued with the publication of Jensen and Meckling's (1976) paper that focused on developing a theory on the ownership structure of the firm, as well as Grossman and Hart's (1980) paper. The Berle and Means publication led to a large literature on agency costs and governance. The model raised concerns with diffuse ownership as it believed that it facilitated entrenchment by managers. More recently however, several studies began to question the empirical validity of this model. Demsetz and Lehn (1985) cast doubt on the Berle-Means model as they found no significant relationship between ownership concentration and accounting profit. These results are further validated by Himmelberg, Hubbard and Palia (1999) who build on the Demsetz and Lehns (1985) study and conclude with the same findings. Further papers by Shleifer and Vishny (1986), Morck, Shleifer and Vishny (1988) and Holderness and Sheehan (1988) practically do away with the Berle-Means model.

2.2 Causes of ownership dispersion in the U.S.

The study of ownership structure takes a different direction around this time, with La Porta, Lopez-de-Silanes and Shleifer's (1999) paper focusing on ownership dispersion across the world and what causes the differences between countries. Their findings showed that the level of dispersion in the United States could be attributed to the legal protection of minority shareholders. Roe (2003) argue that the dispersion is due to the differences in politics and Bolton and von Thadden (1998) argue in favour of greater market liquidity and ease of takeovers in the US and UK. The opinions on the matter differed between researchers and became a topic of debate. These conflicting findings formed the base for the paper of Helwege, Pirinsky and Stulz (2007). They argued that previous research sought to determine why, at a point in time, insiders owned a larger fraction of the shares in some firms but not others and that the evolution of ownership as firms mature had not been the subject of study in the United States.

In studying the evolution of ownership, they believed to be able to provide a better understanding of why companies in the United States have a larger dispersion in comparison to other countries. With this goal in mind, they investigated key variables that fall within the explanations of moral hazard, adverse selection and timing.

2.2.1 Moral Hazard

Moral hazard refers to the agency theory of Jensen and Meckling (1976). Their research shifted the classical model of the firm to one that is comprised as a set of contracts among factors of production. Their research predicted that controlling insiders would have a higher concentration of ownership if they find it easier to take advantage of minority shareholders. Taking advantage of shareholders can come in the form of shirking, investing in pet-projects and tunneling funds. The costs of these acts are bared by all the shareholders whilst the benefit is only gained by the controlling insiders. Higher inside ownership means the actions of controlling insiders become better aligned with the interests of the minority shareholders as they bare more of the cost of their actions. If the agency view is correct then firms are expected to become dispersed if there are either new ways of managing the agency problems or if these agency problems become less important.

2.2.3 Adverse Selection

Adverse selection takes on the model of Leland and Pyle (1977). Retention of shares by insiders is seen as a signal of firm quality. This is stringent upon the idea that information asymmetries are high, as is often the case with firms before their IPO. This is due to the relative lack of coverage and monitoring of the company in comparison with public firms. Due to the information asymmetries, insiders maximize the wealth of existing shareholders when they choose to sell equity at an advantageous price; making the sale of equity costly when information asymmetry is high. Subrahmanyam and Titman (1999) show that the advantage of dispersed ownership comes about when the value of insider knowledge becomes less important.

2.2.4 Timing

Timing the market for a firm's shares is the norm. Insiders decide when to sell the shares for their own account and can therefore pick the most advantageous timing. Clarke et al (2004) shows that on secondary equity issues, insiders gain from timing the sale of their shares. Selling shares or issuing equity when the value of a firm's shares are high allows a firm to utilize this price to their benefit such as in the takeover of other companies. This effect is known as the window of opportunity hypothesis. The market for a firm's shares becoming more liquid allows for the larger sale of blocks of shares without swaying the price significantly. This means that ownership can become more dispersed when the liquidity for a firm's shares is higher

2.2.5 Findings

The results of Helwege, Pirinsky and Stulz (2007) showed that insider ownership falls steadily over time from the time of the IPO. Half of the sample firms had less than 20% insider ownership just 10 years after the IPO. Their results further showed that dispersion in the U.S. was driven by the following factors: insider ownership, venture capital, firm size, hard assets, market to book ratio, funding needs, capital expenditures, R&D spending, dividends, leverage, and volatility.

2.3 Measures of ownership dispersion

To fully understand the effects of the independent variables on changes in ownership structure, we must establish that the measure of the dependent variable is the best one possible. The initial measure adopted in this study is the percentage change in insider ownership measured at the time of and the day after the IPO. This variable is used in the study of Helwege (2007) to study the effects of ownership dispersion over time and as an extension, adopted in this paper's study of variables affecting ownership at the time of the IPO. Though a valid measurement of ownership structure, it does not provide much information about the dispersion of ownership after the IPO, only by how much holdings of insiders have changed. Overland, Mavruk and Sjögren (2012) study ownership concentration measure and compare twenty measures used in published journals; finding that measures have strong correlation, but vastly different underlying distributions for each concentration measure. The effect of this is that papers can have conflicting results. Morck et al. (2000) and Thomsen and Pedersen (2000) find a significant positive relationship for the performance of a firm and ownership concentration, whereas Lehmann and Weigand (2000) and Leech and Leahy (1991) find ownership concentration negatively related to the performance of a firm. These differing results come from a varied measure of ownership concentration adopted by these papers. Overland, Mavruk and Sjögren (2012) find that these differences are explained by three properties. These properties are: contextual settings, data quality and methodology. Ultimately methodology is the driver of these conflicting results. Different ownership concentration measures are distributed differently to each other and therefore the results differ when substituting measures for each other. This finding is supported by Edwards and Weichenrieder (2009).

The outcome of Overland et al's (2012) paper shows that any measure used to measure ownership concentration has to be grounded in the scope of the research, meaning that the measurement of the variable is well defined and related to the study. Moreover, they find that measures such as the Banzhaf index provides a strong measurement for ownership concentration, but also find support that a measure such as the Shapley-Shubik can be swapped out for a more accessible measure such as largest shareholder. In using largest shareholding however, no distinction is made between the ownership concentration of the rest of the shareholders. It is reasonable to assume, as Overland et al (2012) state given that there is one large shareholder already, that voting rights divided among many separate shareholders are not treated as equally as the presence of another large shareholder. A company with two large

shareholders is more concentrated than a company with one large shareholder and a plethora of smaller shareholders. The measure of largest shareholder does not take this into account, therefore this paper resolves this by testing as the second regression the change in holdings from the five largest owners of a company. This provides a simple enough measure for which data is present, whilst being a good measure of ownership concentration. The addition of this dependent variable run as a second regression allows a more detailed understanding of the factors that don't just affect insider holdings, but ownership concentration in general. This provides a more detailed picture of the change in ownership at the time of the IPO.

2.4 Ownership structure at the time of IPO

Leland and Pyle (1977) investigate information asymmetry in markets and the role of equity retention. Informational asymmetries are particularly pronounced in financial markets. Borrowers know their collateral and have information pertaining to their project that only they know. This inside information is unlikely to be shared with the lender if the information does not benefit the borrower. To investigate or verify information provided by borrowers can be a costly operation and therefore does not offer a solution to the asymmetry. This means the market value reflects the average quality of projects. Higher average quality can lead to borrowers taking advantage of the uninformed market. By retaining very little or no equity, they exit their project and make a profit. Leland and Pyle (1977) show that the willingness of insiders to invest in the project or firm helps combat the asymmetry. This action signals to the market the quality of the project.

Information asymmetries are inherent to the IPO market. Lack of coverage of firms as well as the cost of monitoring mean that the insiders must find a different way to signal the quality of their firm to the market. The model of Leland and Pyle (1977) is one way in which information is broadcast through the number of shares retained by insiders. Downes and Heinkel (1982) further investigate the effectiveness of this signaling as well as investigating the effect of dividend policy on signaling information. Their results show that for unseasoned new equity issues ownership retention is positively related to firm value. However, dividend policy is not.

Helwege, Pirinsky and Stulz (2007) pioneered an understanding of the factors that caused dispersion over time and what their individual contribution was, however, their research only focused on understanding the dispersion from a year after the IPO of a company. Significant

information is contained in both the retention of insider holdings and dispersion of ownership by a company through an IPO as found by Leland and Pyle (1977) and a better understanding of dispersion and ownership structure can be obtained from the study of this area.

2.4 Hypotheses

A number of factors play a crucial role in the decision-making process regarding the number of shares to issue during an IPO. These factors, mentioned above, include: insider ownership, venture capital, firm size, hard assets, market to book ratio, funding needs, capital expenditures, R&D spending, dividends, leverage, volatility, and level of trending. Most of these factors were used in Helwege, Pirinsky and Stulz (2007) with exception to the trending factor, and were used to identify why large decreases in insider ownership took place over many years. As this paper also investigates a change in ownership structure of companies but at a different time, we can hypothesize that these factors will have an effect in the same direction as found by Helwege, Pirinsky and Stulz but with differing magnitudes. The inclusion of each of the factors is explained, and underneath this paper presents the hypothesis that will be tested in relation to each variable.

2.4.1 Venture Capital

Helwege, Pirinsky and Stulz (2007) found venture capital (VC) backing of a company to result in a company experiencing substantial decreases in insider ownership relative to those that were non-VC funded. This paper believes the effect of VC funding to be much the same in that it will reduce insider ownership, but the magnitude to be different. The reason for this is the IPO marks as a substantial exit strategy for a VC (Black and Gilson (1998)) and cause a significant change in ownership structure; however, VC funding years after an IPO will still alter ownership structure, but not to the same extent as during an IPO.

H1 i) VC backing will result in higher dispersion of ownership after the IPO than for non-VC backed companies.

2.4.2 Firm size

Firms that are larger at the time of their IPO will have fewer information asymmetries as they are monitored more by the market. It is well documented that larger firms are better monitored by institutional shareholders, analysts and the press. Using as the proxy for firm size the

logarithm of total assets this paper expects larger firms to have a larger change in ownership structure at the time of the IPO, making the coefficient of firm size positive.

H1 ii) Larger firms are more likely to have a higher level of dispersed ownership after the IPO than smaller firms.

2.4.3 Market to Book ratio

Firms with lower market to book ratios are ones with more intangibles on the balance sheet. A low ratio can however also indicate that the firm has high growth opportunities. This in turn means more information asymmetries as there is more discretion on the part of insiders. If this is the case, then higher insider holdings are likely. However, Stulz (1990) shows that this could also be different if the insiders are empire-builders, meaning higher market to book ratio would lead to lower insider holdings. This paper's hypothesis posits that of IPOs it is more likely that companies look to be diffusely held when market to book ratios are high.

H1 iii) Firms with a higher market to book ratio are more likely to have a higher dispersed ownership after the IPO.

2.4.4 Tangible Assets

Using the ratio of property, plant, and equipment to total assets as a measure for tangible assets this paper hypothesizes that firms with more tangible assets are less able to be discretionary with their spending. This means that firms with higher ratios for tangible assets are more likely to be held diffusely.

H1 iv) Firms with more tangible assets will experience larger amounts of dispersed ownership after the IPO than firms with less tangible assets.

2.4.5 Free Cash Flow

A firm that has less operating cash flow but higher expenditures has a greater need for funding. Firms in need of funding can expect insider ownership to drop. Using EBITDA as a measure of cash flow this paper hypothesizes that firms with greater free cash flow have a higher optimal

insider ownership level than those with lower free cash flow. Higher EBITDA means higher free cash flow.

H1 v) Firms with greater free cash flow will have less dispersed ownership after the IPO.

2.4.6 Research and Development

R&D spending inherently comes with greater managerial discretion and greater information asymmetries. What exactly the funds are being spent on is usually not known by the public. Firms that have higher R&D spending have higher optimal insider ownership according to Himmelberg (1999). Using R&D spending divided by total assets as a more indicative measure for R&D spending, this paper hypothesizes a higher level of this factor to lead to higher insider ownership and therefore a negative coefficient.

H1 vi) Firms with higher levels of R&D spending are less dispersed after their IPO.

2.4.7 Dividends

Dividend payments result in lower agency problems as less resources are left to the discretion of management. Dividend payments from newly offered firms are rare, however, if included dividend paying firms are likely to have lower optimal insider holdings making the coefficient negative.

H1 vii) Firms that pay dividends will be more dispersed after their IPO.

2.4.8 Leverage

As studied by Jensen (1986) and Stulz (1990) leverage reduces agency problems between outside shareholders and inside owners. Higher levered firms will therefore aim to be more diffusely held. Highly levered firms are also more likely to issue equity and dilute insider ownership leading to a larger amount of shares issued and a positive coefficient. An argument can also be made that insiders that wish for a larger proportion of control over the firm will choose higher leverage to be able to afford a large proportion of the firm's cash flow which would lead to a negative coefficient, however this is less likely to occur in practice.

H1 viii) Firms that have higher leverage will be more dispersed after their IPO.

2.4.9 Volatility

High volatility causes a reduction in expected utility of insiders and is likely to increase the number of shares offered. However, volatility could have the opposite effect as information asymmetries are more important. This paper uses the percentage difference in price of the IPO's stock up to a year after offering, measuring from the high to the low models for the volatility in the market at the time.

H1 ix) Higher values of volatility in the market will have a near zero impact on dispersion.

2.4.10 Trending factor

This paper includes a new variable called "trending", not utilized in the study of ownership structure before. This variable is based on data from Google Trends; which shows how often a particular term is searched relative to the total volume across regions, languages and time. Entering the search term provides a graph that allows the user to see at which points the search-term were the most popular. The image below shows this data for Facebook.

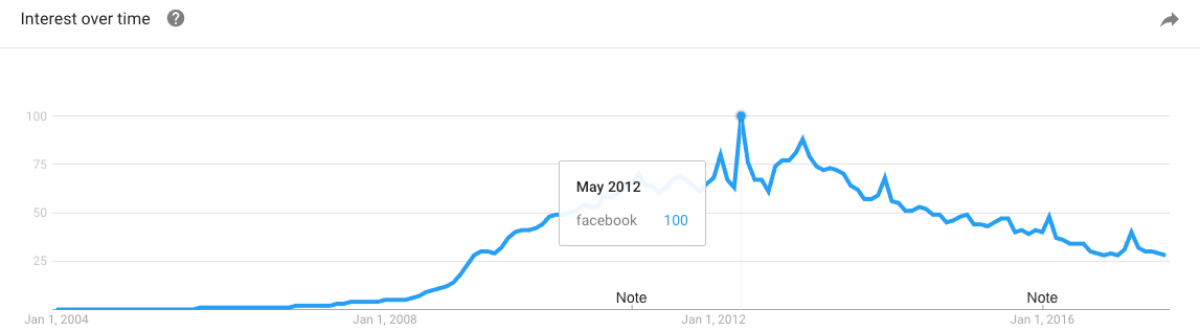


Figure 1 Google Trends search for Facebook under Google Finance

The graph shows that the term "Facebook" searched under the Finance category of Google peaked in May 2012. Facebook held their IPO on the 18th of May 2012. The inclusion of this term is important in understanding how companies decide to issue the number of shares during their IPO as increased searching combats the information asymmetries experienced by a company before its IPO. Over time, larger quantities of data are being collected by companies,

the use of which is applied commercially. This variable aims to use the data available in an academic setting and make larger quantities of data available.

Google's search engine has 1.17 billion frequent users (statista.com) and its search data provides an accurate representation of interest in a topic by the market. This data can be filtered to include only the most relevant categories of searches such as Finance or News rather than every search for the term "Facebook" done at the time; giving an accurate representation of potential market participants in the trending factor. Keim and Madhavan (1996) show that when the market for a firm's shares becomes more liquid, the demand for shares becomes more elastic so that it is less costly for insiders to sell shares in the market. They further show that the cost of selling a block of shares for a small firm can be very large. This cost is relevant at the time of the IPO and the inclusion of this factor will allow a more detailed understanding for the reasons behind the change in insider ownership. Greater levels of trending before the IPO will result in higher diffusion of ownership; this is because a higher level of trending indicates higher interest in a firm by market participants and proxies for the demand for the firm's shares. This means a firm is able to decrease inside ownership at relatively little cost.

H1 x) The more popular a firm is before IPO, the more ownership will be dispersed after IPO

3 Methodology

3.1 Regression type

This paper focuses on a single multivariable regression that models the relationship of independent variables on the dependent variable. Further focusing on a number of key variables for different companies at the time of their IPOs as well as including controls for industry and year of IPO through dummy variables. The outcome of the regression will show the effect that an increase in the independent variable will have on the dependent variable. An example of this is that a 1% increase in the leverage held by the firm could result in a 0.5% increase in the change of insider holdings. To check the robustness of the results, a test to determine the presence of autocorrelation is run. In doing this the multivariable regression will show the marginal effect of each independent variable on the dependent variable. Furthermore, an alternate measure for the level of dispersion in insider ownership is explored.

3.2 Regression

3.2.1 Variables

The multivariable regression uses a number of independent variables regressed onto the dependent variable. The study focuses on which IPO conditions affect a firm's ownership structure so the dependent variable is equal to the change in insider ownership from before and after the IPO measured as a percentage. The value of dependent variable is changed to an alternate measure of dispersion in the form of the change in percentage of ownership of the largest 5 owners, and the regression is rerun; the reason for this is discussed later. The data for the dependent variable is taken from Thomson Reuters and shows that many companies enter the IPO at insider ownership substantially less than 100%. This is paired with data from Datastream for the independent variables. The data collected represents various balance sheet items for each company at the time of their IPO. Importing the data into the statistical program SPSS allows for the building of the multivariate regression.

3.2.2 Correlation

An issue that can arise from the use of these variables is strong correlation between a select few of them. For example, the research and development variable can be closely related to the tangible assets variable as increases in R&D could also be increases in tangible assets. By having a relatively large sample size any correlation between variables is likely to be an accurate representation. An identifying mark of correlation between variables is a large standard error. If this is the case, then the regression is unable to attribute the effects of the independent variables on the dependent variable as they are also correlated between themselves. To combat this the regression can be rerun with one of the variables omitted. This will likely result in a standard error closer to the order of magnitude of the coefficient of the independent variable and solve the issue.

A larger issue in the sample could be the presence of autocorrelation. This paper uses a Durbin-Watson test to test for autocorrelation in the sample. If autocorrelation is present, then the error terms do not fully represent the values for the population. This can be solved through the use of the Generalized Least Squares (GLS) method.

Name of variable	Definition	Measurement
Change in insider ownership (INSD)	INSD is the change in % ownership by insiders before and after the IPO.	the change in the percentage of insider ownership measured before the IPO and again a month after going public. The difference between these values equals the change in insider ownership
Change in size Largest 5 owners (5owner)	The change in shares held by the largest 5 owners of a company	Percentage change of shares held by the largest 5 owners measured the month of the IPO and three months later
Venture Capital (VC)	IPO is backed by venture capital	Dummy variable equal to 1 if the company has received VC funding before IPO.
Firm size (FSIZE)	Size of the firm	Natural log function of the book value of assets, originally measured in dollars
Tangible Assets (TANG)	Physical items with clear value	Natural log function of property + plant + equipment all divided by the total assets. Both tangible assets and total assets are measured in dollars
Market to book equity ratio (MBE)		Market value divided by the book value of company
Free cash flow (FCF)	Cash generated after spending to maintain/expand asset base	EBITDA over sales
R&D (RD)	Cash spent on research and development as a fraction of assets	Natural log function of Research and Development
Dividend (DIV)		Yearly dividend payout in dollars
Leverage (LVRG)	Amount of debt used to buy assets	Total liabilities over book value of assets measured in dollars
Volatility (VOLATILITY)	Dispersion of returns for a market index	Measure of a stock's average annual price movement to a high and low from a mean price for each year.
Trending (TRND)	Amount a term is searched through Google	Trending values are normalized and ranked out of 100. Difference between rankings at the time of IPO and two months prior are taken

3.3.3 Final equation

By regressing all mentioned independent variables on change in insider ownership this paper will be able to list the factors that explain the percentage change in insider ownership of the company and to what extent each contributes. The second regression run allows a more detailed study of the dispersion after the IPO. The two equations will be as follows:

$$INSD\ or\ largest\ 5\ owners = \alpha + \beta_1 VC + \beta_2 LnFSIZE + \beta_3 TANG + \beta_4 BME + \beta_5 FCF + \beta_6 LnRD + \beta_7 DIVIND + \beta_8 LVRG + \beta_9 \Delta VOLATILITY + \beta_{10} \Delta TRNDING + \beta_{11}\ DummyYear + \beta_{12} DummyIndusry + error\ term$$

The majority of the independent variables are chosen due to their proven relationship in previous research. Helwege, Pirinsky and Stulz (2007) use all but the trending factor in their paper to explain why companies become more dispersed over time. Due to this it can be assured that these variables by extension play an important role in dispersion at the time of IPO. As a further measure, the R squared contribution of each variable will be checked to see if a variable is helping explain a portion of the dependent variable that none of the other variables are. Through these measures we can be sure that the results obtained show a causal effect on the number of shares issued.

3.4 Cleaning up regression

After initially running the regression, it is clear that multicollinearity is present. Investigating this issue shows that the inclusion of tangible assets as well as the log function of assets is the root cause. Opting to drop the tangible assets variable and re-running the regressions without this variable.

4.0 Data and descriptive statistics

The data used in this paper comes from a variety of sources. Starting with Thomson Reuters under equity and capital markets, a filter is used to find the number of companies between 1999 and 2015 that have had an IPO in the United States. The data filtered for time period and region is important in the use of this paper. From this Thomson Reuters returns 7016 companies.

Thomson Reuters is then further used to find all available data for insider ownership of the companies in the initial list. These variables cover insider ownership before and after the IPO. The description of insider ownership used by Thomson Reuters is the same as the one adopted in this paper; meaning no transformation of the data is necessary. This eliminates a number of companies from the list and a smaller sample of 3300 different companies remain. The data of the largest five owners is also collected through Thomson Reuters using the database through WRDS allows for the gathering of data on the entire list of companies. Using the ISIN identifier codes from Thomson Reuters, the data of companies can be entered into Datastream to gather the data relevant to all but one of the independent variables. The relevant data for the last independent variable, TrendingData, is discussed further on. At this point the sample size is further reduced in size, giving a final sample size of 768 firms for the main variables in the regression.

This paper develops a script to efficiently utilise an existing program, allowing researchers to gather data from Google Trends in much larger quantities than possible before.

4.1 Google trends data

Within academic research the addition of Google Trends data as a variable in studies is very new. Papers discussed earlier have incorporated the data to a limited extent and the importance of the mass data available is slowly being recognized.
This paper, during the course of its investigation, encountered problems that hindered the availability of this data. This section will highlight the issues and how they are resolved before explaining why only a certain amount of data was used to explore the power of the variable.

4.2.1 API Query

The trending data is taken from the database available at www.Google.com/trends. To extract data from a database a user must make a request for the specific data they require in the form of a query. The database returns data that matches the specifics of the request such as data between a certain time period.
The first issue researchers will run into in the use of Google Trends as a variable in their research is that there is no Google Trends API. Whether this is done purposefully by Google or whether they are working actively on a version is difficult to determine as no active information is public about this, meaning that building a trending variable holds a lot of value

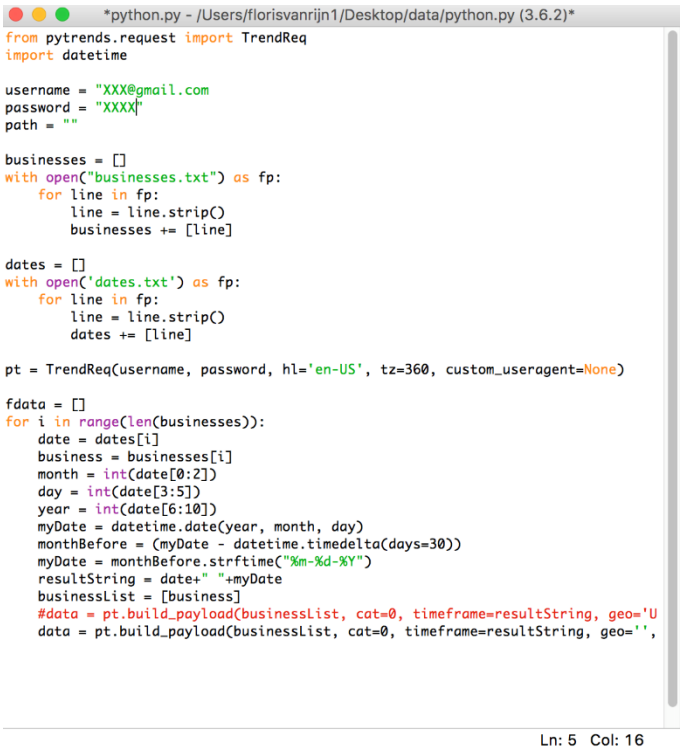
for future research. Having no API currently means a user is unable to query for a range of data and cannot collect data in large quantities. This paper develops a python script that circumvents this issue utilizing the code from an open-source project. Figure 2 below shows an image of the script, which is accompanied by an explanation of its use.

4.2.2 Requirements

Before executing the script, the user must make sure to download and have access to Python (version 2.7), as well as installing the program 'Pytrends' through the command "pip install pytrends" in the command prompt of their computer. Pytrends provides the code that goes through the iterations of querying Google Trends for the data and is an open source collaboration between a number of individuals. Pytrends automates the gathering of the raw data from Google Trends. The script below provides researchers with a format that allows them copy and paste data into two documents and receive the necessary data from Google in return in a workable format.

4.2.3 How does the script work?

Figure 2. Python script



Following along the lines, the role of each line of code is explained. Python makes sure to import the right modules from the already downloaded pytrends. After this the user replaces the |XXX" in lines 4 and 5 with their valid gmail account username and password. The program must use a Google account to access the data.

The following two sub-sections called "businesses" and "dates" refer with their command lines "with open ('dates.txt') as fp:" to two documents the researcher must fill in. In one called businesses a .txt file of all the companies relevant for the query are saved as individual line items. In the file saved as "dates" all the corresponding IPO dates of the companies used in the

"businesses" file are saved. The resulting product is that the first line in the businesses text file is the company, and the first line in the dates text file is the corresponding IPO address.
The following line, in the script is seen from "pt = TrendReq …… custom_useragent=None)" refers to the program taking the login credentials and using them to sign in on behalf of the researcher into Google. In this line hl stands for "home language" and tz is "time-zone". After logging in the pytrends program executes the code following from "fdata". In the next 12 lines it collects the data from the two text files and makes sure to identify the date in proper format. This script also defines a variable called "monthBefore" which takes the date of the IPO for each company and takes the date 30 days before the IPO. This can be altered to the specification requirements by simply changing the "30" to a different number of days. This paper utilizes 60 days prior to the IPO date. The interest in trending variable comes from the change in interest of the company as it approaches its IPO. The program then queries Google for each individual datapoint making sure to take the values for a month before and the date of the IPO and returning the value for interest that Google has recorded.

4.2.4 Limitations

This script and the addition of the pytrends program automate the data collection. However, availability of the Google trends data is subject to limitations, this means the user must use batches of requests, in the region of 1000 companies per use. This can be repeated every hour and is significantly more efficient and less laborious than data collection by hand.

Futhermore, the pytrends program is constantly being updated and improved. This means that if Google updates its Trending database, the individuals who have built the pytrends program must update the software to work on the new database. This is only a limitation in terms of waiting time. Updating the program may take the individuals some time and for this reason the method may be down for extended periods of time. This limitation was encountered by this paper, meaning the collection of larger numbers of data was not possible as Google had drastically changed the accessibility of their database.

It is hoped that Google will release an API for their Trends in the future such that this workaround would not be necessary. However, until that moment, this script allows researchers an automated method of access to data that they would not be able to use otherwise.

4.3 Descriptive Statistics

Table 1.1 Descriptive statistics for variables used in regression. The top row of the table includes the names of descriptive statistics and the first column refers to the variables included. Excluded are the dummy variables used regression. The units of measurement are given for the individual variables. Those without units are ratios or natural logarithms

	N	Min	Max	Mean	Std. Deviation
%change in ownership	769	0.89%	40.15%	15.09%	10.78%
Change in holdings of Largest 5 owners	618	-726.14%	96.48%	-75.29%	182.75%
Market To Book	769	-4.31 (-355)	16.73 (2337)	4.308	4.48
Log Assets	758	1.79	17.18	10.36	2.05
Log R&D	505	0.69	13.89	8.26	1.49
Leverage	769	$-17064.83	$16608	$-27.22	$1313.08
Freecashflow	733	$-140809	$6582	$-1342.57	$8957.85
Vol	364	13.80	94.58	45.47	14.82
Dividends	762	$0	$50.36	$0.22	$2.59
Trending	101	0	100	50.69	24.735
Total Valid N	32				

The table breaks down the descriptive statistics of the variables used. In the first column the variables can be found. It can be seen from the table that the highest amount of data for any one variable is 769 individual points of data. This sample of data is true for the change in ownership, market to book and leverage. Other variables have similar sample sizes, where dividends has 762 points of data. From this we can conclude that the regression that contains the largest sample of data includes market to book, log assets, leverage, free cash flow and dividends. 3 other regressions must be run to include volatility, log R&D and trending. If all the variables were regressed in the same equation, then the overall sample size would be reduced to 32. This is not large enough to form a conclusive analysis and therefore separate regressions are run.

4.3.1 Windsorizing data

Originally the range of data for market to book (as shown in the brackets below the current value) was much higher. Calculating the 5th and 95th percentile of the variable showed that a few points of data were much larger than the rest of the sample. Windsorizing the data, which replaces the top and bottom 5% of data with the next highest value, returns the range to a more expected level without distorting the end result significantly. This practise was also done with both the dependent variables; change in insider ownership and change in holdings of largest 5 owners. For the variable change in insider ownership, the original range of 0% - 80% was not indicative of the sample as the 95th percentile value was equal to a 40% change in ownership. These three variables have been Windsorized. The minimum and maximum values reported for both variables are equal to the cut-off points before Windsorizing.

4.3.2 Understanding the values

The outcome of the descriptive statistics suggests that on average, a company going public experiences a 15% reduction in insider ownership. This is in line with what Helwege (2007) finds, stating that companies are generally not widely held straight after the IPO. For the percentage reduction in holdings of the largest 5 owners, generally after the IPO the holdings increase. This is shown by the negative mean as seen in the descriptive statistics. The variable measures the percentage decrease in holdings. For example, a value of -75% means that on average after the IPO, for this sample of data the largest 5 owners increase their holdings by 75% of their original number of shares. The measure of largest 5 owners is done in this way as a positive value indicates dispersion of ownership. This has the effect that any variables that positively affect the decrease of holdings indicate that ownership has become dispersed. For the book value the range lies between -4.3 times the market to book ratio, and 16 times this ratio with a mean of 4.3. This means on average in the sample of data a company has a market value four times greater than its book value. There is no cause for concern in this value as many items that add to the value of a company do not show up on the balance sheets and are therefore not incorporated in the calculation. These items such as goodwill, brand awareness and intellectual property make the market value much higher. The largest market to book value is equal to 16.73. In this sample, this value is most often part of companies in the Hightech or

Healthcare sector. Companies in this sector have a large part of their value tied up in intellectual property.

On average a company in the sample size has a natural log of assets equal to 10. Initially the minimum of the sample size was 0. For a natural log this means the lowest amount of assets reported had to equal 1 dollar. Upon further investigation this single data point is significantly below the otherwise reported minimum of 1.79. Due to this, the error is removed and the minimum natural log of assets is equal to 1.79. Removing this data point does not significantly alter any of the other descriptive statistics, as the mean remains constant and the standard deviation falls from 2.08 to 2.05.

For research and development it can be seen that on average in the sample companies spend close to 4000 dollars. The maximum of the range is much higher, surpassing a million dollars. Due to the different industries included in the sample, the mean is likely to be much lower than the range. Companies in the healthcare sector invest to a large extent in R&D, whereas those in the industrial sector invest much less. This is mainly due to the inclusion of pharmaceuticals in the healthcare category. Using the natural log of spending makes the variable linear and easier to study in the regression. In terms of the regression, a change in Ln(R&D) from 1 to 2 will have the same effect as that from 11-12 meaning that the effects of R&D spending are diminishing.

The debt used to finance assets in the company, as represented by the leverage variable in the table, ranges from - $17,064 to $16,608 with a mean near zero. Negative values represent a situation where a company has negative debt, which is equivalent to net cash.

The volatility variable represents the change from the mean to a high and low in the year after the IPO. It proxies for the volatility in the market at the time of the IPO. Across our sample it shows that volatility is relatively high with an average change of 45%. Though this is a measure for volatility, caution must be taken as companies experiencing IPOs often suffer from under-pricing meaning this variable does not fully capture the volatility.

Not included in these descriptive statistics are the dummy variables used in the regression. Their purpose is to control for factors including year of going public, VC backing and sector of operation.

5.0 Results

5.1 Change in insider ownership results

Table 1.2 provides the outputs of the different regressions run. Labelled 1 through 4 are the regressions following the constantly stricter form. This means the first regression run includes the most data points. This can be seen from the 'N' row which is equal to 717 observations for (1). The constant in regression (1) is equal to the change of insider ownership that would take place if all the other variables regressed were equal to zero. In real terms this means that the value in having an IPO allows the insider to reduce their ownership by 26%. With a relatively low standard error, this value is significant at the 1% level.

Contrary to initial hypotheses, various variables including the market to book ratio and free cash flow are insignificant. This is true across all forms of the regression. In reality this would mean that none of these variables factor into the decision making process for inside owners to reduce their levels of ownership. Helwegen's (2007) paper studying dispersion over time finds values contradicting this. Though this study focuses on the time of the IPO, it is unlikely that these values play no part in the decision to reduce insider ownership. This means that the equation does not fully capture the variables that affect the decision process.

This regression finds that a positive value for venture capital backing reduces the insider ownership by 3.6% and is significant at the 1% level. This contradicts our initial hypothesis that VC funding will result in higher levels of insider ownership change as an IPO represents a significant exit strategy for the firm. In reality however, VC backing reduces the levels of ownership change by close to 4%. This is likely explained by the inclusion of lock-up periods

Table 1.2) Output of variables regressed on change in insider ownership

This table looks at the determinants of ownership structure changes (column 1-4) for IPOs in America between 1999-2015. Each column under the numbers 1, 2, 3 and 4, represents the output of a more restricted version of the equation. The addition of variables means the regression contains fewer data points. This can be seen by the 'N' at the bottom of the table. The table also reports the significance levels of the variables where 10%, 5% and 1% are denoted as *, ** and *** respectively. Included in the list are only the dummy variables that were significant in any one of the regressions.

	Dependent Variable: % Change In Insider Ownership			
	(1)	(2)	(3)	(4)
	Original regression	Original regression + R&D	Original regression + Volatility	Original Regression + Trending
Constant	25,947***	25,357***	22.620***	27.969***
	(2,952)	(3.291)	(5.768)	(7.736)
Log R&D		-0.770**		
		(-0.372)		
Volatility			-0.020	
			(0.046)	
Trending				0.015
				(0.041)
Venture Capital	-3,658***	-4.584***	-3.235**	-7.915***
	(0.929)	(1.041)	(1.369)	(2.539)
MarketToBook	-0,118	-1.02	-.160	-0.217
	(0,087)	(0.092)	(0.140)	(0.222)
Log Assets	-0.426*	0.249	0.114	-0.525
	(0.221)	(0.312)	(0.357)	(0.583)
Dividends	0.068	-0.202	0.276	0.257
	(0.146)	(0.176)	(0.223)	(0.179)
Dummy_2002	-10.226**	-6.759	-13.942**	NA
	(3.838)	(4.143)	(5.417)	
Dummy_2003	-5.219*	-0.871	-7.085*	NA
	(2.866)	(4.095)	(3.943)	
Dummy_2007			-5.266*	
			(2.973)	
Dummy_2013	-3.549*	-2.118	-16.718**	
	(1.979)	(2.057)	(6.677)	
Dummy_2014	-5.173**	-2.622	-22.120**	
	(1.854)	(2.011)	(8.092)	
Dummy_2015	-8.039***	-6.410**	-17.909*	
	(1.955)	(2.047)	(10.871)	
DI_HighTech	-3.808**	-3.691**	-4.055*	-5.027
	(1.593)	(1.717)	(2.410)	(3.397)
DI_ConsumerStaples	-4.932*	-10.242*	-8.195**	-16.467*
	(2.752)	(5.279)	(3.979)	(8.421)
DI_Industrial		4.834**	1.952	7.465*
		(2.417)	(2.872)	(4.371)
N	717	461	355	98
Durbin-Watson	1.909	2.018	2.093	2.373
Adj. R Square	0.147	0.187	0.121	0.425

after an IPO. To instill confidence in the market for the recently public company, many large blockholders sign contracts that limits the level at which they may reduce their holdings.

Hypothesis II deals with firm size; at the 10% level of significance this paper rejects the hypothesis that firms with larger real values of assets will have a more dispersed ownership just after the IPO. The relation between the natural log of the value of assets and percentage change in ownership is found to be equal to -0.426. This means that for every increase of 1 in the log value of assets, the amount of insider ownership change drops by 0.426%.

Stulz (1999) found that the market to book ratio contains relevant information. A low ratio indicates there may be high growth opportunities and therefore higher levels of information asymmetry. Though the sign of the coefficient is in-line with this paper's hypothesis, the fact that it is insignificant even at the 10% level means hypothesis iii of this paper is rejected. The exclusion of tangible assets from the regression to deal with the multicollinearity means hypothesis iv is also rejected. Furthermore, hypothesis v that deals with the levels of free cash flow is also rejected. The various regressions found no levels for which free cash flow impacted the change in insider ownership.

Looking at the stricter form of the regression, which includes R&D spending, we can see the calculated correlation of LnR&D on change in insider ownership is equal to -0.77, meaning that for every increase of 1 in LnR&D, the insider ownership changes by -0.77%. The total number of observations is reduced from 717 to 456, however this is still large enough to draw significant conclusion upon. The finding for R&D is line with this paper's hypothesis and we can accept the null hypothesis at the 5% significance level. An economic explanation for this is that R&D spending suffers from information asymmetry as discussed earlier. Insiders of the company are better versed in understanding the potential outcome of the R&D; indicating that insiders value the potential outcome of R&D highly. This in turn means they reduce their holdings by a lower amount than companies that have no R&D spending. The value at which insider holdings change from R&D spending is however subject to diminishing returns. To increase the value of Ln(R&D) from 1 to 2 requires less spending on R&D than to increase the value from 10 to 11. This means that the initial spending on R&D is valued higher by insiders than increases in spending given an already high amount of R&D investment.

Hypothesis vii states that firms that pay out dividends are more likely to have dispersed ownership after their IPO. This paper rejects this hypothesis. Though in all but one of the regressions, the coefficient is in-line with theory, this study does not prove that the level of dividends that will be paid is a significant factor in deciding the level of insider ownership.

Hypothesis viii, stating that higher leverage will result in a higher level of dispersion (lower level of insider ownership), is rejected. No relationship between leverage and insider ownership at the time of the IPO is found.

Hypothesis ix of this paper stating that higher values of market volatility are likely to have a near zero impact on the change of insider ownership cannot be proven. The size of the volatility coefficient at -0.02 means that even if volatility were at its maximum value according to the measurement used (equal to 100), it would only have a 2% effect on the insider ownership. The hypothesis is rejected as the coefficient is not significant at any level of the t-statistic.

Finally, the variable this paper introduces in the form of a trending factor cannot be proven to have significant impact on the levels of insider ownership. The maximum measurement of the trending factor is equal to 100, meaning that if the variable was significant, the impact this significance would have on the levels of insider ownership is negligible. This is however limited to the fact that the trending factor regresses on 98 data points. Meaning that the sample is not large enough to deduce whether this variable can be used in future research. The tool developed by this paper still opens up a cohort of data not previously easily accessible.

5.1.2 Control variables

This paper utilizes two main control variables. These come in the form of an industry control and a date control. Dummy variables for each of these are set up and table 1.2 shows the outcomes only of those that are significant. For the each of the variables one of the cases is left out to avoid multi-collinearity. For the year this is 1999 and for the industry this is consumer products and services. The initial regression shows that various years are significant in the change of insider ownership. At the 5% significance level the year 2002 and 2014 are significant, and at the 1% level of significance the year 2015 has an impact on the dependent variable. Each of these variables reduces the change in insider ownership to a large extent. Companies going public in 2002 for example, would see a 10% reduction in their change in

insider ownership than those going public in 1999. Furthermore, the outcomes of controlling for industry show that companies in the high-tech industry are less likely to reduce insider holdings during an IPO. This is explained by the information asymmetry issue. Companies in technology operate in often times leading technology, meaning the understanding of the scope of their product is not understood as well by the market as it is by the managers. In markets with lower asymmetric information the opposite result is expected. This is seen in the regression including R&D which shows that at the 5% level of significance, companies in the industrial sector reduce insider holdings by close to 5%. Lastly this paper finds that the dummy variable for consumer staples, which refers to essential products such as food etc, reduces the change in insider ownership by 5%. Significant at the 10% level, whether the smaller change in ownership is due to information asymmetries, lack of demand or lock-up periods is hard to deduce.

All the regressions included a Durbin-Watson test for which the value is reported in table 1.2. This checks for serial correlation and from the value ranging between 1.9 and 2.3 can be considered normal. It indicates some level of autocorrelation may be present but that this is no cause for concern.

5.2 Change in holdings of largest 5 shareholders

Table 1.3 represents the output of regressing the balance sheet items against the change in holdings of the largest 5 shareholders. Overall the adjusted R squared values of the output are very low. The variables are not telling of the reasons behind the dispersion of ownership. However, the regression does produce significant values for many of the dummy variables, and in regression (3) even significant levels on the independent variables. In regression (1), other than the constant, none of the regressed variables are considered significant at the 10% significance level. The dummy variables for the year 2005 and 2013 are. This means that given IPOs in the year 2005, top 5 shareholders would have increased their holdings by 50% of the original value, all else equal. Furthermore, the industry of the IPO is very important for the levels of dispersion. IPOs in the industrial sector as well as the Telecom sector saw the largest 5 owners reduce their holdings by 76% of their original value within 3 months of the IPO. This value is significant at the 10% level. IPOs in the healthcare sector saw a similar result where the largest 5 owners reduced their holdings by 77% within 3 months, with the value being significant at the 5% level. A more drastic reduction can be seen in the holdings of Energy and

Power, and Telecom sector companies, where the 5 largest shareholders reduced their holdings by 85% and 84%, respectively, of the original value. These four sectors show a large level of dispersion after IPO. This is in line with the theory as these sectors are ones with lower information asymmetry than sectors such as Hightech. The coefficient of Hightech reflects this as the largest 5 holders of shares increase their holdings by over 100% after the IPO.

5.2.1 Hypotheses outcomes

Using the (3) output from the table we can test the initial hypotheses as values of the independent variables are significant. The danger of this is that the sample size is reduced from an original 577 to 291, to include the volatility variable and therefore the variables may not be an accurate representation of the population.

From this regression, the variables significant are the VC dummy, market to book value and free cash flow. The first hypothesis regards the VC backing, stating that VC backing results in higher dispersion of ownership. At the 5% significance level we find that VC backing results in an increased holding of the 5 largest shareholders after the IPO meaning we can reject the null hypothesis. The third hypothesis regarding the market to book ratio states that higher values will result in higher dispersion. A negative value of -6.092, significant at the 5% level, means that the top 5 shareholders increase their holdings after the IPO for higher values of the market to book value. Hypothesis V states that firms with greater free cash flow are less likely to have dispersed ownership, the coefficient from the regression is positive, meaning that increased values of free cash flow result in a reduction of holdings from the top 5 shareholders. At the 10% level of significance we can reject the null hypothesis.

The data fails to accept the other hypotheses as the variables are deemed not significant in the decision making process for the largest shareholders to determine their change in holdings.

Table 1.3) Output of variables regressed on change in holdings of largest 5 shareholders

This table looks at the determinants of the change in holdings of largest 5 shareholders (column 1-4) for IPOs in America between 1999-2015. The dependent variable is measured as the reduction in holdings of the largest 5 shareholders, meaning a negative number is equal to increased hodings. Each column represents the output of a more restricted version of the equation. The addition of variables means the regression contains fewer data points. This can be seen by the 'N' at the bottom of the table. The table also reports the significance levels of the variables where 10%, 5% and 1% are denoted as *, ** and *** respectively. Included in the list are only the dummy variables that were significant in any one of the regressions.

	Dependent Variable: % Change largest 5 shareholder holdings			
	(1)	(2)	(3)	(4)
	Original regression	Original regression + R&D	Original regression + Volatility	Original Regression + Trending
Constant	-141.472**	-239.275**	-140.658	-182.866
	(65,716)	(88.495)	(102.962)	(165.929)
Log R&D		4.943		
		(8.766)		
Volatility			-0.100	
			(0.855)	
Trending				0.434
				(0.848)
Venture Capital	-28.509	-23.298	-49.714**	63.662
	(18.873)	(25.227)	(24.739)	(57.667)
MarketToBook	-0,382	1.303	-6.092**	-8.889*
	(1.772)	(2.235)	(2.530)	(4.767)
Free cash flow	-0.001	0.001	0.004*	
	(0.001)	(0.001)	(0.002)	
Dummy_1999			105.613*	NA
			(54.251)	
Dummy_2002		-175.886*	-210.608**	NA
		(99.947)	(102.752)	
Dummy_2005	-50.799*			
	(30.371)			
Dummy_2012		-87.143**		
		(42.8643)		
Dummy_2013	-132.177***	-127.457**	-221.279*	-134.970**
	(32.178)	(36.634)	(125.083)	(67.353)
DI_HighTech	-81.490**	99.329**		
	(33.213)	(42.545)		
DI_ Energy and Power	85.404*	157.120*		
	(48.283)	(91.271)		
DI_Healthcare	77.526**	89.004**	92.619**	
	(34.802)	44.060	(46.795)	
DI_ConsumerStaples	-103.115**			
	(56.870)			
DI_Industrial	79.654*	123.735**	90.564*	
	(41.756)	(60.114)	(51.732)	
DI_Materials		144.625*		
		(82.725)		
DI_Telecom	83.957*	112.016*		
	(47.800)	(63.844)		
N	577	395	291	85
Durbin-Watson	1.867	1.725	1.890	1.781
Adj. R Square	0.034	0.039	0.050	-.077

6) Conclusion

This paper studies the variables that affect the level of insider ownership, as well as a more detailed measure of dispersion at the time of the IPO. For insider ownership, the variables VC backing, R&D spending as well as dummy variables for timing and industry are significant in explaining why insiders reduce their holdings after the IPO. Surprisingly, results show that VC backing reduces the change of insider ownership at the time of the IPO and is significant at the 1% significance level. This means that companies with VC backing maintain higher levels of insider ownership than those that are not backed. This is a significant result as research focused on studying dispersion over time shows IPOs to be an exit strategy for most VCs.

For the dependent variable of largest 5 shareholders, the variables VC backing, free cash flow and market-to-book ratio are significant, but only in a restricted version of the regression that uses a smaller sample size. Otherwise none of the hypothesized variables are significant in explaining the change in holdings of the largest 5 shareholders.

This paper fails to prove the significance of the market to book ratio, free cash flow, tangible assets, leverage, volatility and trending factor for the change in insider holdings as well as largest 5 shareholders. The variables studied were taken from Helwege's (2007) paper on dispersion over time. These variables are not significant in deducing the changes in insider ownership at the time of IPOs or in explaining dispersion according to Overland's (2012) measure. Future research can focus on expanding the variable list to deduce the variables that are significant. A high intercept of around 25% as well as relatively low values of the adjusted R squared means that variables can be found to better explain the changes in insider ownership during the IPO phase. This is also true for the measure of intercept and R squared recorded for the largest 5 owners as dependent variable. The results of this paper push the understanding of levels to which different variables influence the change in ownership at the time of the IPO. Furthermore, though the relevance of the trending factor cannot be proven with the sample size in this paper, the tools created for utilizing this variable remain useful for future researchers to take advantage of.

References:

Baumol, W. J. (1959). Business behavior, value and growth. New York, 32.

Berle, Adolf A., and Gardiner C. Means. The Modern Corporation and Private Property. New York: Macmillan, 1932.

Black, B. S., & Gilson, R. J. (1998). Venture capital and the structure of capital markets: banks versus stock markets. *Journal of financial economics*, 47(3), 243-277.

Bolton, P., & Thadden, V. (1998). Blocks, liquidity, and corporate control. *The Journal of Finance*, 53(1), 1-25.

Brealey, R., Leland, H. E., & Pyle, D. H. (1977). Informational asymmetries, financial structure, and financial intermediation. *The journal of Finance*, 32(2), 371-387.

Clarke, J, et al, 2004, The long-run performance of secondary equity issues: A test of the windows of opportunity hypothesis, *Journal of Business* 77, 575–603.

Coles, J. L., Lemmon, M. L., Meschke, J. F., Bessembinder, W. T. H., Bhagat, S., Hadlock, C., ... & Perry, T. (2003). Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance. Working Paper.

Demsetz, H, (1983), The Structure of Ownership and the Theory of the Firm, J*ournal of Law and Economics*, 26, issue 2, p. 375-90,

Demsetz, H., & Lehn, K. (1985). The structure of corporate ownership: Causes and consequences. *Journal of political economy*, 93(6), 1155-1177.

Demsetz, H., & Villalonga, B. (2001). Ownership structure and corporate performance. *Journal of corporate finance*, 7(3), 209-233.

Downes, D. H., & Heinkel, R. (1982). Signaling and the valuation of unseasoned new issues. *The Journal of Finance*, 37(1), 1-10.

Grossman, S. J., & Hart, O. D. (1980). Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics*, 42-64.

Helwege, J., Pirinsky, C., & Stulz, R. M. (2007). Why do firms become widely held? An analysis of the dynamics of corporate ownership. *The Journal of Finance*, 62(3), 995-1028.

Himmelberg, C. P., Hubbard, R. G., & Palia, D. (1999). Understanding the determinants of managerial ownership and the link between ownership and performance. *Journal of financial economics*, 53(3), 353-384.

Holderness, C. G., & Sheehan, D. P. (1988). The role of majority shareholders in publicly held corporations: An exploratory analysis. *Journal of financial economics*, 20, 317-346.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of financial economics, 3(4), 305-360.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American economic review*, 76(2), 323-329.

Keim, D. B., & Madhavan, A. (1996). The upstairs market for large-block transactions: Analysis and measurement of price effects. *The Review of Financial Studies*, 9(1), 1-36.

La Porta, R., Lopez-De-Silanes, F. and Shleifer, A. (1999), Corporate Ownership Around the World. *The Journal of Finance*, 54: 471–517.

Leech, D., & Leahy, J. (1991). Ownership structure, control type classifications and the performance of large British companies. *The Economic Journal*, 101(409), 1418-1437.

Lehmann, E., & Weigand, J. (2000). Does the governed corporation perform better? Governance structures and corporate performance in Germany. Review of Finance, 4(2), 157-195.

Marris, R. (1964). The economics of capital utilisation: A report on multiple-shift work (No. 10). CUP Archive.

Michael C. Jensen, William H. Meckling, (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure, In *Journal of Financial Economics*, Volume 3, Issue 4, 1976, Pages 305-360,

Morck, R., Shleifer, A., & Vishny, R. W. (1988). Management ownership and market valuation: An empirical analysis. *Journal of financial economics*, 20, 293-315.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of financial economics*, 58(1), 215-260.

Overland, C., Mavruk, T., & Sjögren, S. (2012). Keeping it real or keeping it simple? Ownership concentration measures compared. In The 4th workshop of the Nordic Corporate Governance Network, Reykjavik University June 15, 2012.

Roe, M. J. (1994). Strong managers, weak owners. Princeton, NJ: *Princeton University Press*.

Roe, M. J. (2003). Political determinants of corporate governance: Political context, corporate impact. *Oxford University Press on Demand.*

Shleifer, A., & Vishny, R. (1986). Large Shareholders and Corporate Control. *Journal of Political Economy*, 94(3), 461-488.

Stulz, R. (1990). Managerial discretion and optimal financing policies. *Journal of financial Economics*, 26(1), 3-27.

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

Thomsen, S., & Pedersen, T. (2000). Ownership structure and economic performance in the largest European companies. *Strategic Management Journal*, 689-705.

Williamson, O. E. (1964). The economics of discretionary behavior: Managerial objectives in a theory of the firm. *Prentice-Hall.*