



Digging for Fool's Gold

*An Empirical Study on Factors Affecting Initial Stock Performance of
Venture Capital-Backed IPOs*

Kaja Røttingsnes & Jonas Lier Gjørum

Supervisor: Trond M. Døskeland

Master thesis, Economics and Business Administration

Major: Financial Economics and Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

This thesis is written as part of our Master of Science in Economics and Business Administration at NHH, with majors in Business Analytics and Financial Economics. It has been a challenging and exciting semester, and we are grateful we had the opportunity to spend an entire semester researching a topic of our choice.

We would like to thank our supervisor, Trond M. Døskeland, for valuable guidance and feedback throughout the semester. As this thesis marks the end of our time at NHH, we would also like to thank NHH, NHHS and all the amazing people that have made these years so memorable.

Abstract

We analyze the performance of venture-backed IPOs on the New York Stock Exchange and Nasdaq between 2011 and 2019. Throughout this period, a large number of venture-backed tech companies with billion-dollar valuations have gone public, and many have experienced significant valuation cuts during their first months of trading.

By using multiple regression analysis and the Mann-Whitney U test, we find evidence of a positive relationship between offer size and first-day returns. We also find that tech companies and unprofitable companies achieve higher first-day returns than other companies. When looking at the three months after the first day of trading, the analyses indicate opposite effects, and we find that unprofitable tech companies going public achieve significantly lower returns than other companies. However, results for the three-month time period are in general less conclusive than those for the first-day of trading. Contrary to our hypothesis, the amount of pre-IPO funding does not seem to affect aftermarket performance.

Keywords – NHH, master thesis, initial public offerings, venture capital

Contents

1	Introduction	1
1.1	Motivation and Background	1
1.2	Problem Definition	2
2	Theory	3
2.1	Relevant Theory and Definitions	3
2.2	Literature Review	5
2.3	Hypotheses	9
3	Data Collection and Cleaning	11
4	Descriptive Analysis	17
5	Analyzing Individual Attributes	28
5.1	Method: Multiple Regression Analysis	29
5.1.1	Model Formulations	29
5.1.2	Explanation of Model Variables	29
5.1.3	OLS Violations	31
5.1.4	Regression Results	33
5.2	Findings	34
5.2.1	Amount of Pre-IPO Funding	34
5.2.2	Tech Companies vs. Non-Tech Companies	36
5.2.3	Offer Size	38
5.2.4	Profitable Companies vs. Unprofitable Companies	39
5.3	Summary and Discussion	40
6	Analyzing Combinations of Attributes	42
6.1	Rephrasing the Hypothesis	42
6.2	Method: Mann-Whitney U Test	43
6.2.1	Test Results	44
6.3	Findings	45
6.3.1	Tech Companies With Large IPOs	45
6.3.2	Unprofitable Tech Companies	46
6.3.3	Unprofitable Companies With Large IPOs	47
6.3.4	Unprofitable Tech Companies With Large IPOs	48
6.4	Summary and Discussion	49
7	Conclusion	50
	References	52
	Appendix	54
A1	Fuzzy String Matching	54
A2	OLS Violations	55

List of Figures

2.1	Typical Funding Story of a Successful Startup Firm	3
4.1	Distribution of First-Day Simple Returns	20
4.2	Distribution of Simple Returns Three Months After IPO	20
4.3	Distribution of Dataset	21
4.4	Average Simple Returns, by ISIN Industry Section	22
4.5	Distribution of Offer Sizes	23
4.6	Plotting Simple Returns Against Offer Size	23
4.7	Distribution of Pre-IPO Funding Amount	24
4.8	Plotting Simple Returns Against Pre-IPO Funding Amount	24
4.9	Distribution of Company Age at IPO	25
4.10	Plotting Simple Returns Against Company Age at IPO	25
4.11	% of Venture-Backed IPOs with Positive Profit Margin year before IPO	26
4.12	Distribution of Profit Margin Year Before IPO	27
4.13	Plotting Simple Returns Against Profit Margin Year Before IPO	27
5.1	Indexed Average Returns, Much vs. Not-So-Much Funding	35
5.2	Indexed Average Returns, by Industry	36
5.3	Indexed Average Returns, Tech vs. Non-Tech IPOs	37
5.4	Indexed Average Returns, Large vs. Not-So-Large IPOs	38
5.5	Indexed Average Returns, Profitable vs. Unprofitable Companies	39
6.1	Overlap Between Large IPOs and Heavily Funded Companies	42
6.2	Overlapping Characteristics	43
6.3	Indexed Average Returns, Tech Companies With Large IPOs	45
6.4	Indexed Average Returns, Unprofitable Tech Companies	46
6.5	Indexed Average Returns, Unprofitable Companies With Large IPOs	47
6.6	Indexed Average Returns, All Three Attributes	48
A2.1	Distribution of Residuals, u_{hat}	58
A2.2	Distribution of Non-Logged Variables	59
A2.3	Distribution of Log-Transformed Variables	59

List of Tables

3.1	Most Active Investors	13
4.1	Descriptive Statistics, by Year	18
4.2	Average Simple Returns	19
5.1	Regression Results Against Abnormal First-Day Return	33
5.2	Regression Results Against Abnormal Three-Month Return, Excluding the First Day of Trading	34
6.1	Mann-Whitney U Test: First-Day Return	44
6.2	Mann-Whitney U Test: Three-Month Return	45
A1.1	Stock Splits Identified for Tickers in Dataset	54
A1.2	Fuzzy String Matching of Company Names	54
A2.1	Correlation Matrix	56

1 Introduction

1.1 Motivation and Background

The initial public offering market of 2019 has been strange. High-profile unicorns¹ such as Uber, Lyft and Slack have gone public this year, with less than stellar performance - all trading well below their offer prices shortly after their initial public offerings (IPOs). Another highly anticipated IPO, WeWork, was delayed indefinitely after public markets basically revolted against their \$47 billion valuation. This was later cut down to \$8 billion (Klebnikov, 2019). As of 30 September 2019, companies such as Groupon, Spotify and Snap were also trading below their IPO offer prices (Carlson, 2019).

These companies share a few key characteristics:

- They all define themselves as “tech companies”
- Before they went public, they were heavily funded by venture capitalists
- Their IPOs were among the largest of 2019
- They were unprofitable when they went public²

It is easy to draw lines between the IPO class of 2019 and the dot-com boom in the late 90’s. Not since then have so many highly valued tech companies gone public in such short time (Grocer and Russell, 2019). Both then and now, tech companies spent enormous amounts on advertisement, bragged about their ability to change the world, but struggled to find a path to profitability. And both then and now, valuations of billion dollar tech-companies have been dramatically reduced in a matter of weeks (Thompson, 2019).

This mildly troubling development forms the backdrop for this thesis, where we will search for patterns for which companies experience similar stock price movements after going public.

¹Term used to describe a privately held startup company valued at over \$1 billion

²Grocer and Russell (2019)

1.2 Problem Definition

As the stock exchanges in the world where most high-profile IPOs take place, we focus on new listings on the New York Stock Exchange (NYSE) and Nasdaq. Since 2010, there have been around 1 600 listings on these two stock exchanges, including all the aforementioned companies. In light of the growing venture capital industry and increasing amounts of private investments in unlisted companies, we will focus on companies that have raised external funding in the private market prior to going public. For the sake of convenience, we will refer to all private funding as venture capital. Specifically, we will investigate whether their stock performance in the first three months of trading can be explained by selected publicly available information, such as the four characteristics mentioned above.

The intention of this exercise is *not* to predict after-market performance of specific future IPOs, which would demand a much higher level of detail and careful consideration of individual companies. Rather, we will address macro trends in the IPO landscape, such as increasing access to private equity financing and seemingly decreasing investor demands for profitability in the immediate future. As such, we aim to identify correlations rather than causalities.

In order to do this, we will use statistical analysis to explore the relationship between stock returns and selected explanatory variables, including amount of pre-IPO funding, IPO size, profitability, and industry. When analyzing post-IPO stock performance, first-day trading gains are most often analyzed separately. Thus, two separate time horizons for stock returns will be examined: returns from the first day of trading after IPO, and from the subsequent three months of trading, excluding first-day trading gains. The three-month time horizon is chosen because it enables investigation of short-to-medium-run returns while also capturing as many of the 2019 IPOs as possible.

2 Theory

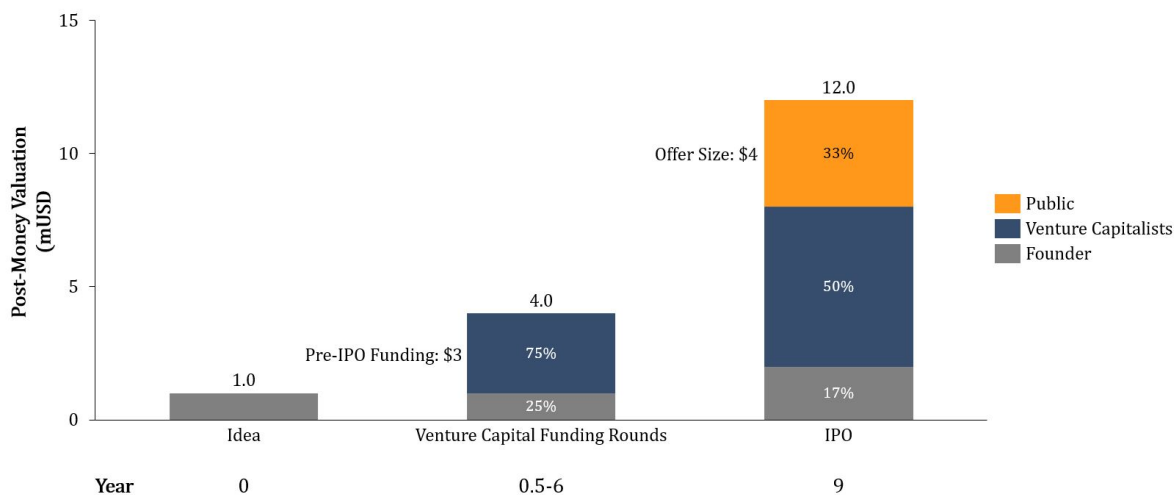
Before diving into the dataset of our thesis, established theory and previous studies will be examined to further understand our problem definition. We will start by introducing the most important financial expressions used in the thesis, before studying the key findings of previous research. These findings, together with the trends mentioned in the chapter regarding our motivation, will form the hypothesis of our thesis.

2.1 Relevant Theory and Definitions

Equity Financing for Private Companies

As the thesis investigates companies that raised external capital before going public, it is relevant to assess the processes for equity financing of private companies. In *Corporate Finance* (2017), Berk and DeMarzo outline the typical equity financing life cycle of a successful startup firm. Figure 2.1 illustrates a typical funding story for a successful firm, from idea to IPO, where all funding rounds have been added together. The length of each bar indicates the firm's post-money valuation, in other words the enterprise value of the firm after investments have been made, or after the firm has gone public. The percentages within each bar indicate the ownership split between the founder, venture capitalists, and the public. In the example below, the firm raised \$3 million in external funding (pre-IPO funding) before going public, and raised another \$4 million in their IPO (offer size).

Figure 2.1: Typical Funding Story of a Successful Startup Firm



When someone starts a business, they usually provide the initial capital necessary by themselves or with help from their immediate family. As the business grows, however, the need for increasingly larger amounts of capital makes it necessary to seek sources of outside capital. This funding can be raised from several potential sources, such as angel investors, venture capital firms, institutional investors, corporate investors, and crowdfunding. In this thesis, we will not distinguish between funding from different sources, but simply refer to all as *pre-IPO funding*.

The first outside capital a company raises is typically from a so-called *angel investor*. An angel investor is often a wealthy individual willing to help new companies get started in exchange for a share of the business. This investment is typically followed by a *seed round*, where the source can for instance be angel investors, groups of angel investors, accredited investors or equity crowdfunding groups (Sarath, 2019). Following the seed round, the next rounds of funding are commonly named alphabetically, and referred to as Series A, Series B and so on.

Although all pre-IPO capital can be described as venture capital, these alphabetically named rounds are considered the stage where *venture capital firms* join the party.

Venture Capital

A *venture capital (VC) firm* is a limited partnership that specializes in raising money to invest in the private equity of young firms (Berk and DeMarzo, 2017). Investors in venture capital funds are typically large institutions, such as pension funds, financial firms, and insurance companies (Zider, 1998).

Over the past 50 years, the venture capital industry has grown enormously. Both when measured in number of deals and in total investment volume, the sector grew steeply in the 1990s, and peaked at the height of the internet boom in 2000. The activity declined significantly after the tech bubble burst, but has since then grown back to late 1990s levels, despite a dip after the 2007-2008 financial crisis (Berk and DeMarzo, 2017).

In order to cash in on their investment, venture capitalists have two main exit strategies to choose from: through an acquisition, or through a *public offering* (Berk and DeMarzo, 2017).

The Initial Public Offering

An initial public offering is the process in which a company sells shares of its stock to the public for the first time. The shares sold in the IPO may either be new shares that raise new capital, or existing shares that are sold by the current shareholders (Berk and DeMarzo, 2017). The first-mentioned is called a *primary offering*, and is typically used to finance further growth. The latter is called a *secondary offering*, and is the action that offers a direct exit opportunity for the venture capitalists at the time of the IPO. Alternatively, the venture capitalists holds on to their share until some time after the IPO, and sell them on the public market(Berk and DeMarzo, 2017).

The amount of capital raised in an IPO can be referred to in several ways. We will mainly use *offer size* or *IPO size*. As a company does not have to list 100% of their shares in their IPO, the offer size must not be confused with the enterprise value of a company.

2.2 Literature Review

The Changing Role of the IPO

The main motive of going public has traditionally been to achieve greater liquidity and better access to capital (Berk and DeMarzo, 2017). Through both the initial and subsequent offerings, companies are typically able to raise larger amounts of capital in the public compared to the private markets. Increased liquidity of the firm's stock, as a consequence of public trading and continuously adjustments of market values, makes the company a less risky investment. This can contribute to easier access to additional capital (Beck, 2017).

However, as mentioned above, public offerings are not only a means of accessing external capital, but also a typical exit strategy for existing investors. In an article published in 2018, Døskeland and Strömberg find that the private capital markets are increasing relative to the public markets, and particularly growth companies in the tech industry stay private longer. Thus, the role of the IPO might be shifting from mainly being a way of raising capital to mainly being a way for venture capitalists to cash in on their investments in startup companies.

Underpricing

When assessing research on aftermarket performance, the phenomenon of IPO underpricing is widely discussed in academic literature. Underpricing can be defined as when the IPO price is lower than the closing price at the first trading day of an IPO (Berk and DeMarzo, 2017), and will often be referred to as *first-day returns* in this thesis. In the time period between 1960 and 2015, the average first-day return in the US market was 17% (Berk and DeMarzo, 2017), which serves as evidence of considerable average underpricing. As company shares are sold at a price below what the market would evidently be willing to pay, underpricing is a substantial indirect cost for the company conducting the IPO (Loughran and Ritter, 2002). In 2002, Loughran and Ritter calculated that the average IPO leaves \$9.1 million on the table, which is twice of the average fees paid to investment bankers.

In the IPO process, the party that buys the securities from the issuing company and sells them on to investors is called an underwriter (Banton, 2019). When underwriters provide firm commitments, they expose themselves to a certain risk: if they overestimate the interest in the market, the underwriter might end up having to sell shares at a lower price than they bought them for, thereby taking a loss (Banton, 2019). To minimize their own exposure to this risk, underwriters hence have the incentive to intentionally underprice IPOs (Berk and DeMarzo, 2017). Loughran and Ritter (2002) stated that only 9% of all US IPOs between 1990 and 1998 experienced a fall in share price on the first trading day, and for 16% the offer price is the same as the price at the end of the first day. In other words, the majority of IPOs in this period experienced a price increase, meaning they were underpriced.

Aftermarket Performance

While IPOs tend to be underpriced at offer, they appear to underperform in the long-run. Ritter (1991) defined “long-run” as a three-year period, calculated from the closing market price on the first day of public trading to the market price on the three-year anniversary. First-day returns are excluded to isolate underpricing from long-run performance. In his studies from 1991, Ritter found that his sample of 1 526 American IPOs in the period from 1975 to 1984 significantly underperformed against a set of comparable firms that

went public longer back in time. Together with Ivo Welch, he later confirmed his findings with the article “*A Review of IPO Activity, Pricing, and Allocations*” published in 2002. Ritter and Welch (2002) investigated IPOs between 1980 and 2001 and found that for the subsequent three years the IPOs underperformed the market by an average of 23.4%.

Like Ritter, we will exclude first-day trading gains when looking at aftermarket performance. However, in light of the disappointing performance of the mentioned IPO class of 2019, we want to investigate more immediate effects of the IPOs. Consequently, this thesis will assess the aftermarket performance for the first *three months* after going public.

Overestimating Growth Opportunities

Ritter (1991) mentions risk mismeasurement and bad luck as possible explanations for the phenomenon of underperformance in the aftermarket, as well as the scenario of “fads and over-optimism”. The latter is particularly interesting, as it is likely to be valid for highly anticipated IPOs such as Uber, Lyft and Slack. The term “fads” was established by Shiller (1990) as “scenarios of firms going public when investors are irrationally over-optimistic about the future potential of certain industries” (Ritter, 1991). Ritter found that firms with high adjusted first-day returns have a tendency to have the poorest aftermarket performance (Ritter, 1991). As this effect is stronger for smaller and younger companies than for more established companies, it might indicate that a tendency of over-optimism exists; growth opportunities of IPOs seem to be systematically overestimated by investors.

There has been found evidence that the company’s age at the time of the IPO strongly correlates with its aftermarket performance (Ritter, 1991). In his research, Ritter found that relatively young growth companies were well represented among the long-term underperformers, especially among those that IPOed in high-volume years. The same group also tends to have the greatest first-day return. Conversely, more established companies experience the opposite effect; less underpricing and higher long-run returns. In their analysis of first-day returns of IPOs in Taiwan, Shen and Goo (2019) found a positive correlation between aftermarket risk and first-day returns. Interpreting age as a proxy for aftermarket risk, where young companies are more risky than older companies, their observations are consistent with the concept that risky issues experience higher average first-day returns (Ritter, 1991).

Similar patterns are shown when assessing the size of the IPO and their IPO performance. Ritter (1991) found that smaller IPOs have the highest average adjusted first-day returns, as well as the poorest aftermarket performance. These findings confirmed his previous work with Randolph P. Beatty in 1986, where they argued that also the offer size could be used as a proxy for risk. Smaller offerings tend to be more speculative than larger offerings, hence experiencing higher first-day return (Ritter and Beatty, 1986).

This finding is particularly interesting when investigating the extreme underperformance of the mentioned unicorns that went public in 2019. As the largest IPOs of today do not involve well-established and profitable companies, but rather growing tech companies without a certain path to profitability, the findings of Ritter and Beatty from 1986 may no longer be valid.

The Effect of Being Backed by Venture Capital

Earlier this year, Crunchbase, a leading platform for company insights, published an article stating that there are more VC funds than ever, that the average transaction size is increasing and that giant tech companies are able to raise private financing rounds of sizes that were never possible before (Rowley, 2019). Yet, there has not been done much research on how the amount of pre-IPO funding relates to first-day or long-term stock performance. However, there has been done research on how venture-backed companies perform compared to non-venture-backed companies. The conclusions vary; whereas Megginson and Weiss (1991) found that venture capital-backed IPOs between 1983 and 1987 were significantly less underpriced than others, Ritter (2019) found the opposite for IPOs between 2009 and 2019. Given the more recent dataset, which also covered a longer period of time, we consider Ritter's findings more relevant for our work.

Concerning long-run performance, Brav and Gompers (1997) provide evidence of better performance for VC-backed IPOs. They emphasize factors such as access to top-tier investment banks, less information asymmetries and a demand of better corporate governance as benefits of being venture-backed. Additionally, Brav and Gompers argue that smaller non-venture-backed IPOs are more likely to be owned by individuals, who are more likely to be affected by fads. The VC effect might be linked to Ritter and Beatty's findings of 1986 regarding a company's size at IPO. They found that VC-backed IPOs

usually are bigger because the amount of funding and guidance received have helped them grow.

2.3 Hypotheses

The overall problem definition of our thesis was introduced in chapter 1. We wish to explore the relationship between stock returns and the following four attributes: amount of pre-IPO funding, IPO size, negative profitability, and status as tech company. The research assessed in this chapter forms the basis of what we expect to find when analysing the mentioned relationships, and therefore define our hypotheses.

There are typically high expectations related to large IPOs, and to IPOs of heavily venture-backed companies. This also applies to tech companies, which can lead to investors being over-optimistic of IPOs with either of these three characteristics. Further, both Ritter and Beatty (1986), and Shen and Goo (2019) have found that risky IPOs in general experience higher first-day returns than other IPOs. As future profitability cannot be taken for granted, companies that have yet to deliver positive net earnings are more risky to invest in than other companies. With all of this in mind, our first hypothesis is as follows:

Hypothesis 1: status as tech company, amount of pre-IPO funding, IPO size, and negative profitability are all positively correlated with first-day return

Because firms with high first-day returns typically underperform in the longer run, our second hypothesis is as follows:

Hypothesis 2: status as tech company, amount of pre-IPO funding, IPO size, and negative profitability are all negatively correlated with three-month return, excluding first-day trading gains

Lastly, we expect companies that hold multiple of the attributes to be even more affected, leading to our third and final hypothesis:

Hypothesis 3: When combining the characteristics, correlations on both first-day return and three-month return will be even stronger than they are individually

We hope to contribute to prior research done on IPOs by testing these hypotheses on a recent dataset consisting only of venture-backed IPOs. To our knowledge, there exists

no literature on this specific group, and neither on the relationship between amount of pre-IPO funding and post-IPO stock performance. In light of the increasing amounts of private investments in unlisted companies, as well as other already mentioned macro trends in the IPO landscape, we hope our work will be particularly relevant to the IPO market going forward.

3 Data Collection and Cleaning

In order to answer our hypothesis, we need data on four key areas:

1. A set of companies going public to examine
2. Pre-IPO funding rounds for these companies
3. Pre-IPO accounting numbers for these companies
4. Post-IPO stock performance for these companies

In this chapter, we will explain where this data was found, as well as how it was cleaned and compiled.

Initial Public Offerings

In order to investigate post-IPO stock performance, we need information about all the IPOs that have taken place at NYSE and Nasdaq within the relevant time frame. The Securities Data Company (SDC) Platinum database is ideal for this purpose. SDC Platinum is an online historical financial transactions database, providing detailed information on financial transactions such as new issues, mergers and acquisitions and private equity. As we will use accounting data from the year before a company's IPO for our analyses, and the accounting data goes back to 2010, we use collect data from SDC Platinum on all IPOs after 1 January 2011. Because they are critical for our analyses and data handling, we remove entries where either offer price per share, offer size, or *International Securities Identification Number (ISIN)* is missing. This returns a dataset with 1 531 IPOs.

Adjusting IPO Offer Prices for Stock Splits

To boost the liquidity of their shares, a company can divide its existing shares into multiple shares. This action is called a *stock split*, and reduces the value of each share by the factor induced by the *stock split ratio*. Conversely, companies can also conduct *reverse stock splits* to increase the price of the stock. While the source from which we fetch historical stock prices adjust for stock splits in their database, SDC do not. Because we intend to use the offer price as a basis for calculation of returns, we hence need to adjust the offer prices from SDC Platinum for stock splits.

Historical data on stock splits is available at Investing.com’s Stock Split Calendar. In order to efficiently extract data on all stock splits since 1 January 2011, we use a Python script based on the *Beautiful Soup* library to scrape the Stock Split Calendar. This returns a total of 3 230 stock splits, with split ratios ranging from 1:25 000 to 1 333:1. Each ratio is then converted to a multiplier by finding the inverse of the ratio’s numerator and denominator. We create an offer price multiplier by calculating the product of all the multipliers for the same stock.

We then merge the stock split dataset with the IPO dataset. As companies can change tickers over time, we might experience cases where a ticker belongs to different companies in the two datasets. To overcome the possibility of mismatching offer price multipliers and offer prices, and thus corrupting our dataset, we use a fuzzy string matching algorithm. After merging the datasets using the company ticker as key, we test the similarity between the company name column from each dataset. The results from this exercise are displayed in table A1.1 in appendix A1.

Using the ticker as key, we found 13 matches between the stock split dataset and the IPO dataset, and fuzzy matching uncovered 1 case where the ticker had changed owner since the IPO. Thus, the offer price of 12 companies was adjusted for stock splits (and reverse splits), with multipliers ranging from 150 to 0.17. When cross-checking against Seeking Alpha, widely considered a reliable source of information about financial markets, we find no indications that these multipliers are incorrect, and thus keep the adjusted offer prices in the dataset.

Funding Rounds

As for IPOs, we collect data on funding rounds from SDC Platinum. One of their products, VentureXpert, contains detailed information on investments in unlisted companies. As we intend to investigate the relationship between pre-IPO funding and post-IPO stock performance, we download data on pre-IPO investments in companies later listed on either New York Stock Exchange or Nasdaq.

For each investment, we collect the name and ticker of the company that was invested in, disclosed and estimated investment amount, round number, and name and type of investor. For investments where the disclosed investment amount is not available, we use

the estimated investment amount as replacement. In order to capture as much information as possible, we collect data on all investments available in the database, dating back to 1965. This yields a dataset of 54 362 individual investments in 4 512 different companies. Investments where neither disclosed nor estimated amount is available are removed from the dataset, returning 48 676 individual investments in 3 902 different companies. Table 3.1 provides an overview of the most active investors in our dataset, measured by the value of investment rounds they have participated in. As SDC do not distinguish between types of investors, both venture capital and private equity investments are included.

Table 3.1: Most Active Investors

Investor	Number of Investments	Sum of Investment Rounds Participated In (mUSD)
Carlyle Investment Management	64	33 053
Thomas H Lee Partners	56	18 383
New Enterprise Associates	620	14 641
Sequoia Capital Operations	402	12 033
Bessemer Venture Partners	276	6 662
Technology Crossover Ventures	129	5 968
Energy & Minerals Group	3	4 284
RBC Capital Markets Corp	2	4 274
Boyu Capital Advisory Co Ltd	2	4 150
Polaris Growth Management LLC	220	3 903

As mentioned in chapter 2, investments in non-listed companies often happen in rounds where the company raises funds from multiple investors at once. When a funding round is announced, it is typically the total amount raised from all investors which is disclosed, and not the amount invested by each individual investor. In order to avoid funding round duplications, we therefore group the funding rounds by company name and round number, returning a final dataset with 14 713 funding rounds, still for 3 902 companies.

As investment amounts are registered in nominal US dollar value, we use the *cpi* Python library to adjust investment amounts for inflation in accordance with the US Consumer Price Index. For later analyses, we then compute the total amount of funding raised pre-IPO for each company, and merge with the IPO dataset.

Accounting Data

In order to investigate the relationship between pre-IPO profitability and post-IPO stock performance, we collect accounting information from the Orbis database. According to themselves, Orbis is “the world’s most powerful comparable data resource on private companies”. Because the accounting data from Orbis will be matched and merged with the IPO dataset, we gather as much data from Orbis as possible in order to maximize the number of matches. Thus, we download historical data on revenue, net income, industry code, total assets and number of employees for all current and formerly publicly listed firms in the database, 121 973 in total. This includes companies listed on all stock exchanges in the world. As we will merge this data with data from SDC Platinum using ISIN, we remove rows where ISIN is missing, leaving 105 725 companies.

Orbis provides historical data from 2010 and onwards, limiting us to assess worldwide IPOs from 2011 and onwards. We remove companies with IPOs before this, thereby reducing the number of companies in the Orbis dataset to 18 220. We then calculate profit margin by dividing net income by revenue.

After merging with the IPO dataset, the size of the dataset is reduced to 1 250 companies. At this point, the only data point missing is the target variable.

Stock Data

As stated earlier, we aim to see whether pre-IPO information can be used to explain aftermarket stock performance. Thus, we gather historical stock price development for the companies in the merged dataset from Yahoo Finance, using the *yfinance* Python library. As we will investigate the stock price development following the IPO, we gather data for the first 63 trading days, equivalent to three calendar months. In order to adjust for market development in the same period, we also gather data for the development of the S&P 500 index in the same period. The S&P 500 is a stock market index that measures the performance of the 500 largest companies listed on US stock exchanges, and is widely considered to be a good representation of the US stock market. Thus, we will use this as a baseline to calculate abnormal returns. For first-day returns, where the IPO date is t , this is done as follows:

$$\text{First-Day Return} + 1 = \frac{\text{Closing Price}_t}{\text{Offer Price}}$$

$$\text{S\&P Return} + 1 = \frac{\text{S\&P Closing Price}_t}{\text{S\&P Closing Price}_{t-1}}$$

$$\text{Abnormal First-Day Return} = \frac{\text{First-Day Return} + 1}{\text{S\&P Return} + 1}$$

Similarly, abnormal returns for the three months subsequent to the first day of trading are calculated as follows:

$$\text{Three-Month Return} + 1 = \frac{\text{Closing Price}_{t+62}}{\text{Closing Price}_t}$$

$$\text{S\&P Return} + 1 = \frac{\text{S\&P Closing Price}_{t+62}}{\text{S\&P Closing Price}_t}$$

$$\text{Abnormal Three-Month Return} = \frac{\text{Three-Month Return} + 1}{\text{S\&P Return} + 1}$$

An abnormal return above 1.00 can be interpreted as IPOs outperforming the S&P 500. Conversely, an abnormal return of less than 1.00 indicates that the IPO was outperformed by the S&P 500.

Out of the 1 250 companies in the dataset, historical stock price data is available for 871, due to for instance mergers and acquisitions, or companies going private. After investigating the most extreme observations of returns, we manually change the data points that are incorrect and we find the correct data. In other cases, for instance when the data from Yahoo Finance is sporadically missing or inconsistent, we remove the data point from the dataset.

Like we did for stock splits, we also perform a fuzzy string matching exercise in order to avoid mismatching of companies and stock returns, leading to removal of 32 observations. Table A1.2 in appendix A1 shows examples of fuzzy string matching results from this exercise.

Finally, we manually remove observations where key data is missing or obviously incorrect. Examples of obviously incorrect data include IPO dates and founding dates that are set to the future, and negative amounts of pre-IPO funding. This exercise brings the number of

observations down to 644. Because we intend to only examine companies that have raised external capital in the private market before going public, we remove companies without any information on this. This leaves us with a dataset consisting of 180 companies, which we will use for our analyses.

Selection Bias

For our analyses to be reliable, it is essential that the observations in our dataset are representative for the group we intend to analyse, namely venture-backed companies that have gone public. Our data is collected from sources that are widely considered as reliable, and we gathered all IPOs for the period we are investigating. The dataset we started with should therefore be representative. Throughout the data cleaning and compilation process, however, the number of companies in our dataset has been massively reduced. The boiled-down dataset could be subject to selection bias, caused for instance by higher degree of data availability for high-profile companies. Nevertheless, filtrations are unavoidable in order to carry out the desired analyses.

4 Descriptive Analysis

In this chapter, we will present descriptive analyses of our compiled and cleaned dataset. After providing a general overview of the characteristics of the dataset, we will examine distributions of the data points which will later be used for statistical analyses. We will also plot selected independent variables against post-IPO returns, to see whether there are any clear patterns that can be seen with the naked eye.

Note that whereas we calculated abnormal returns to be used as target variables in the statistical analyses, we use simple stock returns for the descriptive analyses. As these analyses are meant to be only indicative, simple returns are preferred because they are easier to relate to and compare. Simple returns for the two time frames are calculated as follows:

$$\textit{First-Day Simple Return} = \frac{\textit{Closing Price}_t}{\textit{Offer Price}} - 1$$

$$\textit{Three-Month Simple Return} = \frac{\textit{Closing Price}_{t+62}}{\textit{Closing Price}_t} - 1$$

General Descriptive Statistics

As already mentioned, we have data on 180 venture-backed IPOs, dating back to 2011. The companies raised a median of \$111.2 million in their offerings, and raised \$140.0 million in pre-IPO funding. The median age of the companies when they filed for IPO was 10.9 years, 22.8% of the companies in the dataset were profitable the year before their IPO, and 31.1% were tech companies. Table 4.1 displays annual averages of these metrics, where average offer sizes and pre-IPO funding amounts are adjusted for inflation. There are no clear historical trends, perhaps with the exception of profitability the year before IPO, although the lowest observations are in the middle of the time frame. This is likely as a result of the narrow time frame in scope, as previous research has shown that companies going public today are in general older, larger and to a lesser extent profitable compared to companies that went public a few decades ago (Ritter, 2018).

Table 4.1: Descriptive Statistics, by Year

Year	Number of IPOs	Median Offering Size (mUSD)	Median Pre-IPO Funding (mUSD)	Median Age at IPO (Years)	% Profitable Year Before IPO	% Tech Companies
2011	7	152.4	366.5	12.3	42.9%	42.9%
2012	7	82.0	133.2	7.6	57.1%	28.6%
2013	33	112.6	147.5	8.4	36.4%	18.2%
2014	45	106.1	113.4	9.3	22.4%	31.1%
2015	21	107.9	112.3	9.5	4.8%	33.3%
2016	12	95.0	146.3	8.4	8.3%	25.0%
2017	22	111.3	122.4	10.7	22.7%	45.5%
2018	18	177.3	174.5	8.5	11.1%	33.3%
2019	15	175.8	1 166.2	9.3	13.3%	33.3%
Total	180	111.2	140.0	10.0	22.8%	31.1%

With regards to potential selection bias, these figures are roughly in line with previous statistics. Ritter (2018) examined 834 VC-backed IPOs from between 2011 and 2017 and found the median age to be between 10 and 12 for each of these years. Similarly to our dataset, he also found the percentage of companies that were profitable when going public to vary significantly, ranging from 16% to 64%. A likely explanation for why his numbers are higher than ours is that Ritter measured profitability the year of IPO rather than the year before, giving the companies more time to achieve profitability. Concerning offering size, Renaissance Capital³, a company providing pre-IPO institutional research, reported in 2018 annual median offer sizes of US IPOs to be slightly lower than those in our dataset. Contrary to us, however, Renaissance Capital do not single out VC-backed IPOs. Before removing non-backed IPOs from our dataset, we saw that they were on average smaller than backed IPOs. For tech status and pre-IPO funding, we have not found any data that could be used to validate our dataset against selection bias. Considering that the other attributes are aligned with comparable datasets, there is nonetheless reason to believe that our dataset is free of selection bias.

³Not to be confused with Renaissance Technologies, the hedge fund

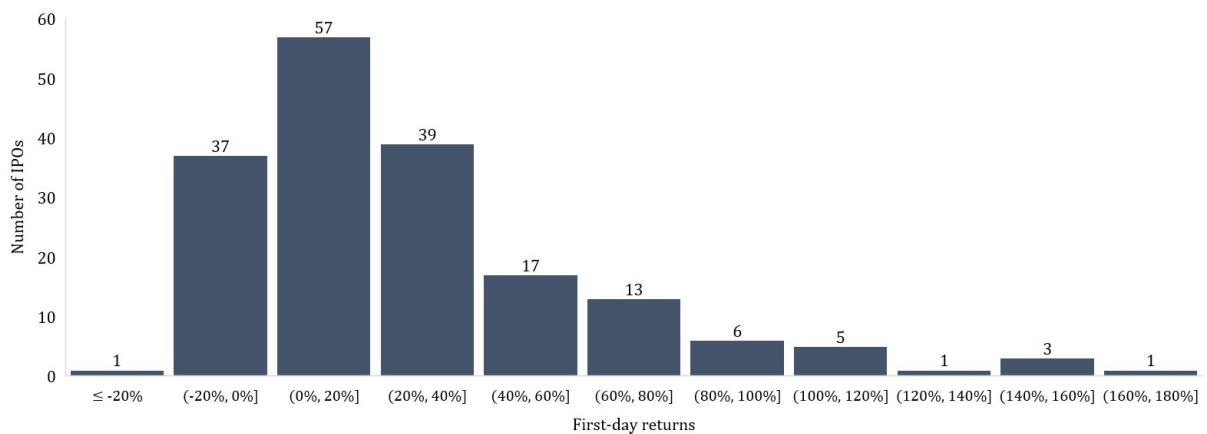
Returns

Calculated from the offer price, the median return for the IPOs in the dataset (table 4.2) was 17.7% on the first day of trading, and 29.7% after three months. However, when the first day of trading is excluded, the median return is lower, 7.4%. Although weak, the trend seems to be that average first-day return has increased over the observed period.

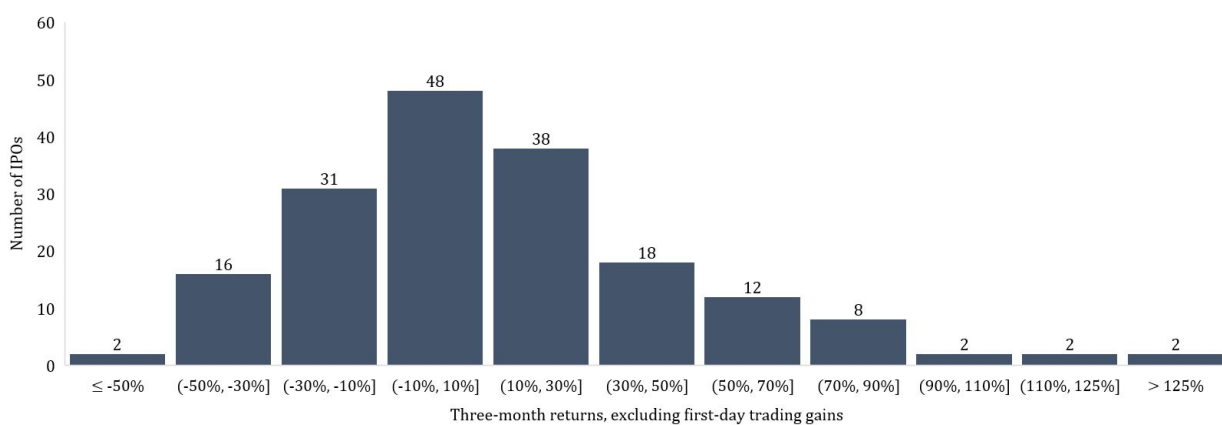
Table 4.2: Average Simple Returns

Year	1 Day	3 Months, Including First Day of Trading	3 Months, Excluding First Day of Trading
2011	14.8%	13.8%	2.5%
2012	8.4%	24.1%	12.5%
2013	28.6%	53.9%	21.0%
2014	27.8%	42.5%	14.1%
2015	32.1%	18.7%	-9.2%
2016	30.7%	77.7%	35.2%
2017	17.7%	42.3%	18.2%
2018	33.0%	38.1%	2.8%
2019	44.6%	87.0%	16.0%
Total	28.2%	45.1%	12.8%

Figure 4.1 shows the distribution of first-day trading gains of the IPOs in our dataset. It is evident that most companies filing for IPO experience an appreciation in their share price compared to the offer price of the IPO. This is in line with the phenomenon of underpricing, as discussed in the literature review. On the negative end, the most extreme observation in the dataset is Chegg Inc., who experienced a share price decline of 23% on their first day of trading. On the positive of the scale is Beyond Meat, the producer of plant-based meat substitutes whose share price rose by 163% on their first day of trading in May 2019. This was the best first-day performance of a US IPO larger than \$200 million since the dot-com boom (Murphy, 2019).

Figure 4.1: Distribution of First-Day Simple Returns

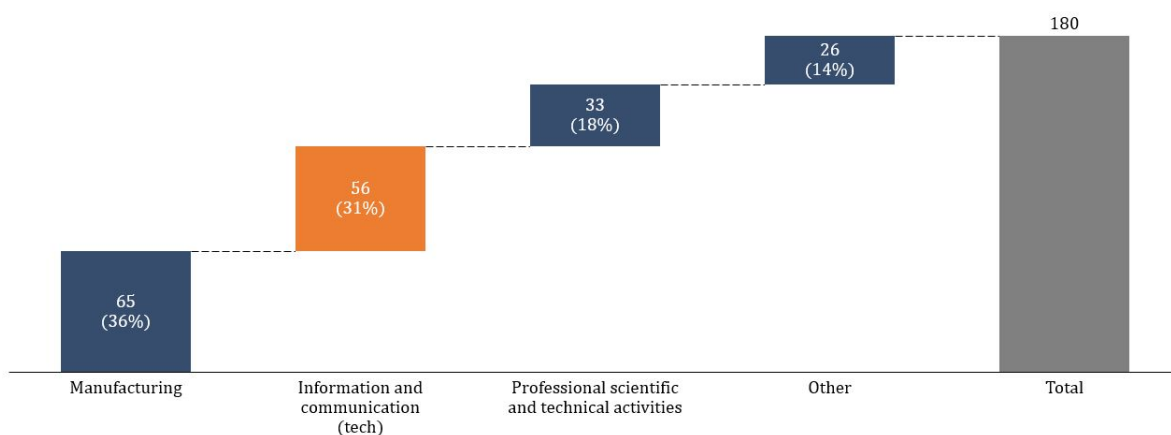
When looking at the three-month horizon, the distribution is more spread out across the axis, as illustrated in figure 4.2. Also here, Beyond Meat is on the far right. Even after their remarkable price increase at the day of IPO, the Beyond Meat stock continued to rise, tripling in value over the next three months of trading - totalling a value increase of 686.0% from their offer price. At the other end of the axis is Castlight Health, a San Francisco-based healthcare navigation company who experienced a stock price depreciation of 57.6% in the three months following their IPO in March 2014.

Figure 4.2: Distribution of Simple Returns Three Months After IPO

Industries

By using the EU industry standards classification system (NACE) code of each company, we can investigate whether there are patterns to be found when dividing companies by industry. NACE divides companies into 21 broad sections. The top three most represented sections in our dataset are Manufacturing (65 IPOs), Information and Communication (56 IPOs), and Professional Scientific and Technical Activities (33 IPOs), leaving 26 IPOs in other sections. This distribution is visualized in figure 4.3.

Figure 4.3: Distribution of Dataset



Within the Manufacturing section, 32 out of 65 companies are manufacturers of pharmaceutical preparations (NACE code 21.20), 11 are manufacturers of medical and dental instruments and supplies (code 32.50), and 9 are manufacturers of electronic components (code 26.11).

Within the Information and Communication section, 33 out of 56 companies are classified as “publishers of other software” (code 58.29), 11 belong to “Other information technology and computer service activities” (code 62.09), and 9 belong to “Computer programming activities” (code 62.01). This section includes heavily funded companies such as Uber, Snap, Facebook, Pinterest and Twitter, all among the top 10 most funded IPOs in our dataset. For the sake of convenience, companies belonging to the Information and Communication section will be referred to as “tech companies”, and other companies will be referred to as “non-tech companies”.

Within the Professional Scientific and Technical Activities section, 28 out of 33 companies operate within research and experimental development on biotechnology (code 72.11).

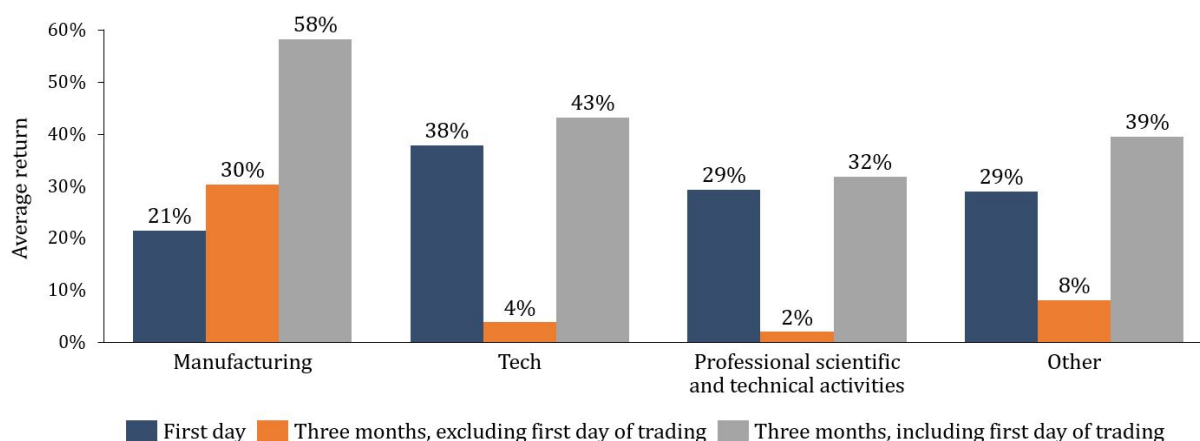
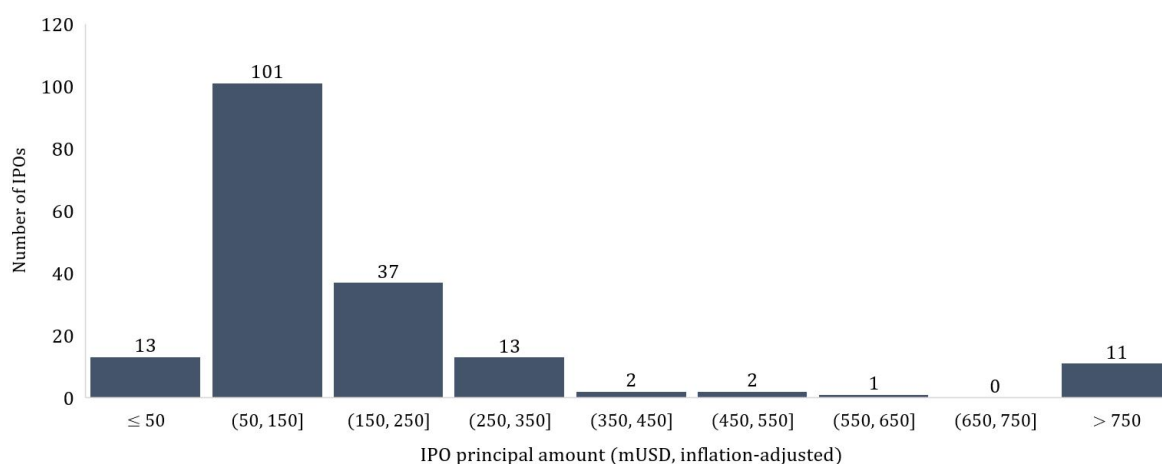
Figure 4.4: Average Simple Returns, by ISIN Industry Section

Figure 4.4 shows average returns for the three sections described above, as well as for the 26 companies in other sections. Apart from Manufacturing companies, average return is higher on the first day of trading than in the following three months for all industries. Tech companies experience the highest first-day returns, and nearly the lowest returns in the following three months, only “beaten” by Professional Scientific and Technical activities. Despite their slow start, manufacturing companies’ high returns in the following months make them outperform the other industries when looking at three-month return.

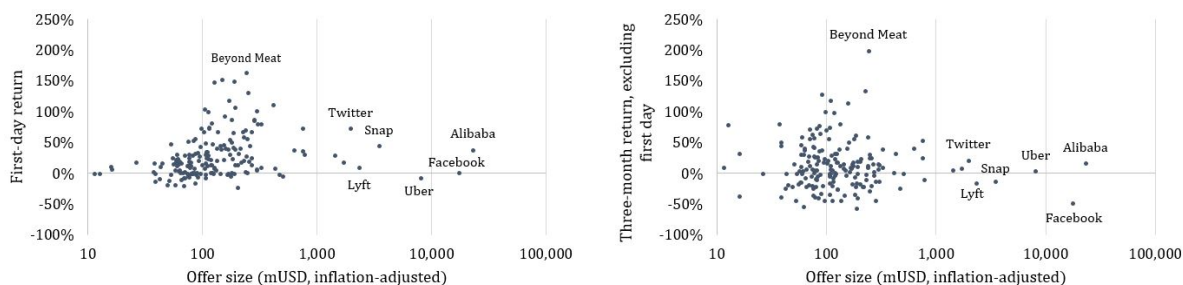
Offer Size

When a company is filing for an initial public offering, they offer a number of shares to the public for a given price per share. As mentioned in the theory section, the *offer size* of an IPO is quite simply the number of shares offered to the market multiplied by the offer price per share.

Figure 4.5 displays the distribution of inflation-adjusted offer sizes in our dataset. The majority of the observations raised between \$50 and \$150 million when they went public, and only 13 companies raised less than \$50 million. This aligns with the median offer size of US IPOs in 2014, which according to Berk and DeMarzo (2017) was \$100 million. By significant margin, the largest IPO in the dataset are Alibaba and Facebook, raising \$23.1 and \$17.5 billion, respectively. Adding Uber’s \$8.1 billion IPO, the top three raised more than the other 177 companies in the dataset combined - \$48.7 vs. \$36.4 billion.

Figure 4.5: Distribution of Offer Sizes

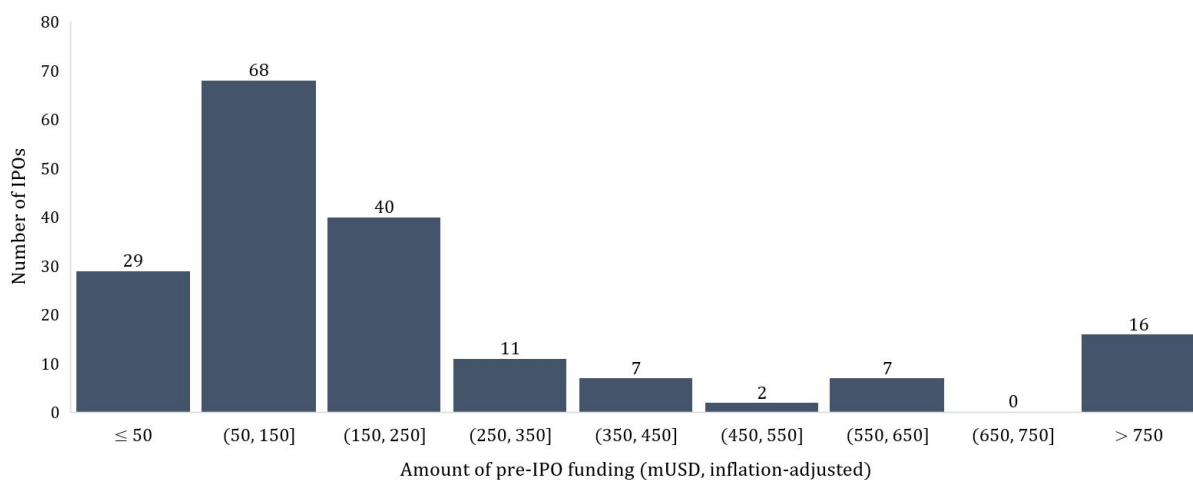
In figure 4.6, offer size is plotted against first-day and three-month returns, respectively. In order to adjust for skewness in offer sizes, a base-10 log scale is used for the horizontal axis. There seems to be a slight correlation between first-day return and offer size, although some of the largest IPOs in the dataset achieved modest or even negative returns on their first day of trading.

Figure 4.6: Plotting Simple Returns Against Offer Size

For three-month returns, on the other hand, it is not possible to identify any clear pattern when plotting against offer size.

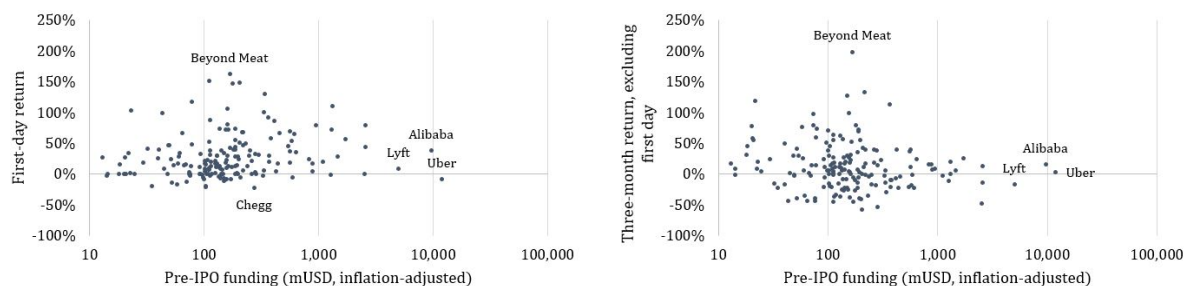
Amount of Pre-IPO Funding

Figure 4.7 shows the distribution of pre-IPO funding amount. The shape of this histogram is fairly similar to the one displaying offer size distribution, likely indicating a covariance between offer size and pre-IPO funding amount.

Figure 4.7: Distribution of Pre-IPO Funding Amount

Slightly over a third of the companies in the dataset raised between \$50 and \$150 million in pre-IPO funding, and three quarters raised less than \$250 million. 16 companies raised more than \$750 million, and the range within this group is enormous: the three companies that raised the most funding were Uber (\$11.9 billion), Alibaba (\$9.6 billion) and Lyft (\$5.0 billion). Their combined pre-IPO funding amount, \$26.5 billion, constitutes 38% of all the funding amounts in our dataset added together.

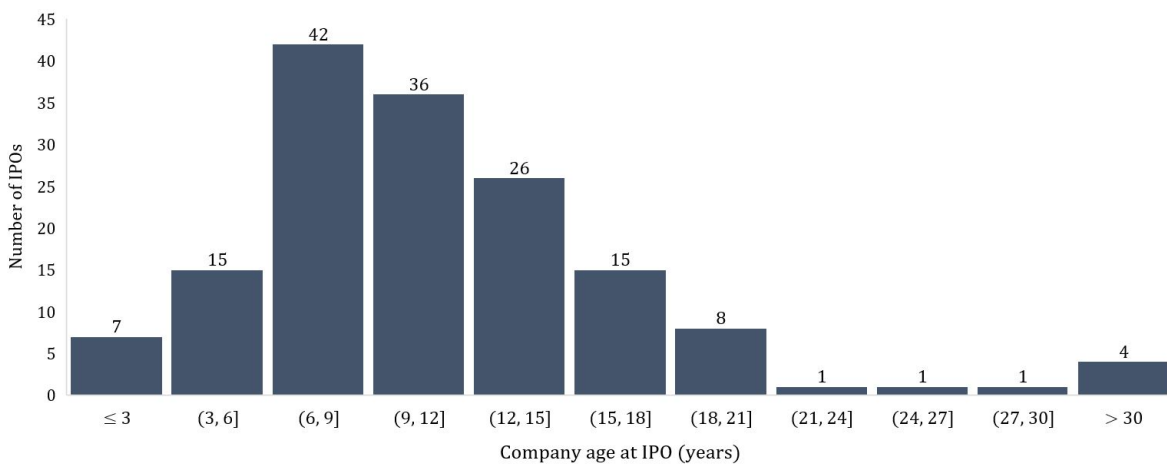
When pre-IPO funding amounts are plotted against returns, as illustrated in figure 4.8, it is difficult to spot any clear trends apart from a more scattered distribution of returns with the longer time horizon. As for offer size, a base-10 log scale is used for the horizontal axis.

Figure 4.8: Plotting Simple Returns Against Pre-IPO Funding Amount

Company Age at IPO

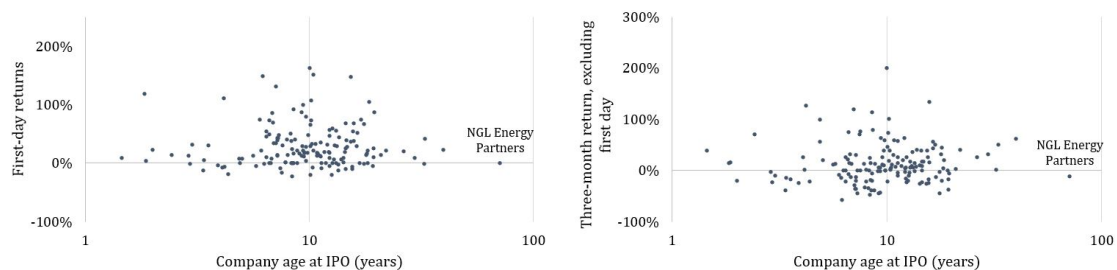
Figure 4.9 shows distribution of company age at IPO. The younger end of the scale is filled predominantly with pharmaceutical companies, with Spark Therapeutics as the youngest company going public in our dataset. Spark develops gene therapies, and filed for IPO in January 2015, only two years after the company's inception in 2013. The oldest company for which we were able to find reliable data was NGL Energy Partners, who waited 71 years before going public in 2011.

Figure 4.9: Distribution of Company Age at IPO



Similarly to earlier, it is difficult to spot any clear patterns when age is plotted against returns, as displayed in figure 4.10. Thus, a statistically significant linear correlation between returns and company age at IPO is unlikely. As earlier, a base-10 log scale is used for the horizontal axis.

Figure 4.10: Plotting Simple Returns Against Company Age at IPO



Profit Margin Year Before IPO

The most recent year for which Orbis provides accounting data is 2018. In order to include IPOs from 2019, we will therefore use profit margin the year before IPO as the measurement of profitability for our analyses. Thus, when discussing “unprofitable companies”, we will refer to companies with negative profit margin the year before IPO.

As mentioned earlier, one of the main motives of an IPO has traditionally been to raise capital in order to finance growth (Berk and DeMarzo, 2017). Driven by the increased access to private growth capital provided by venture capitalists, companies today typically wait longer before they file for IPO (Døskeland and Strömberg, 2018). Thus, one would expect that companies going public today are more mature than they used to be. However, this does not appear to be reflected when looking at the profitability of companies filing for IPO in recent years. In fact, according to Rooney (2019), Goldman Sachs are forecasting that less than a quarter of the companies going public in 2019 will report positive net income this year, which is the lowest level since the tech bubble of the early 2000s. A similar trend can be spotted in our dataset, as illustrated in figure 4.11.

Figure 4.11: % of Venture-Backed IPOs with Positive Profit Margin year before IPO

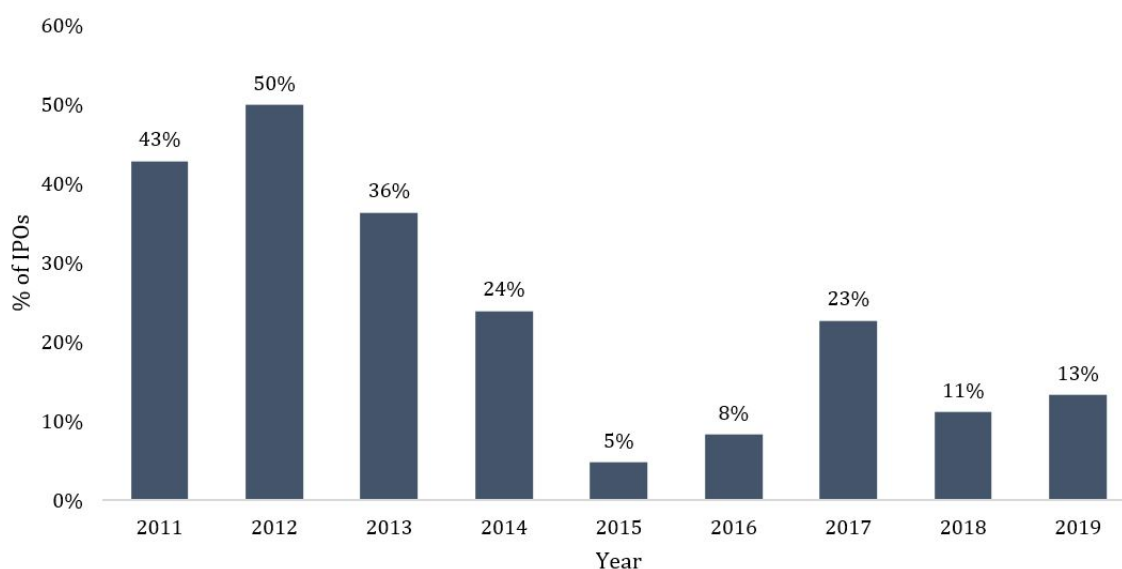
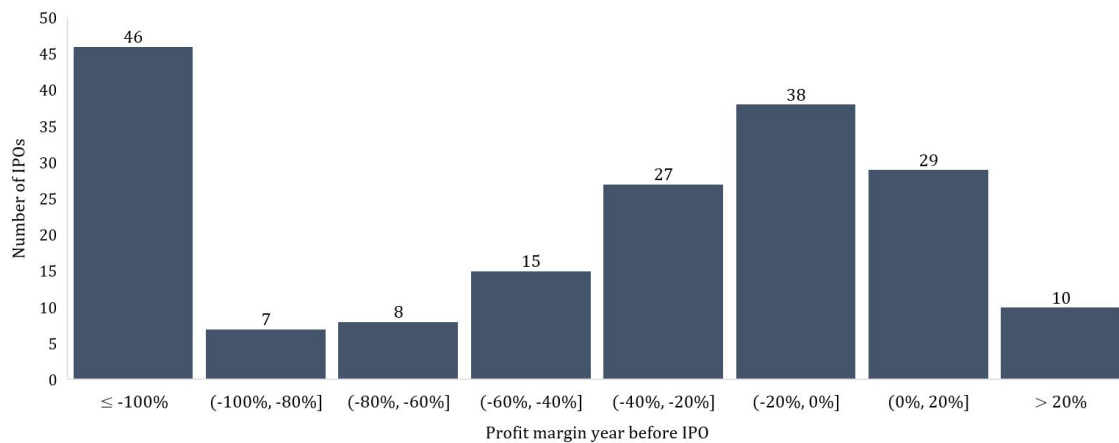
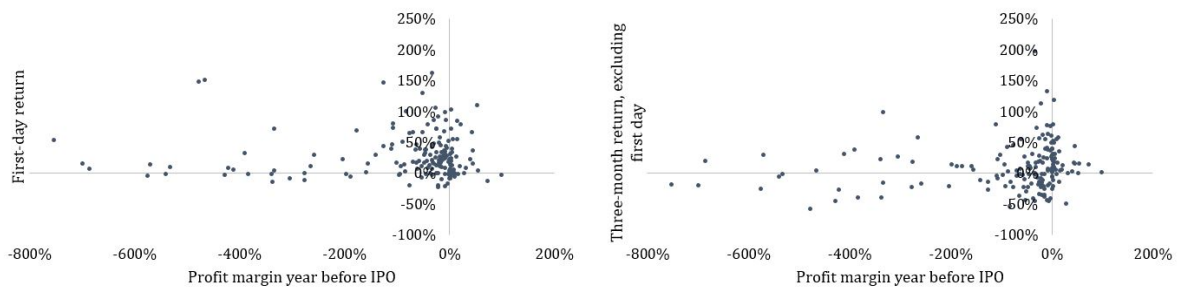


Figure 4.12 shows the distribution of profit margins the year before IPO. It is evident from the figure that most of the companies in our dataset were unprofitable the year before they filed for IPO, as already mentioned in the introduction. Out of the 180 companies, only 39 (22.8%) reported a positive bottom line the year before their IPO.

Figure 4.12: Distribution of Profit Margin Year Before IPO

Moreover, it stands out that 46 companies (25.6%) had profit margins lower than -100%. 40 of these companies are either biotechnology or pharmaceutical companies, both typically research and development-heavy industries.

Figure 4.13 displays returns plotted against profit margin year before IPO, excluding the nine companies with profit margins below -800%. As earlier, there are no clear patterns suggesting a linear correlation between returns and profit margin year before IPO.

Figure 4.13: Plotting Simple Returns Against Profit Margin Year Before IPO

5 Analyzing Individual Attributes

While we introduced and explored the dataset in the previous chapter, more targeted analysis will be done in this chapter. The aim is to answer our first two hypotheses:

Hypothesis 1: status as tech company, amount of pre-IPO funding, IPO size, and negative profitability are all negatively correlated with first-day return

Hypothesis 2: status as tech company, amount of pre-IPO funding, IPO size, and negative profitability are all negatively correlated with three-month return, excluding first-day trading gains

In order to successfully answer our hypotheses, we will make use of both non-statistical and statistical analysis. Whereas the former will be used to visually search for patterns in the performance of companies holding each of the attributes, the latter will be used to examine whether or not the patterns are statistically significant.

The non-statistical analysis will be executed by dividing the dataset into groups based on the four characteristics, which will be compared in order to see if there are any clear patterns. Similarly to for the descriptive statistics, we will use simple returns for these analyses.

To determine the statistical significance of the patterns, multiple regression analysis will be used to investigate whether there are linear correlations between post-IPO stock performance and the independent variables in the dataset. As stated in the data chapter, we will use abnormal returns for these analyses.

In this chapter, each of the four characteristics in scope will be analyzed independently, as outlined above. The next subchapter explains how the regression analysis was carried out, before proceeding with analysis results in the subsequent subchapter.

5.1 Method: Multiple Regression Analysis

To investigate whether there are statistically significant relationships between post-IPO stock performance and the attributes, multiple regressions and the ordinary least squares (OLS) method will be used. This method chooses the estimates that minimize the sum of the squared residuals to provide as precise estimates as possible (Wooldridge, 2012).

5.1.1 Model Formulations

In our regressions, the first-day and three-month returns serve as the dependent variables. In addition to the attributes that form part of our hypothesis, which are the independent variables, we add certain data points from our dataset to serve as control variables. Following, the regression models are formulated, before the different variables will be further explained in the next sub-chapter.

Regression Model for the First Trading Day

$$\begin{aligned} \log abnormalreturn1 &= \beta_0 + \beta_1 \log totalraised + \beta_2 \log offer size + \\ &\beta_3 \log number of employees + \beta_4 \log age at ipo + \beta_5 D_{profitnegative} + \beta_6 D_{over250} + \beta_7 D_{tech} + \varepsilon \end{aligned}$$

Regression Model for the First Three Months of Trading

$$\begin{aligned} \log abnormalreturn90 &= \beta_0 + \beta_1 \log totalraised + \beta_2 \log offer size + \\ &\beta_3 \log number of employees + \beta_4 \log age at ipo + \beta_5 D_{profitnegative} + \beta_6 D_{over250} + \beta_7 D_{tech} + \varepsilon \end{aligned}$$

5.1.2 Explanation of Model Variables

In this section, we will provide brief explanations of the variables used in our regression models. Most of the continuous variables are log-transformed to improve linearity. Further details concerning this can be found in appendix A2.

Dependent Variables

First-day abnormal return: *logabnormalreturn1*

The dependent variable *logabnormalreturn1* is the log-transformed abnormal return of the first day of trading. The S&P 500 index is used as a benchmark to calculate the abnormal return.

Three-month abnormal return: *logabnormalreturn90*

The dependent variable *logabnormalreturn90* is the log-transformed abnormal return for the first three months of trading, excluding first-day trading gains. The value is calculated by adjusting the return to the S&P 500 index for the same time period.

Independent Variables**Adjusted offer size: *logoffersize***

The offer size is the log-transformation of number of shares offered multiplied with the price per share. As our dataset contains information on IPOs from 2011 and onwards, the offer size has been adjusted with the US Consumer Price Index to make sure the data is comparable.

Adjusted total raised before IPO: *logtotalraised*

The value of the variable is the log-transformed total amount of funding the company has raised before their IPO. The different funding rounds are adjusted with the US Consumer Price Index from the time the funding was raised.

Company age at IPO: *logageatipo*

The company's age at IPO is found by subtracting the year when the company was founded from the year they went public. As the age of a few companies were significantly higher than most other companies, it is log-transformed.

Number of employees the year before IPO: *lognumberofemployees*

The variable is the log-transformed number of employees the company had the year before they went public.

Dummy variable for tech companies: *Dtech*

The dummy variable has a value of one when the company is classified as a tech company, and a value of zero for non-tech companies. As explained earlier, "tech companies" are defined by using the EU industry standards classification system (NACE) code of each company, where companies classified as "Information and communication" are set to be "tech companies".

Dummy variable for negativ profit margins the year before IPO: Dprofitnegative

The dataset contains the profit margins for each company the year before they went public. The dummy holds a value of one if the company had a negative profit margin, and a value of zero if they were profitable. As stated in previous chapters, we will simply refer to these companies as “profitable” and “unprofitable” companies.

Dummy variable for companies that raised over \$250 million before IPO: Dover250

Instead of only investigating whether there is a linear relationship between the amount of funding each company have raised before IPO and their returns post-IPO, we will investigate if the most funded companies perform differently than others. The dummy variables will therefore hold a value of one if the company raised \$250 million or more. These companies will be referred to as “heavily funded companies”. The dummy variable will hold a value of zero for all other companies.

5.1.3 OLS Violations

There are certain assumptions that must hold true for the regression models to be accurate (Wooldridge, 2014). The main assumptions for the ordinary least squares method are:

1. Linearity in the parameters
2. Random sampling
3. No perfect collinearity
4. Zero conditional mean
5. Homoscedasticity

To determine the validity of our analysis, violations of these assumptions have been investigated. Variables are log-transformed to improve the linear relationship between the dependent and the independent variables, and the variance inflation indicators and a correlation matrix have been calculated to reveal potential multicollinearity. Further, the RESET test and the White test are run to make sure the zero conditional mean assumption is valid and to identify potential homoscedasticity. The tests show no sign of violations

of these OLS assumptions. However, as discussed in chapter 3 and 4, the assumption of random sampling is harder to confirm, as our data cleaning might have biased our data. Still, through thorough assessment of the dataset, we feel comfortable assuming that the dataset is a representative sample of the whole population of venture-backed IPOs.

Consequently, we assume that the dataset does not violate any of the OLS assumptions. The Gauss-Markov theorem is therefore valid, and our OLS estimators are the Best Linear Unbiased Estimators (BLUEs) for the model (Wooldridge, 2012). Moreover, normally distributed unobserved error terms are preferred for exact statistical inference (Wooldridge, 2012). We have therefore analysed the residuals, as well as the distribution of the dependent variables. With as many as 180 observations, the central limit theorem concludes that the OLS estimators satisfy asymptotic normality, meaning we can assume we have an approximate normal distribution (Wooldridge, 2012).

Further details concerning model validation can be found in appendix A2.

5.1.4 Regression Results

Table 5.1 and table 5.2 show the results of our regression analysis, which will be explained and discussed in subchapter 5.2.

Table 5.1: Regression Results Against Abnormal First-Day Return

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
logtotal-raised	0.0255*	0.0119	0.0145					0.0226*		
	(0.051)	(0.405)	(0.480)					(0.079)		
logoffer-size		0.0416**	0.0264	0.0494*	0.0533***	0.0381**	0.0461**		0.0448**	0.0564***
		(0.030)	(0.380)	(0.070)	(0.009)	(0.030)	(0.010)		(0.025)	(0.001)
lognumberof-employees			-0.005	-0.008						
			(0.786)	(0.610)						
logageatipo			-0.054		-0.026					
			(0.127)		(0.379)					
profitnegative				0.130**	0.0841*		0.0834*			0.0952*
				(0.014)	(0.070)		(0.061)			(0.033)
over250				0.0008	0.0144				0.0185	
				(0.990)	(0.774)				(0.714)	
techyn						0.0946**	0.0850**	0.111***		
						(0.020)	(0.036)	(0.005)		
_cons	0.085	-0.053	0.164	-0.086	-0.062	-0.007	-0.109	0.066	-0.015	-0.142
	(0.195)	(0.558)	(0.216)	(0.506)	(0.640)	(0.933)	(0.280)	(0.311)	(0.871)	(0.156)
N	180	180	124	124	180	180	180	180	180	180
R²	0.0213	0.0472	0.0502	0.0745	0.0724	0.0723	0.0907	0.0637	0.0441	0.0677

*P-values in parentheses, Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

The regression table reports the coefficients and p-values (in parentheses) from the regressions run with log-transformed abnormal first day returns as the dependent variable.

Table 5.2: Regression Results Against Abnormal Three-Month Return, Excluding the First Day of Trading

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
logtotal-raised	-0.027 (0.116)	-0.022 (0.236)	-0.038 (0.174)					-0.024 (0.151)		
logoffer-size		-0.013 (0.597)	0.008 (0.849)	-0.015 (0.690)	-0.024 (0.366)	-0.017 (0.471)	-0.025 (0.293)		-0.016 (0.555)	-0.034 (0.137)
lognumberof-employees			-0.004 (0.864)	0.006 (0.789)						
logageatipo			0.052 (0.279)		0.022 (0.575)					
profitnegative				-0.101 (0.158)	-0.087 (0.158)		-0.086 (0.146)			-0.097 (0.101)
over250				-0.096 (0.246)	-0.050 (0.455)				-0.054 (0.422)	
techyn						-0.086 (0.112)	-0.076 (0.161)	-0.09* (0.086)		
_cons	0.178** (0.038)	0.222* (0.064)	0.088 (0.617)	0.176 (0.317)	0.196 (0.266)	0.157 (0.167)	0.262* (0.052)	0.194** (0.024)	0.137 (0.269)	0.293** (0.029)
N	180	180	124	124	180	180	180	180	180	180
R²	0.014	0.015	0.032	0.034	0.028	0.022	0.033	0.030	0.011	0.022

*P-values in parentheses, Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$*

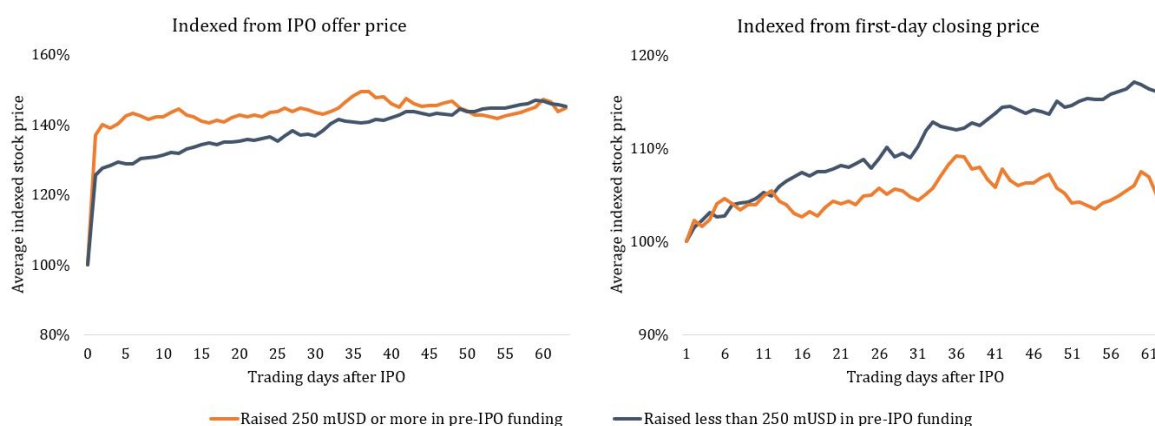
The regression table reports the coefficients and p-values (in parentheses) from the regressions run with log-transformed abnormal three-month return as the dependent variable.

5.2 Findings

5.2.1 Amount of Pre-IPO Funding

In the descriptive analysis, we saw few signs of a linear relationship between pre-IPO funding and post-IPO stock performance. In order to investigate this relation in a slightly different way, the companies in the dataset are divided into two groups, based on whether or not they raised more than \$250 million in pre-IPO funding. This specific amount is chosen in order to separate the top quartile from the rest of the dataset, rounded to the nearest \$50 million. There are 43 companies in the group that raised \$250 million or more in pre-IPO funding. As already mentioned, these will for the sake of convenience be referred to as “heavily funded companies”.

Figure 5.1 displays post-IPO stock performance for companies that raised external funding before their IPO.

Figure 5.1: Indexed Average Returns, Much vs. Not-So-Much Funding

The graph on the left displays indexed average post-IPO stock performance with the IPO offer price as starting point. It is evident that both groups on average are underpriced, and that heavily funded IPOs were on average more underpriced than other IPOs. However, over the following three months, the heavily funded companies were outperformed by the other companies. This is particularly evident when first-day gains are excluded, as illustrated in the graph on the right. Although both groups appreciated in value, the average nominal return for companies that raised less than \$250 million is approximately three times those of companies that raised more than \$250 million (15% vs. 5%).

If the part of hypothesis 1 about pre-IPO funding and its effect on underpricing is true, we would expect the variables in the regression models related to funding to be significantly positive. However, the regression results, displayed in table 5.1, do not confirm our hypothesis. The continuous variable “logtotalraised” would be significant if there was a linear relationship between the amount of funding raised and the first-day return. Although two out of four regressions show significant results for this variable, it is only at a 10% significance level, and the p-values for the two remaining coefficients are far from significant. It is also worth noticing that the variable is only significant when it is not combined with the significant variable for offer size. It is therefore reason to believe that the significant results for the “logtotalraised” variable in regression (1) and (8) hold information of the offer size and hence provide biased results.

Even though a linear relationship between the first-day return and the amount of pre-IPO funding seems unlikely, the most funded companies might still perform differently than others on their first day of trading. To investigate this relationship, the regressions include

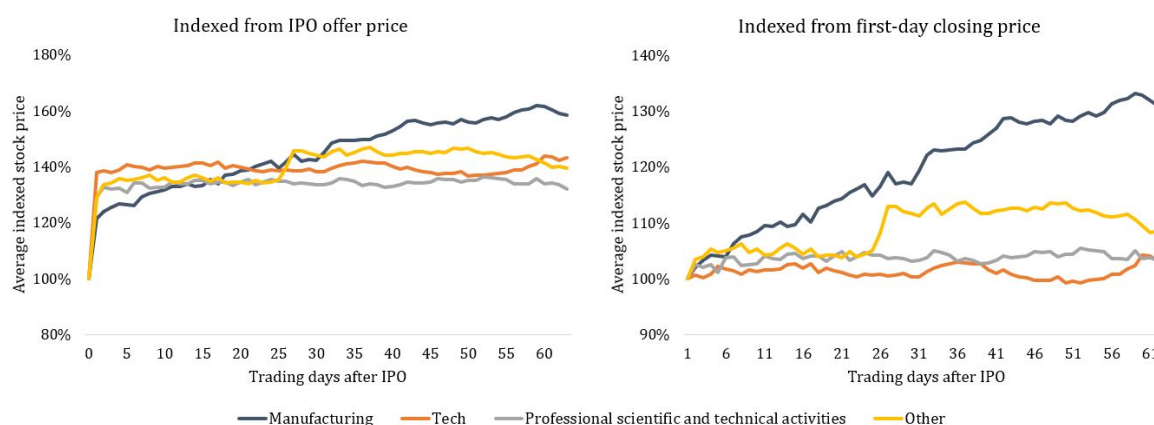
a dummy variable for heavily funded companies. As our hypotheses are partially motivated by the performance of the aforementioned unicorns that went public in 2019, and that the increased access to private funding might have affected their performance, we would expect the dummy variable to be significant. However, the regressions reveal no significant differences in first-day return for the companies in the “over250”-category. Based on these results, the most funded companies do not seem to be more underpriced at IPO than other companies, and our hypothesis seems to be rejected.

For the three-month time horizon, one of our hypotheses is that pre-IPO funding and abnormal returns are negatively correlated. Thus, we expect to find that the coefficients for the explanatory variables related to pre-IPO funding have negative prefixes and are statistically significant. Although the regression results, as displayed in table 5.2, imply a negative relationship for both the continuous and the binary variable, the correlations are not statistically significant. Thus, this hypothesis can be rejected.

5.2.2 Tech Companies vs. Non-Tech Companies

As described in the descriptive analysis chapter, the companies can be divided into industry categories by using the EU NACE code. Figure 5.2 shows average indexed returns for the three largest categories, as well as the rest of the companies in the dataset.

Figure 5.2: Indexed Average Returns, by Industry

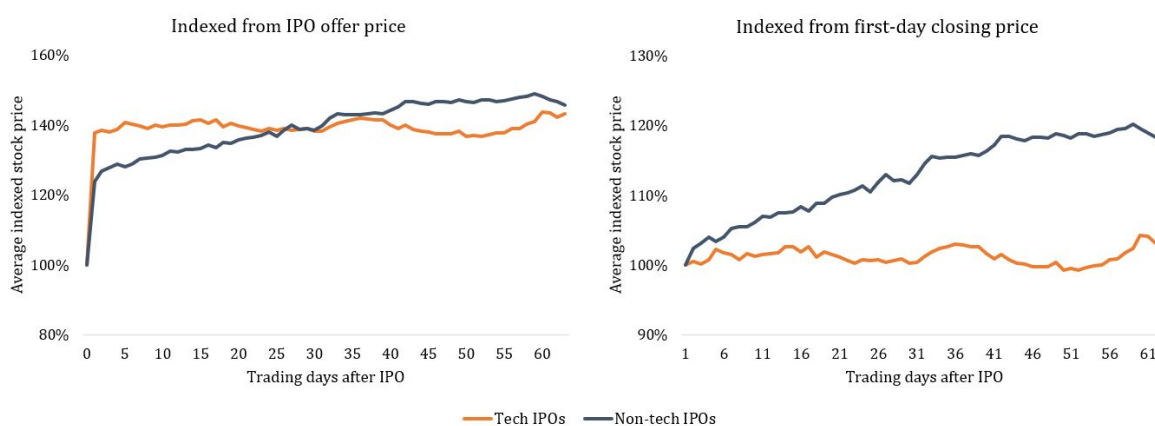


It is evident that the tech companies were more underpriced than other companies in our dataset. Over the first three months after the IPO however, manufacturing companies on average outperform other companies. This applies regardless of whether or not gains from the first day of trading are included, but is particularly evident when they are

excluded. From the right graph in figure 5.2, it also appears that tech companies and companies within professional scientific and technical activities on average experienced very modest returns in the first three months following their IPO, on average 3.8% and 2.0%, respectively, when first day gains are excluded.

When instead dividing into “tech” and “non-tech” companies, the picture is very similar to for pre-IPO funding: tech companies are more underpriced than other companies, but perform worse over the next three months (figure 5.3). After 3 months, and excluding first-day trading gains, non-tech companies delivered average returns of 17.7%, almost five times as high as tech companies (3.8%).

Figure 5.3: Indexed Average Returns, Tech vs. Non-Tech IPOs



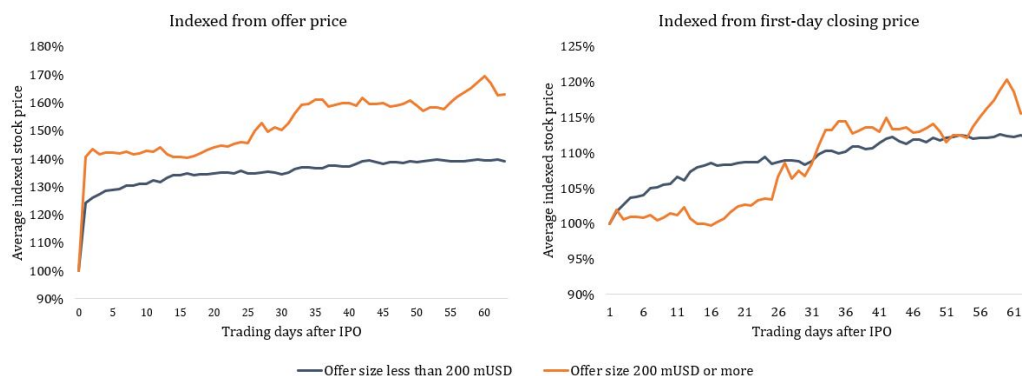
In the multiple regression results (table 5.1), the tech label seems to have a significant positive impact on the first-day return. This is consistent throughout all three regressions, and the results show that if a company is categorized as a tech company, first-day abnormal return increases with an average of about 9.7%. The scenario of “fads” may be the case for these companies, as the substantial underpricing can be a sign of investors being overoptimistic about future potential of companies in the tech industry.

As for the amount of pre-IPO funding, we also expect to find a significant negative relationship between three-month returns and the tech label. Although the covariation is consistently negative, the correlation is only significant in one of the three models, and on a 10% level. This is also the only statistically significant result of the regressions using three-month returns as the target variable. Thus, we cannot unambiguously conclude that tech companies perform worse than other companies on the three-month horizon.

5.2.3 Offer Size

Like for amount of pre-IPO funding, we separate the dataset into two groups based on whether or not companies are among the top 25% with regards to size of their IPO. This split is done at \$200 million, and average indexed stock price development for these two groups are illustrated in figure 5.4.

Figure 5.4: Indexed Average Returns, Large vs. Not-So-Large IPOs



As expected, IPOs larger than \$200 million are on average more underpriced than smaller IPOs. Whereas the average first-day return for smaller IPOs is 24.1%, the average for the largest IPOs is 40.7%. Contrary to the other characteristics in scope, however, the stocks of the largest IPOs in our dataset perform slightly better than others also on the three-month term, in spite of modest returns in the first month.

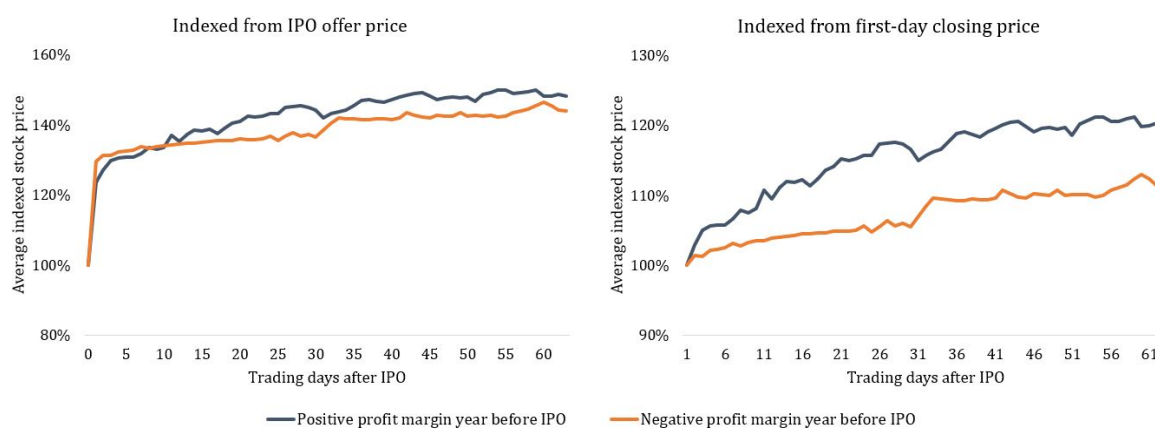
The regression analyses clearly demonstrate that offer size is correlated with first-day returns, with a significance level of 5% or less in 6 out of 8 regressions. This means that larger IPOs tend to be more underpriced than smaller IPOs. We find that for a one percent increase in offer size, the first-day abnormal return increases with an average of 4.3%.

For three-month returns, both expectations and results are similar to the previous characteristics. With one exception, the coefficient prefixes are negative, but none are statistically significant. Thus, we cannot confirm our hypothesis that larger IPOs are outperformed by smaller IPOs in the first three months of trading.

5.2.4 Profitable Companies vs. Unprofitable Companies

Similarly to before, we divide the companies in two groups, this time based on whether or not they were profitable the year before their IPO. As explained in previous paragraphs, we will refer to these companies as “profitable” or “unprofitable” companies. There are 41 profitable and 139 unprofitable companies in the dataset. This division seemingly tells a similar story to earlier, only with different main characters: the unprofitable companies were on average slightly more underpriced than profitable companies, but were clearly outperformed by the profitable companies in the following three months of trading (figure 5.5). After three months of trading, and excluding first-day trading gains, profitable companies delivered on average a return of 19.8%, whereas unprofitable companies delivered an average return of 11.2%.

Figure 5.5: Indexed Average Returns, Profitable vs. Unprofitable Companies



As discussed earlier concerning age, companies with a negative profit margin can also be considered risky investments, leaving an expectation of high first-day returns. One could therefore expect the dummy variable “profitnegative” to be relevant when analyzing first-day returns. All regressions show significant results at a 10% or 5% level for “profitnegative”, and the coefficients hover around 9%. The interpretation of our results is therefore that companies with a negative profit margin the year before IPO are on average 9% more underpriced than others.

Unsurprisingly, the coefficient indicates that unprofitable companies are outperformed by profitable companies on the three-month horizon. Equally unsurprising, this relationship is not statistically significant - although it is very close to be significant in one of the four models, where it is significant on a 10.1%-level. Nevertheless, we cannot confirm

our hypothesis that unprofitable companies going public are outperformed by profitable companies going public in the first three months of trading.

5.3 Summary and Discussion

In this chapter, we analyzed for relationships between post-IPO stock performance and four key attributes: status as tech company, amount of pre-IPO funding, IPO size, and negative profitability.

In line with our hypothesis, multiple regression results indicated that each attribute is associated with high first-day returns and low returns in the subsequent months. However, correlations are almost without exception statistically significant only for the first day of trading.

Despite the lack of statistical significance, it is worth noticing that practically all coefficients in the regressions change from positive to negative when switching from first-day to three-month returns. The interpretation of this is that factors having a positive effect on underpricing have a negative effect on three-month performance. This is consistent with Ritter's findings, as he finds that firms with high adjusted first-day returns tend to perform poorly in the aftermarket (Ritter, 1991).

While most of our hypotheses are in line with previous literature, those regarding offer size are not. Ritter (1991) found that smaller companies tend to be more underpriced than larger companies. He also found that younger companies are more underpriced than older companies. Ritter argues that growth companies are considered more exposed to aftermarket risk than mature companies, and experience a higher first-day return. This is consistent with the findings of Shen and Goo (2019), that aftermarket risk and first-day returns are positively correlated. However, whereas the largest IPOs used to involve mature, stable companies, the largest IPOs of the last decade have rather involved growth companies with uncertain futures in terms of delivering lasting profits. This change might explain why the relationship between the size of the initial offer and the level of underpricing has turned.

Pre-IPO funding turned out to be the *least* influential of the four attributes. Considering the increasing access to private equity capital for unlisted companies combined with disappointing performance of recent heavily funded companies going public, we expected

to find a stronger relationship. However, the analyses indicated a stronger relationship between offer size and returns. As the companies that caught our attention - such as Uber, Lyft and Slack - are among both the most venture-backed *and* the largest IPOs, it appears that offer size might be a better proxy for hype than funding.

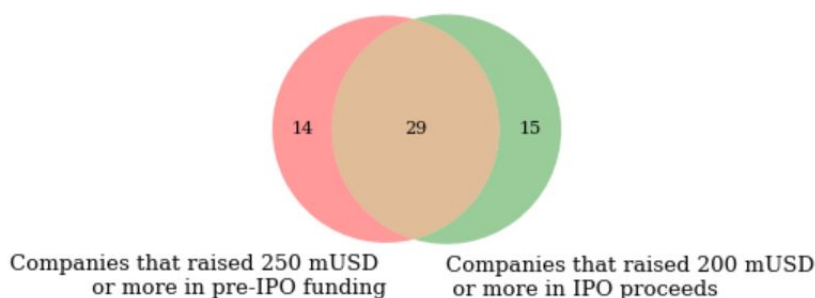
6 Analyzing Combinations of Attributes

In this chapter, we will investigate our third hypothesis:

Hypothesis 3: When combining the characteristics, their effects on both first-day return and three-month return will be even stronger than they are individually

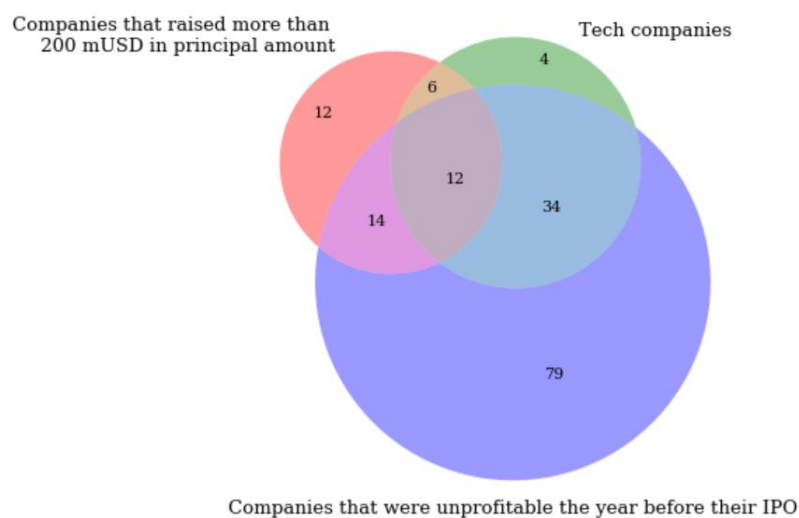
Groups of companies are formed by combining characteristics based on the significant relationships observed in the regressions. Even though the significant results were primarily found for the first-day return, we will investigate whether combinations of these attributes also have an impact on the three-month return. Despite not finding a linear relationship between offer size and the three-month return, the largest IPOs may still perform differently than others. For further analysis we will therefore investigate companies with an offer size above \$200 million, as well as tech companies and unprofitable companies. Companies with an offer size above \$200 million will be referred to as “large IPOs”. The dummy variable for heavily funded companies did not provide significant results in the regressions, and will not be included in further analysis. However, as illustrated in figure 6.1, these groups of companies are, unsurprisingly, largely the same.

Figure 6.1: Overlap Between Large IPOs and Heavily Funded Companies



6.1 Rephrasing the Hypothesis

When we split the dataset into groups based on the four attributes in the previous chapter, the pattern was always the same: companies holding the attribute had on average higher first-day returns and lower three-month returns than other companies. This raises an obvious question: are we looking at virtually the same group of companies over and over? If not, can combinations of variables yield even more conclusive results? The first question can be answered by looking at the Venn diagram in figure 6.2.

Figure 6.2: Overlapping Characteristics

Of the 180 companies in the dataset, 77% (139) were not profitable the year before their IPO. This is more or less consistent when combining with other groups: 82% of tech companies were not profitable the year before their IPO, whereas the number is slightly lower for companies with an offer size above \$200 million (59%). 12 companies in the dataset hold all three attributes.

It is therefore clear that even though a similar pattern can be observed for the individual characteristics, the companies within each category are not the same. This finding indicates that if each attribute affect the aftermarket performance in similar directions, companies holding a combination of the attributes should experience even stronger effects.

With the previous paragraphs in mind, our hypothesis for the combined variables can be rephrased as follows:

When combining the characteristics “tech”, “unprofitable” and “large IPO”, the positive effect on first-day returns and negative effect on returns in the subsequent three months will be even stronger than the effect of each characteristic individually

In this chapter, we will investigate whether this is true.

6.2 Method: Mann-Whitney U Test

Figure 6.2 presents how the relevant characteristics overlap, and thereby how groups of companies holding more than one characteristic are formed. Returns for members

of these groups will be compared with returns for non-members through a two-sample Mann-Whitney U test (also called the Wilcoxon rank-sum test).

The Mann-Whitney U test investigates whether two samples are significantly different from each other. The test does not require normally distributed observations, and can be used for samples of different sizes. As it uses the sum of the ranks of the observations, instead of the mean of each sample, outliers will not gain too much weight in the test.

To further investigate the dataset, the Mann-Whitney U test will be used to test the following groups against the rest of the dataset:

1. Tech companies with large IPOs
2. Unprofitable tech companies
3. Unprofitable companies with large IPOs
4. Unprofitable tech companies with large IPOs

In line with our hypothesis, we expect to find that all groups perform different than others, with a better first-day performance, and a worse performance in the three-month time frame.

6.2.1 Test Results

Table 6.1 and 6.2 show the results of the Mann-Whitney U tests, which will be explained and discussed in subchapter 6.3.

Table 6.1: Mann-Whitney U Test: First-Day Return

Group compared to the rest of the dataset	Significant?	Performance compared to others
Tech companies with large IPOs (18 obs)	Yes, $\alpha = 5\%$	Better in 67.1% of all cases
Unprofitable tech companies (46 obs)	Yes, $\alpha = 1\%$	Better in 69.3% of all cases
Unprofitable companies with large IPOs (26 obs)	Yes, $\alpha = 1\%$	Better in 67.2% of all cases
Unprofitable tech companies with large IPOs (12 obs)	Yes, $\alpha = 1\%$	Better in 73.3% of all cases

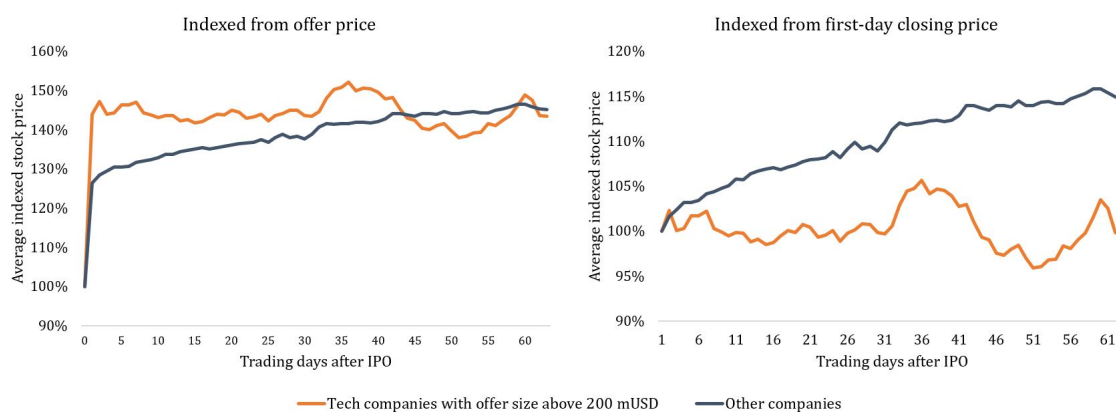
Table 6.2: Mann-Whitney U Test: Three-Month Return

Group compared to the rest of the dataset	Significant?	Performance compared to others
Tech companies with large IPOs (18 obs)	No	Worse in 59.9% of all cases
Unprofitable tech companies (46 obs)	Yes, $\alpha = 10\%$	Worse in 59.1% of all cases
Unprofitable companies with large IPOs (26 obs)	No	Worse in 56.1% of all cases
Unprofitable tech companies with large IPOs (12 obs)	No	Worse in 61.5% of all cases

6.3 Findings

6.3.1 Tech Companies With Large IPOs

Out of 56 tech IPOs in the dataset, 18 were larger than \$200 million. Figure 6.3 displays average stock development calculated from offer price and closing price after the first day of trading for this group compared to the other 162 companies in the dataset.

Figure 6.3: Indexed Average Returns, Tech Companies With Large IPOs

The patterns are similar to those observed earlier: higher first-day returns for the large tech IPOs than for other IPOs, while the large tech IPOs are outperformed in the next three months of trading. Specifically, tech companies with large IPOs had on average 17.5 percentage points higher first-day returns than other companies, and 15.2 percentage points lower returns in the next three months.

The Mann-Whitney U test results, as shown in tables 6.1 and 6.2, confirm that tech

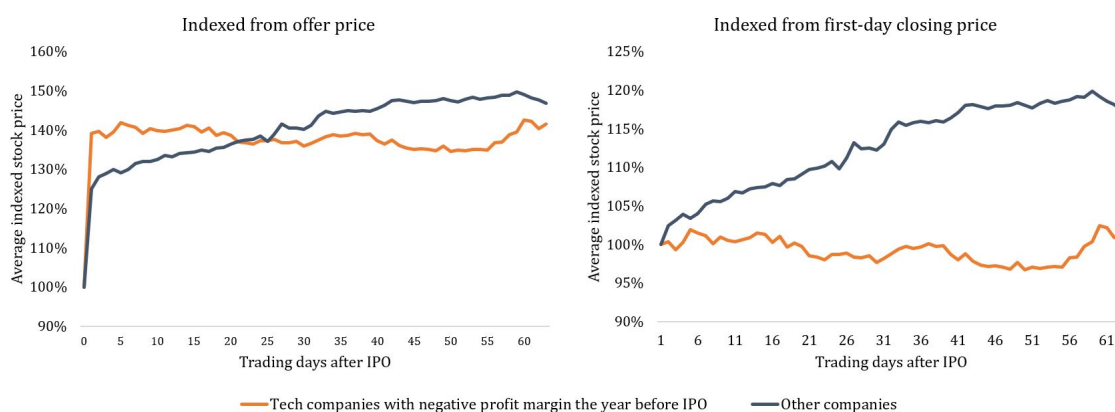
companies with large IPOs outperform others on the first day of trading, with a higher return in 67.1% of all cases. This result is significant at a 5% level, and supports our hypothesis.

On the three-month horizon, large tech IPOs deliver lower returns than other companies 59.9% of the time. However, this underperformance is not significant.

6.3.2 Unprofitable Tech Companies

There are 46 tech companies that were unprofitable the year before IPO, making it the largest overlap of the characteristics that are being examined. These companies had an average first-day return of 39.1%, 14.2 percentage points above other companies. As expected, the next three months generate far lower returns. The stocks of the unprofitable tech companies increased by on average 1.7% - 15.8 percentage points lower than other companies.

Figure 6.4: Indexed Average Returns, Unprofitable Tech Companies



From the regressions in table 5.1 we find that the significance of unprofitable companies is not consistent at a 5% level. However, the unprofitable *tech* companies perform significantly better than others on the first trading day. Companies in this group outperform other companies going public in 69.3% of all cases, leaving them the pairwise combination of attributes that most often generates higher first-day returns than others. This result is significant at a 1% significance level.

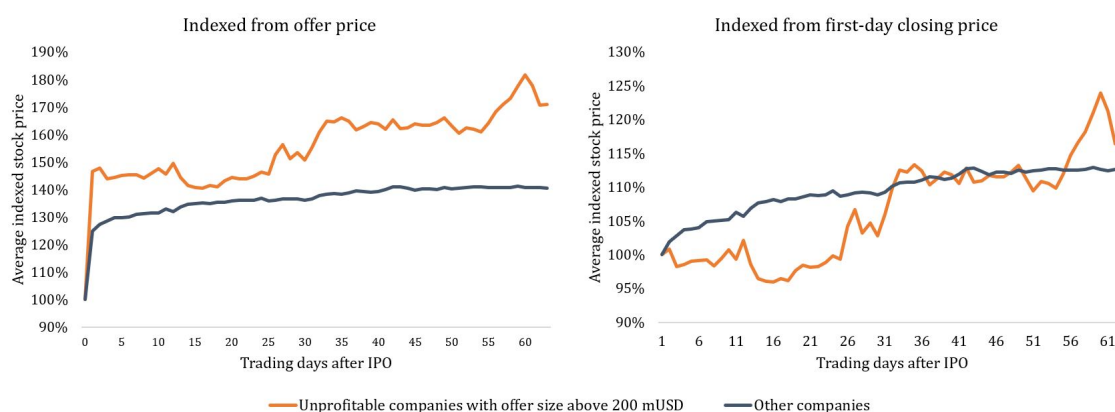
As the only group with significant results in the three-month time frame, the unprofitable tech companies' performance is found to be significantly different than others in the Mann-Whitney U test, at a 10% significance level. This group has a lower three-month

return than others in 59.1% of all cases. Comparing the results from the tests done for the first-day return and the three-month return, the unprofitable tech companies provide the most significant results for both models, but with opposite effects. In other words, our hypothesis seems to be confirmed for this specific combination of characteristics.

6.3.3 Unprofitable Companies With Large IPOs

26 companies with IPOs larger than \$200 million were unprofitable the year before IPO. As shown in figure 6.5, these companies had on average 21.5 percentage points higher first-day returns than other companies, similar to the groups examined above. In the following three months, they also achieved on average higher returns than others, by 4.25 percentage points. Although this irregular pattern might be a consequence of being an average of rather few observations, it is the first assessed group that breaks with the pattern of higher first-day returns followed by lower returns over the next months.

Figure 6.5: Indexed Average Returns, Unprofitable Companies With Large IPOs



As one could expect based on the non-statistic analysis, results from the Mann-Whitney U test are highly significant for first-day returns. It returns a significance level of 1%, and shows that unprofitable companies with large IPOs perform better than other companies in 67.2% of all cases.

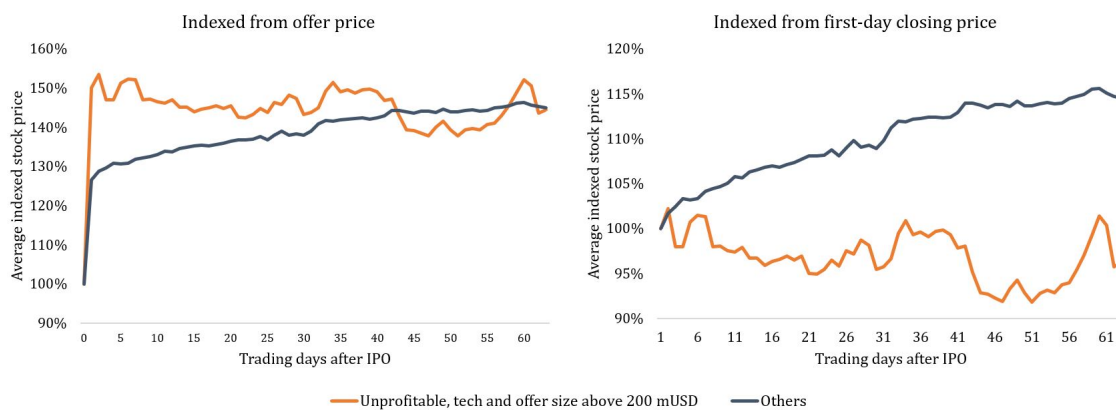
Unsurprisingly, the better performance found for the three-month period is not significant. However, despite having higher average returns than others, the Mann-Whitney U test states that the group performs *worse* than others. A likely explanation is that the group includes for instance Beyond Meat, whose aforementioned sky high returns both on the first day and the next three months of trading pulled the average significantly up. Nevertheless,

with non-significant results, we cannot confirm our hypothesis that unprofitable companies with large IPOs perform worse than others on the three-month horizon.

6.3.4 Unprofitable Tech Companies With Large IPOs

As the Venn diagram in figure 6.2 shows, there are 12 companies that hold all three attributes. As we have found each attribute to impact stock performance, expectations are high that the combination of all three attributes will yield significant results that align with our hypothesis on both time horizons. Figure 6.6 shows that the unprofitable tech companies with large IPOs had a 23.5 percentage points higher average first-day return than other companies. Over the next three months, however, they were outperformed by a margin of 18.3 percentage points.

Figure 6.6: Indexed Average Returns, All Three Attributes



The Mann-Whitney U test confirms that this is the group that most often outperform others the first day of trading, generating higher first-day returns than others in 73.3 of all cases. Unsurprisingly, this finding is significant at a 1% significance level.

Consisting of companies holding all of the three attributes, we expected this group to differ significantly from others also on the three-month horizon. The results from the Mann-Whitney U test states that the group performs worse in 61.5% of all cases, but, slightly unexpected, this result is not significant.

6.4 Summary and Discussion

This chapter has focused on answering whether our hypothesis regarding companies holding more than one of the mentioned attributes hold true. As stated in the beginning of the chapter, our hypothesis was that:

When combining the characteristics “tech”, “unprofitable” and “large IPO”, the positive effect on first-day returns and negative effect on returns in the subsequent three months will be even stronger than the effect of each characteristic individually

To answer whether this is true, both non-statistical and statistical analyses have been used. We found that for the pairwise combinations of the attributes, all three groups perform significantly better than other companies on the first day of trading. Of the pairwise combinations of attributes, the group of unprofitable companies with large IPOs performs better than others in the most cases. Results were even clearer for companies holding all three attributes.

Despite the regressions' lack of significant correlations for the three-month return, we believed that the Mann-Whitney U test would uncover significant patterns. Although the results from the Mann-Whitney U test state that all groups in the test perform worse than other companies in the time frame of three months, only one group provide significant results; the unprofitable tech companies.

In total, all groups deliver higher first-day returns than others, followed by worse performance in the next three months. By the nature of the categories, the high first-day returns may be related to the “hype” of these kinds of companies, or the “fads” as Ritter defined them. As the results are opposite for the performance after three months of trading, excluding this positive first-day trading gain, the differences seem to diminish and the total returns from the initial offer evens out to some extent. This is consistent with Ritter's research. He states that his findings indicate that the underpriced companies are not priced too low, but their market value at the end of the first day is set too high (Ritter, 1991). In our analysis, we see similar trends, but as the three-month results are not consistently significant we cannot draw the same conclusions as Ritter. This might be caused by our shorter time frame of three months, compared to his time frame of three years.

7 Conclusion

We have investigated post-IPO stock performance of initial public offerings on the New York Stock Exchange and Nasdaq that raised private funding before going public. We looked for relationships between stock performance in the first three months of trading and selected publicly available information. Our hypotheses were that status as tech company, amount of pre-IPO funding, IPO size, and negative profitability have (1) positive impact on stock performance on the first day of trading, (2) negative impact on stock performance in the subsequent three months, and (3) that impacts are even stronger for combinations of the attributes. This was tested for by making use of multiple regression analyses, as well as Mann-Whitney U tests.

Our regression analyses provided evidence of a statistically significant positive relationship between offer size and first-day returns. The models also indicated that tech companies achieve higher first-day returns than non-tech companies, and that unprofitable companies achieve higher first-day returns than profitable companies. When looking at stock performance in the three months following the first day of trading, however, the regression analyses did not provide any statistically significant results.

When instead splitting the dataset into groups based on the four attributes, the Mann-Whitney U test partially confirmed that combinations of the attributes yield even more conclusive results. For first-day returns, all combinations yielded significant results. For the next three months, however, only the combination of negative profitability and tech proved to be significant.

Thus, the results partially confirmed our hypotheses, and is in line with previous research done on the subject. Hype and aftermarket risk are typically associated with high first-day returns, which in turn are associated with low returns in the subsequent time period. With increasing access to private equity financing, tech companies' convincing promises to change the world, and decreasing investor demands for profitability in the immediate future, today's companies going public are larger, more hyped and less profitable than ever before. Given our findings, and that these trends in the IPO landscape persist, we can thus expect average stock returns of future IPOs to be higher on the first-day and lower in the longer run than what they are today.

Suggestions for Further Research

Although our analyses provide some significant evidence of factors related to aftermarket performance of venture-backed IPOs, we see several ways of doing further research on the field. Firstly, it is possible to challenge the tech classification. Thompson (2019) argues that several of the high-profile IPOs of 2019 that have depreciated in value are not “real tech companies”. Several of them either sell both hardware *and* software - like the stationary-bike company Peloton - or their core offering is a digital marketplace for transaction of services in the physical world - such as Uber and Lyft. Singling out these “non-pure-tech” companies rather than all tech companies might have produced more significant results. Peloton, Uber and Lyft, as well as Spotify, Groupon and Snap, also have in common that they are consumer-facing, and thus naturally more prone to hype than for instance niche enterprise tech companies - potentially impacting both their pre-market valuations and aftermarket performance.

Secondly, a longer observation period than three months could be applied. We saw for some of the group comparisons that the graphs were seemingly moving further apart from each other over time, potentially indicating that a longer observation period could have yielded clearer results. However, this is not at all guaranteed, and the differences might just as well decrease over time.

Finally, a larger dataset might have provided clearer and more robust results. We were constrained by availability of accounting data, limiting the dataset to IPOs from 2011 onwards. Having access to data from further back in time would increase the amount of observations, thus improving the robustness of our analyses. However, older data might not reflect the aforementioned trends in the IPO landscape. Hence, it could be even more interesting to perform this exercise again at a later stage.

References

- Banton, C. (2019). Underwriter. *Investopedia*.
- Beck, J. (2017). Determinants of ipo underpricing: Tech vs nontech industries. *Major Themes in Economics*, 19.
- Berk, J. and DeMarzo, P. (2017). *Corporate Finance*. Pearson, 4th edition.
- Brav, A. and 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*.
- Carlson, B. (2019). If you think there's something strange about the 2019 ipo market—you're right. *Fortune*.
- Døskeland, T. M. and Strömberg, P. (2018). Evaluating investments in unlisted equity for the norwegian government pension fund global (gpgf). *Regjeringen*.
- Grocer, S. and Russell, K. (2019). Uber is going public: How today's tech i.p.o.s differ from the dot-com boom. *The New York Times*.
- Klebnikov, S. (2019). Wework's valuation plummets to \$8 billion as softbank completes takeover. *Forbes*.
- Loughran, T. and Ritter, J. (2002). Why don't issuers get upset about leaving money on the table in ipos? *Review of Financial Studies*, 15(2):413–444.
- Meggison, W. L. and Weiss, K. A. (1991). Venture capitalist certification in initial public offerings. *The Journal of Finance*.
- Murphy, M. (2019). Beyond meat soars 163% in biggest-popping u.s. ipo since 2000. *MarketWatch*.
- PennState Science (2018). *10.7 - Detecting Multicollinearity Using Variance Inflation Factors*. The Pennsylvania State University, Pennsylvania, USA.
- Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance*.
- Ritter, J. R. (2018). Initial public offerings: Median age of ipos through 2017.
- Ritter, J. R. (2019). Why don't issuers get upset about leaving money on the table in ipos?
- Ritter, J. R. and Beatty, R. P. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *The Journal of Financial Economics*.
- Ritter, J. R. and Welch, I. (2002). A review of ipo activity, pricing, and allocations. *The Journal of Finance*.
- Rooney, K. (2019). This year's ipo class is the least profitable of any year since the tech bubble. *CNBC*.
- Rowley, J. D. (2019). There are more vc funds than ever, but capital concentrates at the top. *Crunchbase*.

- Sarath, C. (2019). Seed funding — things you need to know. *The Startup*.
- Shen, F.-Y. and Goo, Y.-J. (2019). The ipo initial returnsaftermarket risk question revisited: evidence from firms in taiwan. *Investment Management and Financial Innovations*, 16.
- Shiller, R. J. (1990). Market volatility and investor behavior. *The American Economic Review*.
- Thompson, D. (2019). The not-com bubble is popping. *The Atlantic*.
- Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach*. Cengage Learning.
- Zider, B. (1998). How venture capital works. *Harvard Business Review*.

Appendix

A1 Fuzzy String Matching

Table A1.1: Stock Splits Identified for Tickers in Dataset

Ticker	Name from SDC Platinum	Name from Investing.com	Offer Price Multiplier	Fuzz Score	Same Company?
ET	Exacttarget Inc	Energy Transfer	0.25	27	No
HCA	HCA Holdings Inc	HCA	0.22	32	Yes
TTPH	Tetraphase Pharmaceuticals	Tetraphase	20.00	50	Yes
FRAN	Francesca's Holdings	Francescas	12.00	61	Yes
GNCA	Genocea Biosciences	Genocea Bio	8.00	65	Yes
SRC	Spirit Realty Capital	Spirit Realty	2.62	68	Yes
OSMT	Osmotica Pharmaceuticals	Osmotica Pharma	0.50	70	Yes
HTBX	Heat Biologics Inc	Heat Biologics	10.00	88	Yes
RWLK	Rewalk Robotics Ltd	Rewalk Robotics	25.00	88	Yes
WLH	William Lyon Homes Inc	William Lyon Homes	8.25	90	Yes
OBLN	Obalon Therapeutics Inc	Obalon Therapeutics	10.00	90	Yes
ONTX	Onconova Therapeutics Inc	Onconova Therapeutics	150.00	91	Yes
SFBS	Servisfirst Bancshares Inc	Servisfirst Bancshares	0.17	92	Yes

Table A1.2: Fuzzy String Matching of Company Names

Ticker	Name from SDC Platinum	Name from Yahoo Finance	Fuzz Score	Same Company?
EARN	Ellington Residential	Earn	17	No
VLRS	Controladora Vuela	Volaris	26	No
ET	Energy Transfer LP	Exacttarget Inc	30	No
EGRX	Eagle Pharmaceuticals	Eagle	31	Yes
IFRX	Inflarx N.V.	Fireman BV	36	No
PUYI	Puyi American Depository	Puyi Inc	41	Yes
ULTA	Ulta Beauty Inc	Ulta Salon Cosmetic	42	Yes
ARMK	Aramark	Aramark Holdings Corp	50	Yes
FBNK	Facebank Group Inc	First Connecticut Bancorp	51	No
APTV	Aptiv PLC	Delphi Automotive PLC	53	No
TC	Tuanche Limited	Tuanche Internet Info	55	Yes

A2 OLS Violations

Linearity in the Parameters

The assumption of linearity depends on the dependent variable being a linear function of the independent variable (Wooldridge, 2014). This assumption is not very restrictive, as both the dependent and the independent variables can be arbitrary functions of the underlying parameters of interest. Non-linear parameters can be transformed to linear variables in the model through for example functions of natural logarithms and squares. In our dataset, several variables are log-transformed to improve the linear relationship between the dependent and the independent variables.

Random Sampling

For the sample to be representative for the whole population it is crucial that it is collected randomly. Our data is collected from sources we consider reliable, and we gathered all IPOs for the time period we are investigating. The dataset we started with should therefore be representative. However, our data cleaning might bias our data. As data points were missing for many observations, the final dataset only holds information on a fraction of the total number of IPOs in the time period. We cannot be completely sure that the deleted observations were randomly chosen, as there might be reasons why some companies did not have all data points registered. This might bias our results.

No Perfect Collinearity

The assumption of no perfect collinearity states that there can be no constant independent variables, and no exact linear relationship between the independent variables (Wooldridge, 2014). For dummy variables or variables representing shares of a whole, the model cannot include all shares, or mutually exclusive dummies, as they will be perfectly dependent on each other.

While perfect collinearity violates one of the assumptions for the OLS model, multicollinearity can still be included. Multicollinearity can be defined as a high, but not perfect, correlations between the independent variables (Wooldridge, 2014). A model with multicollinearity will provide the best OLS estimates, but it will not be as precise due to larger standard errors. One solution is to remove one of the correlated variables, but it is

important to be aware that it might lead to omitted variable bias. Whether to include the correlated variables or not, is therefore a tradeoff between precision and bias. For research purposes, multicollinearity can create problems, as statistical inference is more difficult. However, when looking for causality, we often favor avoiding the omitted variable bias over good precision (Wooldridge, 2012).

In our dataset, multicollinearity is tested for by calculating the variance inflation indicator (VIF) found in table 1 and through the correlation matrix shown in table 2. Science argue what size a VIF can hold without causing issues in the regressions, due to high correlation. The VIFs observed in table 1 are far below the most commonly used limit of 10, but even when using the more conservative level of 4 (PennState Science, 2018) as our limit, our variables show no sign of severe correlation.

The correlation matrix shown as table 2 confirms the conclusion from the VIF calculations. With correlations higher than 0,8 severe problems with multicollinearity commonly appear, but in the correlation matrix of our independent variables we find no values of such size. The highest values of correlation is found for the variables "over250" and "Intotalraised", with a correlation of 0.7262. This observation seems to be natural, as the dummy variable "over250" is calculated directly from the value of the total amount raised before IPO. The second largest value is found for the number of employees and the size of the IPO, with a correlation of 0.6320.

Table A2.1: Correlation Matrix

	Intotal- raised	lnpri.- amount	ln.- employees	lnage- atipo	profit- negative	over250	tech
Intotalraised	1.0000						
lnoffersize	0.5474	1.0000					
lnno.ofemployees	0.2163	0.6320	1.0000				
lnageatipo	-0.0612	0.0549	0.3089	1.0000			
profitnegative	0.0004	-0.2461	-0.3048	-0.3307	1.0000		
over250	0.7262	0.5265	0.3744	0.0401	-0.1232	1.0000	
tech	0.0636	0.2270	0.3506	-0.1399	0.1039	0.1206	1.0000

Zero Conditional Mean

The zero conditional mean assumption is a key assumption to be able to draw conclusions on causality. It states that the unobserved factors is independent of the independent variable, which means that the error term u have an expected value of zero given any value of the independent variable $E(u|x) = 0$ (Wooldridge, 2012). When this assumption is violated x is often called endogenous.

If the original equation satisfies the zero conditional mean assumption, then no nonlinear functions of the independent variables should be significant when added to the model (Wooldridge, 2012). With this rule as a basis, Ramsey's (1969) regression specification error test (RESET) investigates whether the assumption holds for a given regression model. The RESET test is used to detect functional misspecification in our dataset, and show no signs of omitted variables. The p-value for the regression model with the log-transformed abnormal first day return as the dependent variables is 0.3538, and the regression model with the log-transformed abnormal three-month return as the dependent variable the p-value is 0.2254. Both p-values are above the traditional significant level of 5%, and we cannot reject the null hypothesis of not having omitted variables in the models. We will therefore assume that the variables are exogenous, resulting in unbiased coefficient estimates in the regressions.

Although this may be true, discussions regarding whether the RESET test is an accurate way of detecting omitted variables are present. Wooldridge argues that the test has no power for detecting omitted variables if their expectations are linear in the included independent variables in the model (Wooldridge, 2012). For testing functional form, however, the RESET test is established as a good approach.

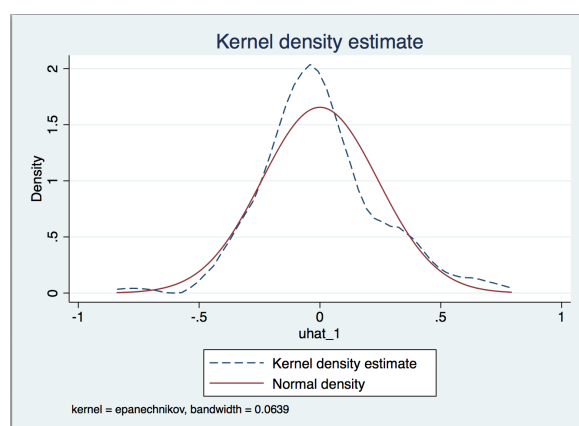
Homoscedasticity

For the assumption of homoscedasticity to hold true, the variance of the unobserved factors, the error term, need to be constant for any given value of the explanatory variables (Wooldridge, 2012). This assumption is important for creating an efficient OLS model, and a violation will result in less precise estimations. However, heteroskedasticity does not cause inconsistency or bias in the estimators of the coefficients, nor affects the goodness of fit (R^2) of the OLS model.

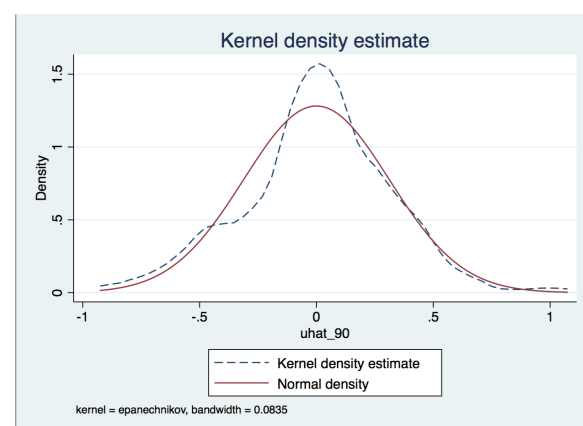
To investigate whether the dataset holds signs of heteroskedasticity, the White's test is used. With first day return as the dependent variable, the regression generates a χ^2 -value of 31,71 and a p-value is 0.4811. For the three-month return the results are almost identical, with a χ^2 -value of 30.56 and a p-value of 0.5397. With p-values far from what would be necessary to reject the null hypothesis of homoscedasticity, we can, according to the Gauss-Markov theorem, assume that the regressions is the best unbiased linear estimator.

The Normality Assumption

For the OLS to be the best linear unbiased estimator there are no requirements of normality. However, normally distributed unobserved error terms are preferred for exact statistical inference (Wooldridge, 2012). For the normality assumption to hold, the unobserved population errors need to be independent of the explanatory variables and normally distributed with zero mean and variance σ^2 (Wooldridge, 2012). As the error term, u , is unobserved, we will investigate the residuals, u_hat , as well as the distribution of the dependent variables. To determine whether the residuals are normally distributed, we have plotted the residuals in a histogram, before using the Skewness and kurtosis test for normality.



(a) Residuals, First-Day Return



(b) Residuals, Three-Month Return

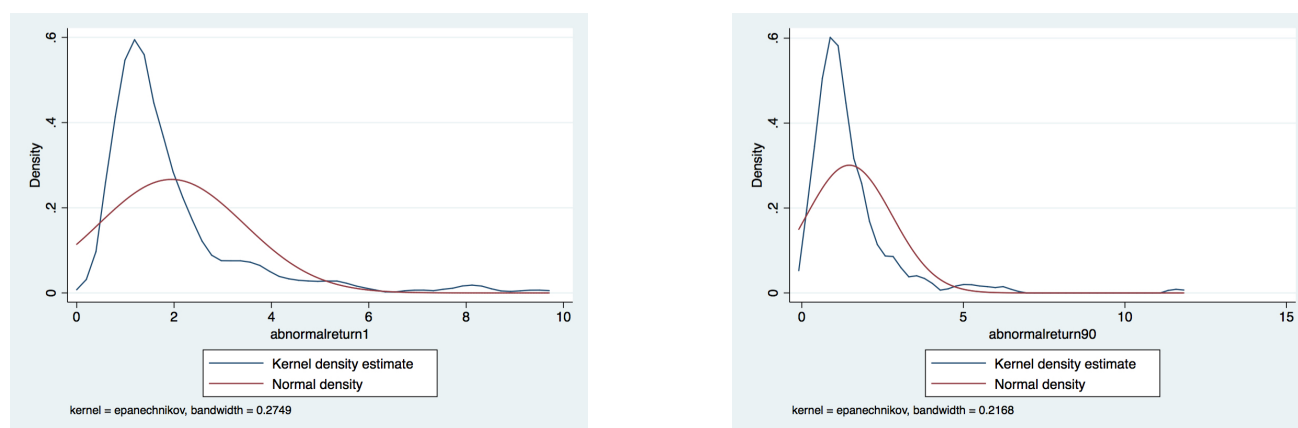
Figure A2.1: Distribution of Residuals, u_hat

Both histograms show that most residuals have values close to zero. The histogram representing the residuals for the regression run with the log-transformed abnormal first day return as the dependent variable seem to have more negative observations than positive, while the histogram for the three months return looks more evenly distributed

around zero. To test whether the skewness is significant enough to reject a null hypothesis of normally distributed residuals, we use the Skewness and kurtosis test for normality.

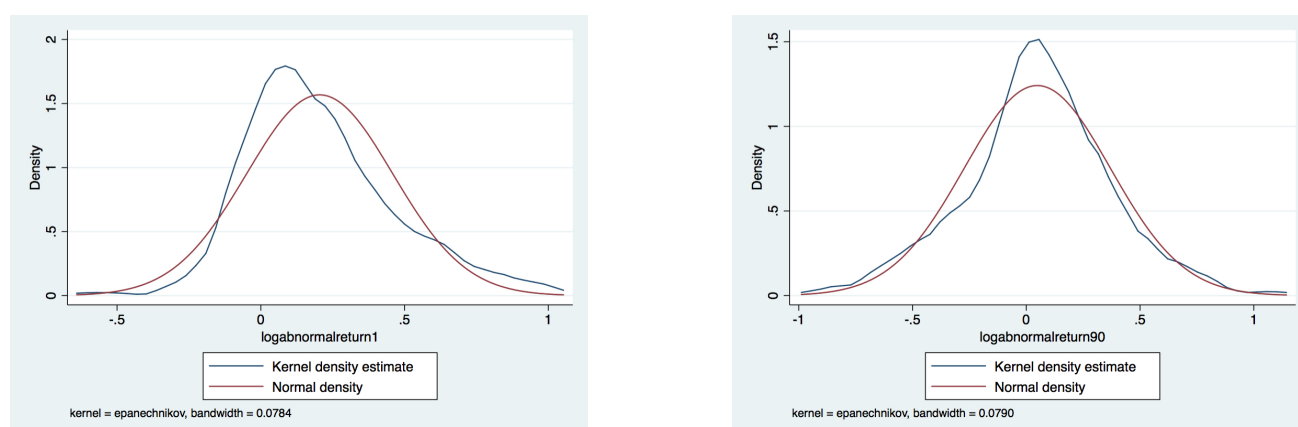
For the first day return regression, the test generates a p-value of 0.0373. As this is less than the 5% significant level, we can reject the null hypothesis of normal distribution of the residuals. However, the three-month regression received different results. The Skewness and kurtosis test for normality generates a p-value of 0.3557 and we cannot reject that the residuals are normally distributed. See appendix 1 for the Stata output.

Further, the dependent variables are assessed. As a starting point, the distribution of the non-logged abnormal return variables are investigated, before looking at the log-transformed versions.



(a) Abnormalreturn1

(b) Abnormalreturn90

Figure A2.2: Distribution of Non-Logged Variables

(a) Logabnormalreturn1

(b) Logabnormalreturn90

Figure A2.3: Distribution of Log-Transformed Variables

The Kernel density estimates for the abnormal returns before logging the variables, show a

distribution that is heavily skewed to the right. Logging them clearly gives better results, with distributions closer to normality. Some skewness is still observable, especially when looking at the first day return where the Kernel density curve has maximum point at a lower value than the normal distribution. The distribution of the three-month return variable seems to follow the normal distribution quite well. To determine whether the skewness is significant enough to reject a null hypothesis of normal distribution, we use the Skewness and kurtosis test for normality.

The p-values for the logged dependent variables confirm our observation that the dependent variables are closer to being normally distributed when they are log-transformed. The p-values of the non-logged variables were 0.0000, confidently rejecting the null hypothesis of normality. When log-transforming the dependent variables, the p-values are of greater size. For the logged three-month return variable, the p-value of 0.2901 is not significant, and we cannot reject the null hypothesis of normal distribution. However, the first day-return still receives a significant p-value of 0.0014, and we have to reject the null hypothesis at the 5% significant level. As with the residuals, normality for the three-month return error seem more likely than for the errors of the first day return model.

We can conclude that the log-transforming dependent variables improved the distribution, but were not sufficient to achieve normality for both models. This may be caused by outliers in our dataset. The most extreme observations are not excluded, but they have been investigated to be sure that the data is correct. However, with a number of observation as high as 180, the central limit theorem conclude that the OLS estimators satisfy asymptotic normality, meaning we can assume we have an approximate normal distribution (Wooldridge, 2012).