Norwegian School of Economics
Bergen, Spring 2022

SPACs off Track

An Empirical Study on Attributes Affecting the Post-Merger Performance of De-SPAC Companies

Erik Hjort & Tom Hoel

Supervisor: Walter Pohl

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis concludes our master's degree in Economics and Business Administration at NHH, with majors in Financial Economics and minors in Business Analytics. Despite the several hurdles along the way, this has been a successful and rewarding semester. We are grateful for the opportunity to spend this semester researching a topic that is important and intriguing to us.

We would like to express our gratitude to our competent supervisor, Walther Pohl, for providing insightful feedback and direction during the semester. Also, thanks to Lars Jonas Andersson, Professor of Applied Statistics, for his advice on model construction and confirmation. We would like to thank all the inspiring people we have met in Trondheim and Bergen throughout this degree. Finally, we would like to thank our families for their continuous support.

Norwegian School of Economics

Bergen, June 2022

Erik Hjort

Tom Hoel

Abstract

Special Purpose Acquisition Companies (SPACs) — acclaimed as a better alternative to the traditional IPO for taking a company public — have been booming since 2020. This thesis analyzes attributes affecting the post-merger performance of SPACs merging between 2020 and early 2022. Throughout this period, SPACs have massively underperformed their benchmarks, except on the first day of trading.

Multiple Linear Regression were used to investigate the relationship between independent variables and the dependent variables first-day and two-month return. We discover a significant negative relationship between performance and the redemption rates (investors withdrawal). Further, we find that the market favored young, profitable, and non-healthcare companies, which outperformed their peers in the short run. Contrary to our initial beliefs, the performance correlates similarly with the designated attributes – independent of the time horizon of interest. Substantiated by the "rise of retail investors" in 2020, we further reveal that two weeks of lagged "hype" has a significant positive relationship with post-merger initial performance.

Based on the obtained results, we suspected that the redemption rate absorbed the effect of the other predictors. This insight was further evaluated using a variety of machine learning models which concluded two things. First, redemption rates indeed absorb the effects in the OLS models. Second, in an Ordinal Logistic Regression, the other variables were able to predict the redemption rate with an accuracy close to 75%.

Keywords – SPAC, IPO, Empirical study, OLS

Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation and Background

Special Purpose Acquisition Companies (SPACs) have become very popular in the actions of taking a company public in the past few years. Since 2020, SPAC IPOs have accounted for more than half of total Initial Public offerings (IPOs) in the U.S., with 55% in 2020 and 63% in 2021. So far this year, 80% of the U.S. IPO market consisted of SPACs. A reverse merger with a SPAC is often considered a "backdoor" for private companies aiming for the public market, forming a so-called "De-SPAC" company.

While SPACs have been booming in recent years, the same cannot be said about their performance. In the paper "A Sober Look at SPACs," the authors find that, on average, SPACs yield overall and significant negative returns compared to their benchmarks: "The declining performance over longer periods suggest a continuous downward adjustment in the market's valuation of post-merger SPACs." (Klausner et al., 2020). The paper further holds the pre-merger dilution in Net Cash per Share accountable for the underperformance.

The motivation that forms the backdrop for this thesis concerns identifying and understanding key attributes driving the performance. Substantiated by the previous findings on certain companies' characteristics and their initial IPO performance, we aim to recognize which elements are crucial for De-SPACs' price movements. These elements concern both fundamental attributes of the SPAC itself and characteristics of the proposed target company. Moreover, as a result of recession-fighting measures in the wake of the Covid-19 pandemic, much of the excess liquidity has found its way into the financial markets. This shift in market dynamics has culminated in a rapid increase in retail investors. These so-called "first-timers" are shown to be different from traditional investors, as they often use alternative platforms as a foundation for their financial decision-making. Despite being regarded as individual investors, their collective influence on the market has proven to be substantial.

In the light of the motivation outlined above, we conduct an empirical study on attributes affecting post-merger performance on SPACs merging between 2020 and early 2022. For the sake of convenience, we will refer to these De-SPACs as our "Merger Cohort."

1.2 Problem Definition

Given the dismal performance of former SPACs, we will concentrate on key attributes that we believe, based on prior research, cause impairment in the value of the De-SPAC companies. We will specifically investigate whether the stock price performance of our Merger Cohort during the first two months of trading can be explained by the hand-picked publicly available information listed below. The first three bullet points concern vital elements of the Net Cash per Share. The latter follows previous empirical research on IPO performance:

- Redemption rates
- PIPE investments
- Total Assets raised
- Characteristics of the target company (profitability, lifespan, sector)

In addition, we seek to quantify to what extent "hype" influences the initial stock performance of our Merger Cohort. To do so, we look at:

- Google Search Hits
- Reddit Mentions

This empirical study intends to address different trends in the SPAC landscape. As such, we aim to identify correlations rather than causalities. Our OLS results show that key elements of Net Cash per Share have a significant effect on both the two-month and initial returns. Particularly, we find that the redemption rate has a detrimental ramification on the stock performance. Furthermore, in conformity with previous IPO studies, we find that the characteristics of target companies are important for equity performance. Status as a healthcare company is shown to be associated with lower returns compared to other industries. In contrast to previous findings, we find that younger companies tend to outperform older ones in our Merger Cohort, achieving both higher returns and more capital from its De-SPAC process. Finally, we find that "hype" associated with search frequency on Google and forum activity on Reddit, is positively correlated with the first-day return.

2 Theory

This chapter will highlight major findings from previous SPAC research. To better understand why firms choose to go public through a De-SPAC, we present relative advantages that favor SPACs over traditional IPOs. First, however, we will go over some key concepts, terminologies, and phrases featured in the thesis. This, coupled with the principles offered in Section 1.1, will serve as the framework for our thesis and establish the associated hypotheses.

2.1 Relevant Theory and Definitions

2.1.1 Organization and Structure of SPAC

A SPAC is a publicly held investment vehicle, which has no commercial operations and is commonly referred to as "blank checks companies." Instead, its conceptual purpose is to raise capital through an IPO, acquire a private company and bringing it public, forming a so-called De-SPAC company (Securities and Commission, 2021). First, SPACs are formed by a sponsor establishing a corporation. Next, the sponsor engages an underwriter to take the SPAC public in an IPO. Further, the SPAC has a timeframe of (usually) 24 months to find a target company. The capital raised in SPAC IPOs is locked up in a trust and invested in government bonds. If the SPAC succeeds in finding a target, shareholders vote for approval of the company in question. Such approval will initiate a reverse merger, where a private company acquires their respective SPAC's shares to become a publicly-traded company. If not, the SPAC is wind-up, and investors are refunded their investments.

2.1.1.1 Sponsors

Sponsors are groups or individuals that form SPACs, typically as limited liabilities companies with their pre-IPO capital. Their primary objective is to engage an underwriter to issue shares to investors in an IPO and attract institutional investors. The sponsors take on a "promote" prior to the IPO, acquiring a block of the shares at a nominal price, typically at 20% of the post-IPO equity (Klausner et al., 2020). The offspring of this promotion serves as compensation for setting up and supporting the SPAC. Additionally, the sponsors can purchase warrants or shares at their estimates of "fair market value."

2.1.1.2 Underwriter

The underwriter is a party that evaluates and assumes another party's risk for payment. The SPAC sponsors engage underwriters to issue shares. The underwriter demands fees of typically 5.5% for risk compensation. In comparison, in a traditional IPO, underwriters typically charge a fee of 5% to 7% (Klausner et al., 2020). Moreover, IPO underwriters make sure that all regulatory requirements are met. Since a SPAC is already public when engaging in the reversed merger with the target company, there is no need for an underwriter to overlook regulatory requirements (Lambert, 2021).

2.1.1.3 Investors

SPAC investors (shareholders) contribute to the capital needed for a merger. Depending on the conversion rate, investors can acquire shares of the merged company equal to the value of their SPAC units. Conceptually, when the SPAC proposes a merger, the shareholders have the right to redeem their shares equal to the initial unit price of $10 (Klausner et al., 2020). Investors that wish to partake in the merger are regarded as non-redeeming shareholders, converting their SPAC units into Class A or B shares of their respective De-SPAC. Moreover, investors in the SPAC's IPO receive warrants and rights included in the units – trading separately alongside the merger. This can be interpreted as compensation for allowing their capital to set the De-SPAC up as a public company. Thus, investors can redeem their shares and keep warrants and rights at no cost – giving them the lucrative advantage of bearing no downside risk.

2.1.1.4 PIPE Investments

SPACs also aims to attract private investments in the proposed merger. These so-called "equity infusions" take the form of private investments in public equity (PIPE) and happen concurrently with the merger (Agarwal, 2021). The benefits of PIPEs are, in short, that they raise a large amount of money quickly by investing directly into the market via stock purchase. Thus, SPAC often seeks additional PIPE investment to complete the merger. Similar to PIPE investments, traditional IPOs attract large institutional investments to raise external capital before going public. However, whereas PIPEs invest in publicly traded shares, underwriters allocate shares to the institutional investors.

2.1.1.5 SPAC Set Up

In the following, we will present the set-up of a SPAC – step by step. This includes the process behind its lifecycle. The figure below summarizes the steps: (1) the public investors buy units in the SPAC's IPO, the sponsor buys share/warrants and receives the promote of 20%; (2) within the timeframe of two years, the SPAC propose the merger of a private company to go public; (3) shareholders either redeem their shares or not; (4) the sponsor purchase shares in PIPEs; (5) the merger take action; (6) the non-redeeming shareholders own a slice of the post-merger company's equity; and (7) SPAC sponsors and other third-party investors owns similarly owns small slices of the equity (Klausner et al., 2020).

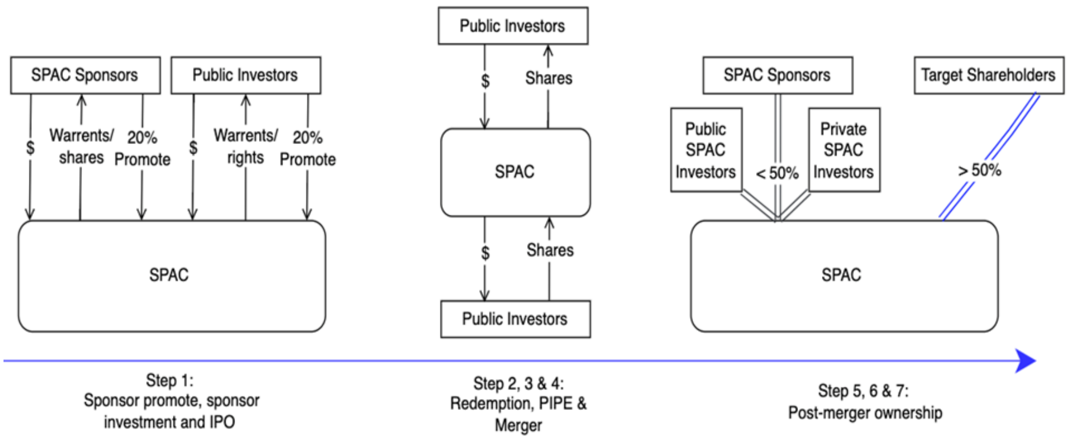

Figure 2.1: The SPAC process

By convention, in its IPO, a SPAC set the price of units at $10.00 (offer price). The unit consists of a share, a warrant, and in some cases, a right to acquire a fraction of a share at no cost when the merger closes. Since 2020, the conversion ratio [1] for warrants and rights are 0.5 and 0.1, respectively (Klausner et al., 2020).

[1]For example, a conversion ratio of 0.5 is interpreted as two warrants for one share.

2.1.2 Relative Advantages of De-SPAC over IPO

This section briefly outlines the commonly stated relative advantages of SPACs over traditional IPOs, from the perspective of an operating (target) company.

2.1.2.1 Speed

First, it often takes less time for an operational firm to arrange a merger with a SPAC and gain shareholder approval. For our Merger Cohort, the average time from deal announcement to merger completion is 21 weeks. As for typical IPOs, the process usually takes six to nine months (PitchBook, 2021). The difference can primarily be attributed to higher requirements, time-consuming roadshows, fund-raising, and financial scrutiny by SEC that companies undergo in preparation for an IPO. De-SPACs, however, avoid these instances as they have already been incurred at their SPAC IPO. It can also be argued that SPAC was structurally more favorable for operating companies to swiftly take advantage of the financial environment, which saw increasingly more tractions and capital inflow from mid-2020 to late 2021 (Klausner et al., 2020). The figure 2.2 illustrates the time from deal annoucement to public debut in weeks.

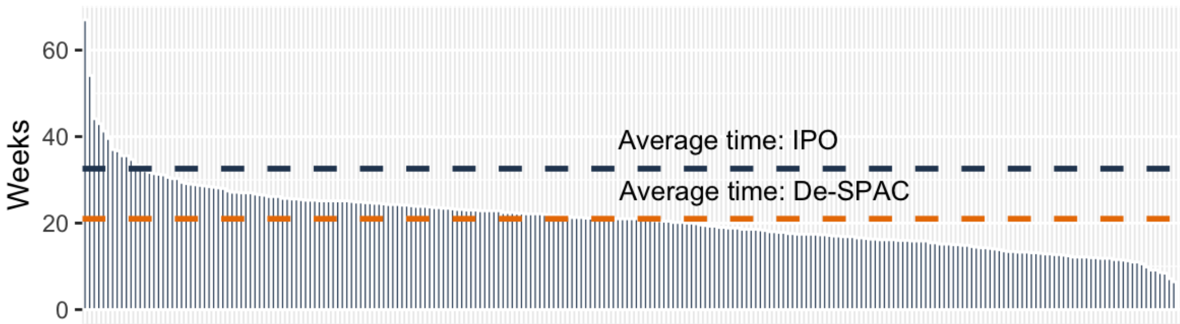

Figure 2.2: Speed to market: SPAC vs. IPO

2.1.2.2 Guidance

Second, target companies gain access to a vast pool of knowledgeable sponsors who can offer guidance and expertise. Comparable to venture capital, they provide both capital and mentorship to the post-merger company (Klausner et al., 2020). Although anyone with adequate resources could establish and carry out a SPAC's objective, many are industry veterans, S&P 500 executives, or fund managers with sought-after experience within their field. Such business insight could be difficult to acquire otherwise without the high inherit compensation that the SPAC structure provides, i.e., 20% of the total money raised. Consequently, both the equity compensation and board membership ensure that sponsors act in the best interest of post-merger shareholders and limit the agent-principals problem, as they now have their "skin in the game."

2.1.2.3 Projections

Third, due to regulatory differences between SPACS and IPOs, a private operating company can issue forward-looking statements such as forecasts and projections without legal precautions during its De-SPAC process (Klausner et al., 2020). Unlike IPOs, a De-SPAC is considered a merger under U.S. laws. Therefore, projections are protected by the regulatory provisions and are, for the most part, shielded from lawsuits in their "safe harbor." Additionally, such projections can be crucial to achieving shareholder approval and attracting investors (PIPEs). Operating companies that seek to optimize their pre-money valuation can therefore greatly benefit from such regulatory arbitrage by merging with a SPAC (Klausner et al., 2020). Often, target companies are categorized as "pre-revenue" or "startups" (Ritter et al., 2021), leaving them with little to no other option than forward-looking statements to communicate their value. Essentially, the SPAC route is an excellent tool for companies that are challenging to correctly price in a typical IPO.

2.1.2.4 Deal & Price Certainty

Finally, it is often stated that SPACs offer better price and deal certainty than its counterpart. Since the funds have already been raised, an agreement between acquirers and target companies can be reached much faster. Although the deal may not go through, all the negotiations between the SPAC and the target company are conducted in a private form. Thus, the firm remains mostly unaffected by negative publicity (Longoni, 2021). In contrast, an IPO is a lengthy process with many unknowns when it is initially initiated. As a result, the firm may wind up with values that do not appropriately reflect the entrepreneur's assessed value. Market conditions might also shift against the firm once the intention to go public is revealed, resulting in the withdrawal of investors' offers.

2.2 Literature Review

This section briefly presents relevant theoretical and empirical literature that we find the most important to our thesis.

2.2.1 The Recent Booming of SPACs

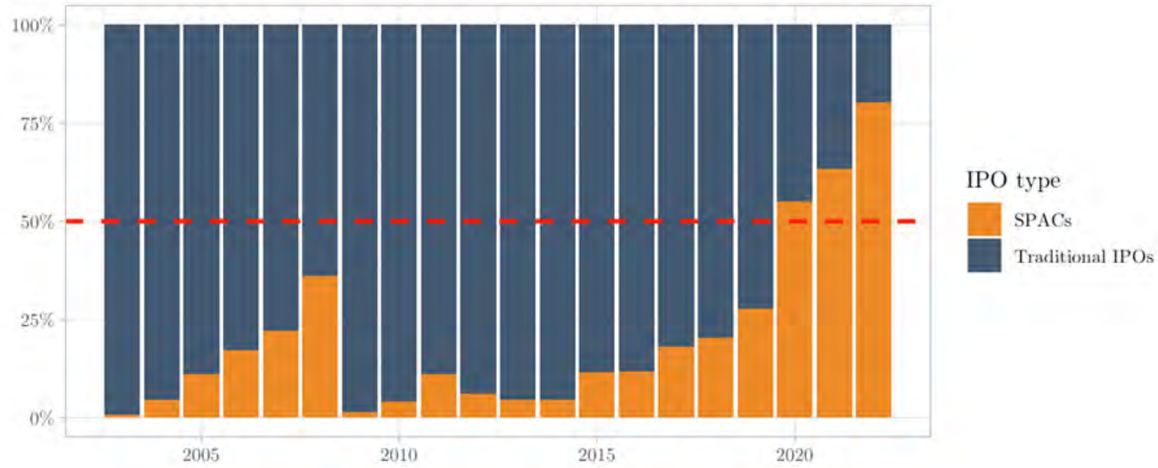

Figure 2.3: IPO composition in the U.S. (2000-2022)

Various factors could explain the surge in interest in the "backdoor" to the public market. From early 2020, we witnessed an increase in volatility levels due to the Covid-19 pandemic. To stabilize the economic landscape, U.S. Federal Reserve enacted quantitative easing by pumping excess liquidity into the market, continuously lowering the Federal Funds Rate, and engaging in Repo operations to support its money markets. Concurrently, the U.S. government began offering stimulus packages for its citizens and businesses. As a result, much of the surplus liquidity floated into the market and was propelled to historic heights after these recession-fighting measures were implemented in mid-2020 (Telford and Siegel, 2021).

Furthermore, in a Bloomberg article, the authors state that "SPACs offered the prospect of better returns that came with some downside protection, thanks to the right of redemption [. . .] the new attention came as venture capital and private equity funds that had pumped money into private companies for a decade were looking for an exit — preferably not an IPO" (Kim and Crystal, 2021). Other research argues that a drop in public companies over the last three decades and the money booming into the market could be a possible explanation. The stock exchange profit from bringing on new companies, thus pushing more SPACs into the market (Panton, 2020).

Moreover, the increase in the amount of capital invested in private equity has been tremendous as well as there has been a significant decline in the number of exits. SPACs serve well as an exit strategy for private equity-backed portfolios, as they allow investors to liquidate their shares at fair market value. Ultimately, the former Financial Times journalist Chris Bryant states that: "SPACs became popular because they offer retail investors the chance to invest in early-stage companies, a domain previously controlled by financial elites." He further concludes that SPACs are the "poor man's private equity" (Bryant, 2022).

2.2.1.1 Underpriced Lunch

Another plausible explanation for the sudden boom can be attributed to SPAC's favorable terms for investors, making it fundamentally more attractive. Even though the terms have remained mostly untouched since 2010, Ritter (2021) argues that the market has not been paying enough attention to this investment vehicle until recently. Moreover, as previously mentioned, a SPAC does not only offer the upside potential to become an early shareholder of the De-SPAC but also a downside risk protection with redemption rights (Klausner et al., 2020).

Participants of SPAC IPOs are usually granted warrants and rights for acquiring De-SPAC shares that are separately traded alongside the SPAC, free of charge. According to the paper "A Sober look at SPACs," the warrants, on average, traded for $1.68 at the time of the merger (Klausner et al., 2020). Hence, on average, investors reap a 16.8% return when liquidating their warrants without any downside risk. Moreover, investors could also achieve additional profits by redeeming their shares after a typical lock-up period of one year, which enjoys additional interest accumulated in the trust. This has proven to be highly profitable for the investors, yielding an average total annualized return of 15.9% (Klausner et al., 2020).

This shows that an investment in SPAC units offers more than its face value of $10, undermining the economic theorem of "there is no such thing as free lunch" (no-arbitrage principle). This underlines the fact that SPACs may be inherently underpriced at IPO. The SPAC market evolved considerably in 2021 towards market efficiency, by offering less attractive warrant terms when possible. Nevertheless, a certain degree of free lunch will still be present until any drastic structural reformations are instated (Ritter et al., 2021).

2.2.2 Aftermarket Performance

2.2.2.1 Two-Months Performance

In "A sober look on SPACs," the authors find that the overall performance in the recent years has been negative when measuring the excess return over different benchmarks. The paper examines SPAC data from 2019 to 2020 and states that "All but one non-high-quality SPAC underperformed, and most underperformed by very large margins. Some high-quality SPACs did well—a few very well—but most others ranged from poor to very poor." In the two-month perspective of interest for our thesis, the paper finds the cumulative excess return over Nasdaq to be approximately equal to negative 10% (Klausner et al., 2020). Interestingly, there seems to be an increase in returns leading up to a peak after approximately one month of trading.

2.2.2.2 First-Day (initial) Performance

While the longer-term performance seems poor, the initial performance yield comparable returns. For instance, Ritter et al. (2021) find minimal first-day returns from 2010 to 2019 and even higher returns for newer De-SPACs. Likewise, Klausner et al. (2020) find that "the average return to investors that bought in at the first day's closing price was 12% above the Russell 2000". Hence, more recent third-generation De-SPACs may have similarities to the famous "IPO-pop," which we will discuss below.

2.2.2.3 First-Day (initial) Performance of IPOs

As to traditional IPOs, the average first-day returns from 2020 to 2021 are equal to 36.85% (Ritter, 2022). These unprecedented gains can largely be attributed to the well-known phenomenon of IPO underpricing (Klausner et al., 2020). Many academics argue that underwriters deliberately underprice shares, in an effort to boost demand and incentivize investors to take on risks in the company. According to Klausner et al. (2020) underpricing, often regarded as an “IPO-pop”, is present regardless of unfavorable fundamental characteristics of the firm. Thus, informed investors can exploit this lucrative opportunity to reap initial abnormal returns, despite companies’ poor quality and prospects.

2.2.3 Net Cash per Share

A central financial ratio for our analysis is the pre-merger net cash per share[2]. Intuitively, the net cash per share ratio can be seen as funds overhanded to the target company after costs, on a per-share basis. Thus, a reverse merger with a SPAC might deliver less cash than $10 per unit. According to previous research, such as Klausner et al. (2020), the net cash per share will dramatically affect the post-merger performance. Below, we will summarize different aspects both causing and preventing dilution in the net cash per share.

2.2.3.1 Attributes Diluting the Net Cash per Share

A vital driver of dilution in the net cash per share is the redemption rate. This is a commonly disputed topic in the literature on SPACs. Various papers, including "A Sober Look at SPACs," claim that high redemption rates cause dilution, ultimately reducing pre-merger net cash per share (Klausner et al., 2020).

Moreover, redundant costs associated with the merger contribute to dilution. Klausner et al. (2020) define these costs as: “value extracted by parties other than the principals in the SPAC transaction – that is, the investors that buy SPAC shares and the target’s pre-merger owners.” Examples of such costs are the sponsor’s promote, warrants and rights to the IPO-stage investors, and underwriting fees and other fees. After accounting for the above-mentioned dilution, the paper finds that: “the mean and median SPACs in our Cohort have just $4.10 and $5.70, respectively, in net cash per share outstanding at the time of their merger” (Klausner et al., 2020).

Although dilution costs in dollars are often non-varying and fixed, the effect of the

[2]
$$Net\ Cash\ per\ Share = \frac{(Total\ Assets * Shares\ Redeemed) + PIPE + Warrants + Cash\ Expenses + Promotes}{Total\ Shares\ outstanding\ after\ Redemption}$$

redemption rate amplifies it on a per-share basis. Initially, 80% of all shares are held by investors, and 20% is granted to the sponsors. A modest redemption rate of 20% recalls 1/5 of the publicly traded shares, depletes the fund, and simultaneously increases the sponsors' share from 20% to 23.81% ($20\%/80\% * 80\% + 20\%$) of the pre-merger SPAC. Moreover, payable costs embedded with SPAC IPO are distributed onto fewer shares than before, further diluting the net cash per share as the redemption rate increases. As a result of these costs, by the time the SPAC merges with a target company, it has far less net cash per share than the $10 attributed to them in the SPAC's merger (Klausner et al., 2020).

2.2.3.2 Attributes Enriching the Net Cash per Share

Given our problem definition, we will examine the following aspects of interest that interact favorably for net cash per share[3]. Previous research finds that the amount of total assets raised and PIPE investments affect the net cash per share. The higher these are, the higher the net cash per share. Klausner et al. (2020) imply that SPACs with low amounts of total assets and PIPE investments will be unable to replace funds lost to redemptions - struggling to reduce dilution.

2.2.4 Characteristics of Target Companies

Aside from attributes regarding the SPAC itself, various features of the target company are of importance for our study. We look to IPO studies to enhance our perspective because the literature on SPACs and target companies is scarce.

2.2.4.1 Profitability

In a study on the stock performance on Venture Capital-backed IPOs from 2011 to 2019, the authors find that "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%" (Røttingsnes and Gjærum, 2019). Contrary, as for first-day returns, the paper finds that companies with negative

[3]Other factors that reduce dilution are waived warrants, lower sponsor promotes, and reduced fees to underwriters. These are, however, not particularly important for our analysis. They will, however, be discussed in Appendix A1 for the purpose of giving the reader a deeper understanding of the mechanisms driving net cash per share.

profitability the year prior to the merger are on average 9% more underpriced than others – suggesting that negative profitability leads to positive first-day returns. Thus, profitability seems to correlate differently with post-IPO performance, depending on which time period one is looking at.

2.2.4.2 Lifespan

In a published article on IPO performance, Ritter (1991) finds that the age of the private company has a statistically significant impact on post-IPO returns. His findings further show that established companies (i.e., older firms) have a strong aftermarket performance compared to younger firms. This is true for the longer-term performance: "the youngest firms [. . .] had exceptionally poor aftermarket performance." As of first-day returns, Ritter states that younger firms: "[. . .] require higher average initial returns and that age is a proxy for this risk". As a result, the first-day performance of young target companies is significantly better compared to more established firms – implying "[. . .] a pattern in the other direction" compared to the long-term performance (Ritter, 1991).

2.2.4.3 Sector

A recent study on industries and their respective performance finds that "Healthcare SPACs deliver highly negative returns in the short term and substantially underperforms small-cap firms and (non-SPAC) healthcare companies" (Gigante and Notarnicola, 2021). In addition, Ritter (1991) finds that the sector of the companies is of importance: "[. . .] the long-run performance of IPOs in different industries varies widely."

2.3 Hypotheses

As mentioned in Chapter 1, we wish to explore the relationship between stock performance and the following five attributes: the number of shares redeemed, amount of PIPE investment, total assets raised, certain characteristics of the target company, and the recent “hype” for SPACs in financial markets. While the first four will form hypotheses 1 and 2, the latter will formulate hypothesis 3. The research assessed in this chapter forms the basis of what we expect to find when analyzing our Merger Cohort. Therefore, we aim to investigate the three hypotheses with a desire to uncover and understand the performance of De-SPACs.

Given that research finds pre-merger dilution in Net Cash per Share crucial for the performance, we investigate how some elements of this affect the two-month stock performance for our Merger Cohort. Moreover, based on IPO research, we believe that the characteristics of the target company are of importance for the returns (Ritter, 1991). Thus, we formulate our first hypothesis as follows:

> **Hypothesis 1 (H1)** *Redemption rates, low PIPE investments and total assets raised, status as a healthcare company, short lifespan, and negative profitability are all* ***negatively*** *correlated with the two-month simple return.*

Unlike prior studies on two-month performance, the first-day return of recent De-SPACs tends to be positive. These returns are comparable to the so-called "IPO-pop," which stems from share underpricing when the issuer’s shares close above the listing price (Klausner et al., 2020). Moreover, the first-day returns of IPOs within the same time period of interest are significantly positive (Ritter, 2022). Thus, following Ritter (1991), we expect the above attributes to have an inverse relationship with the first-day returns (i.e., they are positively correlated). This expectation substantiates our belief that investors put less emphasis on key attributes, instead relying their financial decision-making on the aggregate demand and underpricing of IPOs:

> **Hypothesis 2 (H2)** *Given that previous De-SPACs obtain a “pop” on the first day of trading, the aforementioned attributes are* ***positively*** *correlated with first-day returns.*

Another historical event, dubbed "the rise of retail investors," occurred within the same period. Now, more than ever, the stock market has gained more exposure from unconventional channels, such as social media, alternative news outlets, and internet forums. As a result, we have seen a shift in market dynamics. Thus, we want to investigate whether this has a substantial impact on the performance of the De-SPACs. Particularly, we are interested in whether the positive first-day returns observed in other papers are associated with “hype”.

> **Hypothesis 3 (H3)** *The amount of “hype” associated with the weeks prior to the merger is* ***positively*** *correlated with first-day returns.*

By answering the above hypotheses using recent data, we hope to contribute to prior research on SPACs. Moreover, to our knowledge, there exist no literature on how “hype” from different sources affects De-SPACs. Hopefully, this can shed light on retail investors’ impact on the market sentiment and De-SPAC stock performance.

3 Data Collection and Cleaning

Following our hypotheses, we need data on numerous areas to provide sufficient answers: (1-3) concerns our independent variables, while (4) cover the dependent variables:

1. Pre- and post-merger financial metrics for the SPACs
2. Pre-merger financial metrics for the target companies
3. Measurements of "hype" from different sources
4. Post-merger share prices for the De-SPAC companies

In this chapter, we will explain the collection, cleaning, and compiling of our data.

3.1 Independent Variables

3.1.1 Elements of the Net Cash per Share

The pre-merger dilution in net cash per share is vital for performance. Hence, we wish to explore if certain elements of the equation are strongly connected with post-merger returns. The net cash per share (NCPS) is calculated using the following equation:

$$NCPS = \frac{(\textit{Total Assets} * \textit{Redemption}) + PIPE + Warrants + \textit{Cash Exspenses} + Promotes}{\textit{Total Shares outstanding after Redemption}} \tag{3.1}$$

We are particularly interested in the redemption rate, PIPE investments, and the total assets raised [4]. We collect this data from different SEC filings for the SPAC companies. One shortcoming of the SEC filings is that there is no explicit disclosure of information on SPAC cost, making it difficult to calculate the net cash per share. Thus, some of the data must be laboriously pulled from many different sources in the underlying securities filings. Below, we further elaborate on the variables of interest and where they are collected from.

[4] Since we later in the thesis will present the median net cash per share for our Merger cohort, the collection of the remaining part of the equation is explained in Appendix A2

3.1.1.1 Redemption Rates

The redemption rates can be obtained by analyzing relevant SEC filings, such as the 8-K and 10-K. The latter is helpful in examining the quarterly development of the redemption rate before a merger. We are interested in the final redemption volume, a 8-K filing following a merger includes such information under Section 2.01 "Completion of Acquisition or Disposition of Assets." We extract both the number of shares redeemed and their market value in USD at the merger.

3.1.1.2 PIPE Investments

Since the PIPE (Private Investment in Public Equity) works as an equity infusion concurrently with the merger, the mandatory filing of the 8-K after the merger (item 2.01) is used to obtain the data we need. PIPE investments are extracted manually from the respective filing due to the various inconsistencies in filings prior to 2020 when the term PIPE was yet to be coined. For instance, these equity infusions were referred to as “Private Placements” or “Equity transactions.” These are indistinguishable from PIPEs, and they are regarded as such in our study.

3.1.1.3 Total Assets

Both “cash in trust” and “cash outside trust” are found in the 10-K balance sheet. We define the sum of these as the “total assets,” as these are the cash obtained in the SPAC IPO, which will ultimately be handed over to a target company. Hence, it is crucial for the analysis that we extract numbers from 10-Ks *before* the merger of the target company. Using R, we scrape the SEC website and pull data of interest from the 10-K filings filed prior to a De-SPAC merger.

3.1.2 Characteristics of the Target Company

3.1.2.1 Profitability

The Orbis database extracts relevant key indicators of private companies (i.e., the target companies). According to Orbis themselves, they are “the world’s most powerful comparable data resource on private companies” (Orbis, 2022). Profitability is measured using the target companies’ Return on Equity (ROE) the year prior to the merger. This is because we want to examine the performance of the companies before the merger. ROE is a commonly used measure of profitability, measuring the net income divided by its shareholder's equity (Fernando, 2021).

3.1.2.2 Lifespan

We collect target companies' founding date and merger date using multiple sources such as Google Engine, SPAC Research, and the company's website. The gap between the two metrics defines their age before undergoing a De-SPAC process. In our thesis, we refer to this age as the "lifespan."

3.1.2.3 Sector

SPAC Research had a description of what sector each target company operated in. We cross-validated this against the information found on the Orbis database.

3.1.3 Hype Indicators

3.1.3.1 Google Trends

We need a quantifiable metric that captures the so-called "buzz" to study public interest in specific securities. The buzz indicator, now known as "hype," is defined in this case study as the relative popularity of a specific search query. GOOGLE TRENDS is an online analytics service that provides "[...] the search volume of various keywords over time." (GoogleTrends, 2022). The output is a weekly time series with a popularity metric ranging from 0 to 100. The value 0 indicates that the phrase has not been searched at all or that it is missing data. 100 denotes the term's relative peak in interest. We extracted these time series for all De-SPACs using the R package `GtrendsR`. Since we are primarily interested in the hype leading up to the merger, we extract the average value from week_{t-2} to week_t – as illustrated below:

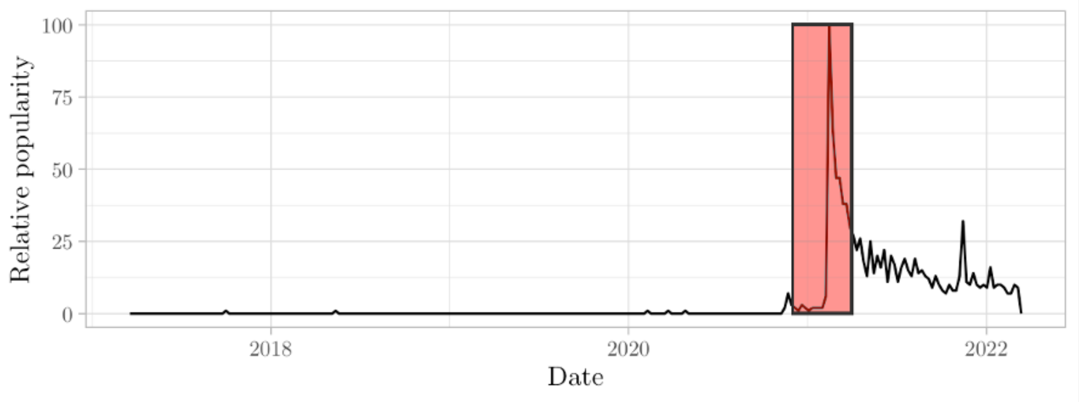

Figure 3.1: Google Trends run-up hype

3.1.3.2 Reddit

Reddit is a self-proclaimed "front page of the internet" that has garnered much traction in recent years. The "rise of retail investors" (Reddel, 2022) can be traced back to this platform, where ideas and investing techniques are frequently discussed in both formal and informal settings. While Google Trends captures the broader popularity in search volume, we are particularly interested in how frequently the De-SPACs are discussed on Reddit in the run-up to the merger. Using the R package `RedditExtractR`, we were able to extract the number of posts and comments related to a search parameter. Here, we focus on the name of a De-SPAC rather than its ticker symbol, as used previously for Google trends.

3.2 Dependent Variables

3.2.0.1 Two-months and first-day simple return

We want to see if the above independent variables can be used to explain the aftermarket stock performance of De-SPAC firms. Thus, we gather historical stock price movement for our Merger Cohort. Although our original "Merger Cohort" included 281 De-SPACs, not all of them had sufficient financial data. For example, the redemption rates and PIPE investments are missing for some of the observations. After accounting for this, as well as missing financial data from Yahoo! Finance, we have a total of 103 De-SPAC companies to analyze. A descriptive summary table of the process is presented at the end of this section.

The dependent variables for our hypotheses will be the two-month simple return and the first-day simple return[5]. While other papers use three-months return to capture the performance after companies' public debut, we instead use the two-months return. The justification for such deviation is based on our desire to study as recent data as possible. Given that we have data from February 2022, any prolonged-time period would have resulted in fewer observations to evaluate. Furthermore, Klausner et al. (2020) finds a considerable reduction in performance around one month after the merger. As a result, the two-month return should be able to capture possibly poor performance. The two-month simple return is calculated using the calculation below:

$$\textit{Two-month simple return} = \frac{\textit{Closing Price}_{t+42}}{\textit{Offer Price}} \tag{3.2}$$

We perform an OLS regression on the first-day return against our predictors to investigate initial performance. The use of first-day returns is motivated by prior research on De-SPACs that discovered favorable first-day returns. Furthermore, research on traditional IPOs reveals that, in recent years, many companies have seen big IPO-pops occur on their first trading day (Ritter, 2022). The following equation is used to compute the first-day return:

$$\textit{First-day simple return} = \frac{\textit{Closing Price}_{t}}{\textit{Offer Price}} \tag{3.3}$$

Later in the descriptive analysis, we analyze the cumulative returns obtained in our time of interest against several benchmarks to assess the performance of our Merger Cohort. In addition to a computed IPO benchmark, the Russell 2000 index is of interest. The enterprises listed on Russell 2000 have a median market capitalization of USD 580 million (Russell, 2021). Given that the median market capitalization for our Merger Cohort is around USD 500 million, we believe that Russell 2000 will serve well as a reference benchmark. Our IPO benchmark is the development of stock prices for all traditional IPOs that went public concurrently with our Merger Cohort. The goal is to have a broader and more comparable understanding of the two listing choices.

[5] In accordance with previous literature on initial performance, the simple returns are indexed at their offer price. Unit price of SPAC can be interpreted as their IPO offer price, which is usually set at $10

In order to compute the excess return for our Merger Cohort later in the descriptive analysis, some extra measures are deemed necessary. The data is transformed to account for differences in merger dates and frequency. In an effort to capture the recent SPAC bubble[6], we decided to use the first quantile of our data as our baseline when conducting excess return calculations. Dates were later defined as days following the merger, such that all De-SPACs essentially became publicly traded on "the same day." In addition, the relevant indexes are extracted from the same baseline date. This procedure is illustrated in the figure below, and the same procedure is applied to our IPO benchmark.

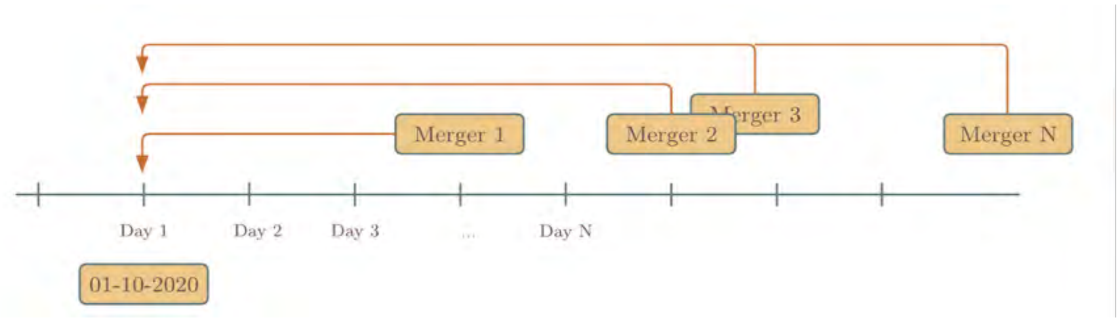

Figure 3.2: Transforming the cumulative returns

The excess return is defined as the difference between our Merger Cohorts cumulative return and the cumulative return of our benchmarks. Excess returns are an important metric that helps a De-SPAC investor to gauge performance in comparison to alternative investments. In general, all investors hope for positive excess return because it provides an investor with more money than they could have achieved by investing elsewhere (Chen, 2021). The cumulative and excess return are defined as:

$$Cumulative\ Returns_{(t_0,t_n)} = (1 + r_{t_0}) * (1 + r_{t_1}) * ... * (1 + r_{t_n}) - 1 \tag{3.4}$$

$$Excess\ return = \overline{CR}_{De-SPACs} - CR_{Benchmark} \tag{3.5}$$

[6] In the Beginning October of 2020, prices surged to a peak of about $11.50 in February of 2021, and then back down to $10.00 in mid-2021 (Klausner et al., 2020)

3.3 Selection Bias

The observations in our dataset must be representative of the De-SPAC companies we aim to evaluate for our thesis to be reliable. Our data is acquired from dependable and credible sources. The number of entities, however, has been significantly reduced throughout the data cleaning and compilation process. This is unavoidable in an effort to conduct the necessary analysis. As a result, the condensed dataset may be prone to selection bias. While we did our best to ensure the accuracy of the data we collected and aggregated here, errors in data collection could be present. However, we are confident that such errors would not have a material impact on the aggregate data that we report.

Observations	**Description**
	Period: 01.01.2020 - 02.10.2022
281	De-SPACs extracted from SPAC Research
194	Excluding firms with missing PIPEs
126	Excluding firms with missing Redemption rate
110	Excluding firms with missing data from Yahoo! Finance
103	**Excluding firms with missing ROE from Orbis**

Table 3.1: Data cleaning process

4 Descriptive Analysis

In this section, we will present our cleaned and compiled dataset using descriptive analysis. First, we will evaluate our Merger Cohort's performance by graphing the cumulative return against relevant benchmarks. Followed by a brief overview of the most relevant aspects of our data, we will analyze the data distributions. This will serve as the foundation for subsequent statistical analysis. Finally, we strive to acquire an overview of patterns evident to the naked eye by graphing post-merger returns on the independent variables.

(Later, in section 6, we will provide a descriptive analysis of our "hype" predictors associated with hypothesis 3)

4.1 General Descriptive Analysis

4.1.1 Two-Months Performance

To elaborate on the performance of our Merger Cohort, we have plotted the two-month performance against Russell 2000.

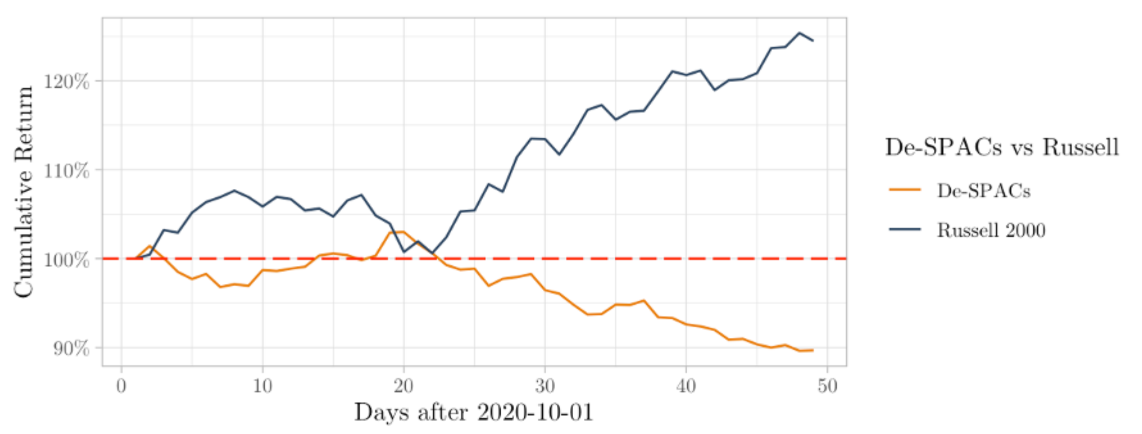

Figure 4.1: De-SPACs versus Russell 2000 Index

According to the figure 4.1, our Merger Cohort performed considerably worse than its benchmark in cumulative terms. The red-dotted line in the plot illustrates the "break-even line." The line representing the De-SPACs (orange) is far below this threshold in the two-month perspective (42 trading days). Moreover, the alternative cost of investing in a "De-SPAC portfolio" rather than Russell 2000 (blue) is even higher. Illustratively, an investor with a two-month buy-and-hold strategy, would (excluded the alternative of investing in the Russell 2000 index) be better off not investing at all. However, we do see a peak in the performance leading up to approximately 20 days of trading. Thus, the performance of our Merger Cohort is in accordance with previous literature such as Klausner et al. (2020). Moreover, we analyze our De-SPAC cohort against its traditional counterpart, IPOs. By graphing their respective cumulative returns, we hope to capture the relative difference in investing in De-SPACs compared to IPOs:

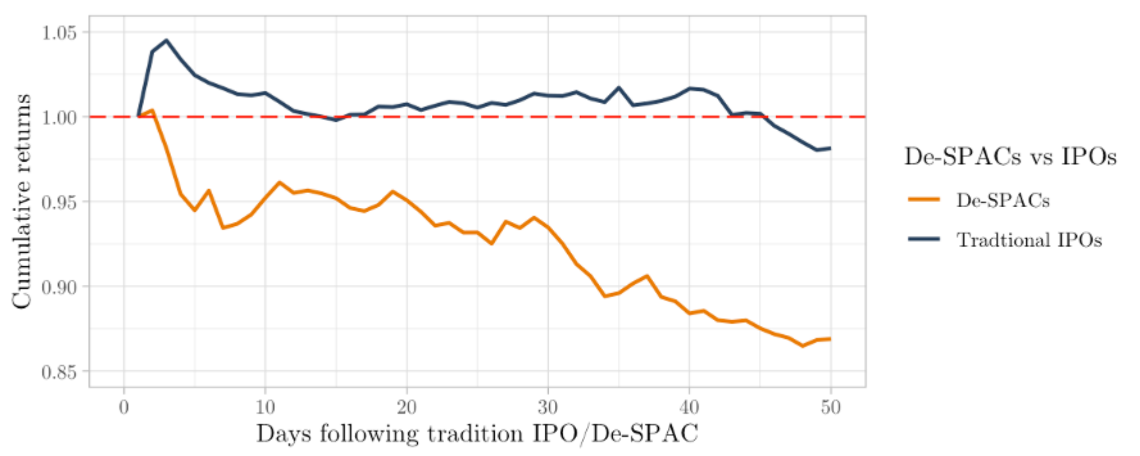

Figure 4.2: De-SPACs versus traditional IPOs

After two months of trading, traditional IPOs outperform our Merger Cohort by 13% $(-0.88 + 1.01)$. However, this is not due to the quality of the IPOs, but rather the lack of it in the De-SPACs. The fact that both perform relatively poorly in a two-month perspective could raise the question of whether going public has been a good strategy in recent years. We will, however, leave that question open for now. Nonetheless, the findings are intriguing and could be pursued further in later studies.

4.1.2 First-Day (initial) Performance

While the cumulative return against benchmarks is focused on a two-month timeframe, the above figure does not adequately display first-day returns. Thus, the figure below is created to depict the average first-day return for our Merger Cohort.

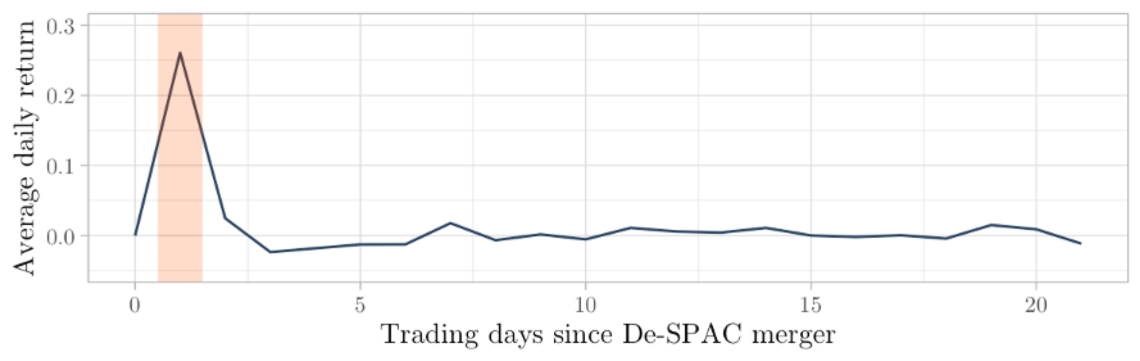

Figure 4.3: De-SPAC "pop"

The figure shows a distinct "pop" on the first trading day, comparable to the previously discussed "IPO-pop" (the IPO-pop is present in our data too, as of figure 4.2). Thus, our Merger Cohort's average first-day return is positive, resulting in initial gains for investors. On the contrary, from the perspective of the target company, the pop could signify an undervaluation of their company.

The table 4.1 below illustrates our Merger Cohort's median first-day and two-month returns. Except for the 2022 De-SPACs, every quarterly statistic on median first-day return yield positive returns. Furthermore, the two-month excess return over Russell 2000 is negative throughout all quarters of interest. In short, it appears that the performance of our data is similar to what prior study has established.

Quarter	Number of Companies	Median first-day	Median two-months Excess Russell 2000
2020 (3)	4	0.32	-0.13
2020 (4)	11	0.01	-0.09
2021 (1)	17	0.64	-0.15
2021 (2)	9	0.16	-0.17
2021 (3)	29	0.20	-0.22
2021 (4)	20	0.00	-0.16
2022 (1)	13	-0.04	-0.28
Total	**103**	**0.11**	**-0.18**

Table 4.1: Descriptive Statistics, by Quarter

4.2 General Descriptive Statistics

4.2.1 Elements of the Net Cash per Share

Year	Number of Companies	Median Redemption Rate	Median PIPE ($mUSD)	Median Assets ($mUSD)	Median Net Cash ($USD)
2020	27	79.00 %	170	239.54	6.11
2021	74	28.30 %	200	258.09	6.66
2022	2	85.55 %	110	164.19	2.09
Total	**103**	**42.70 %**	**190.38**	**251.40**	**6.43**

Table 4.2: Elements of the Net Cash per Share: Descriptive Statistics, by Year

As shown in table 4.2, the median redemption rate is somewhat volatile. Exceptionally high redemption rates are observed in both 2020 and 2022. However, during the "SPAC bubble" in 2021, redemption rates were much lower than in previous years. Our Merger Cohort has a weighted median redemption rate of 42.70 percent. For their 2019-2020 De-SPACs, Klausner et al. (2020) finds this to be 73 percent. Thus, the SPAC bubble seems to impact this figure notably.

The weighted median for PIPE investments is $190.38mUSD. While the years 2020 and 2021 are roughly identical, the companies in 2022 received less equity inflow upon their merger. The amount of total assets raised in 2020 and 2021 are similar, whereas 2022 raised significantly less. The weighted median is calculated to be $251.40mUSD. Nevertheless, this is much higher compared to the proceeds raised in traditional IPOs in the same period, which raised a weighted median of $100.67mUSD (Westenberg et al., 2022).

The combination of these factors results in a weighted median net cash per share of $6.43, implying a $3.57 dilution due to stockholders redeeming their shares and other cost related to the merger. This suggests that 35.7 percent of total assets are drained from post-merger liquidity. While the calculated dilution cost is persuasive, it is significantly less severe compared to the previous SPAC cohort. According to Klausner et al. (2020), the median net cash per share for their respective cohort is $5.70.

4.2.2 Characteristics of the Target Company

In recent years, we have seen a clear shift in market dynamics. Contrary to what Ritter et al. (2021) finds, our target companies tend to be younger, smaller, and less profitable.

Year	Number of Companies	% Healthcare	Median Lifespan (Years)	Median % ROE
2020	27	14.81 %	8	-2.49 %
2021	74	28.38 %	8	-4.52 %
2022	2	50.00 %	1	-38.64 %
Total	**103**	**31.06 %**	**7.86**	**-4.65 %**

Table 4.3: Characteristics of target companies: Descriptive Statistics, by Year

Furthermore, in Ritter's study on De-SPAC mergers between 2013 and 2020, biotech businesses accounted for 37% of traditional IPOs while accounting for only 8% of SPAC mergers (Ritter et al., 2021). In our Merger Cohort (2020-2022), 24% and 31% can be categorized as biotech and healthcare, respectively. The Covid-19 pandemic and the speed to completion of SPAC could be a logical explanation for this unexpected spike. Table 4.3 reveals that the median age of our Merger Cohort is eight years, except for a 2022 outlier. Conversely, Ritter discovers that the median age of companies opting for traditional IPOs during the same period is eleven years (Ritter et al., 2021).

The results of studying the median profitability are fascinating. On average, target companies exhibited negative profitability the year prior to the merger. This is consistent throughout our time period of interest, implying that most target companies had a negative return on equity before going public. In conclusion, it is evident that De-SPACs in our cohort are relatively young, largely healthcare-related, and unprofitable.

Following that, we will go over the above-mentioned independent variables in greater detail by graphing the dependent variables' first-day and two-month returns against them. We aim to find significant trends in our data by doing so.

4.2.3 Redemption Rates

As previously stated, redemption rates have a significant impact on pre-merger net cash per share, which in turn has an impact on De-SPAC post-merger performance. Recall that redemption rates deplete the target company's cash holdings, and severely dilute the eventual shares outstanding. In accordance with general financial theory, this dramatically impacts the stock's performance. Below, we have plotted the first-day- and two-month simple returns against redemption rates. Some of the observations are approximately zero, indicating that "all" shareholders have retained their shares.

As we can interpret from the plot, there seems to be a negative relationship between returns and shares redeemed. While this is in accordance with our hypothesis 1, the opposite seems true for hypothesis 2. Examples of De-SPACs with low redemption rates and high returns are Nikola Corporation ($NKLA) and QuantumScape Corp ($QS). Both firms are related to the production of electric vehicles. Conversely, 180 Life Sciences Corp ($ATNF) has performed very poorly in addition to embed high redemption rate (>90%). Interestingly, this is a healthcare company.

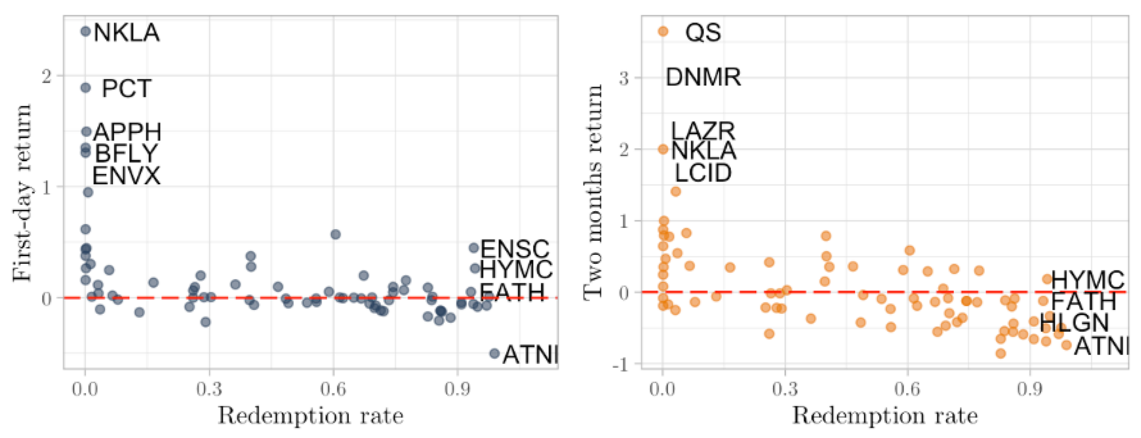

Figure 4.4: Plotting simple returns against Redemption Rate

4.2.4 PIPE Investments

While redemption rates seem to be negatively correlated with two-month returns, a high amount of PIPE investments are expected to be beneficial for the post-merger performance. We plot our dependent variables against PIPE investments in figure 4.5 below. Redemption rates seem to have a more easily identifiable linear relationship with returns than PIPE, which is somewhat harder to recognize. However, it seems to be a slight positive correlation between equity infusions and performance. The well-known De-SPAC, Lucid Motors ($LCID), stands out with high returns and large PIPE investments.

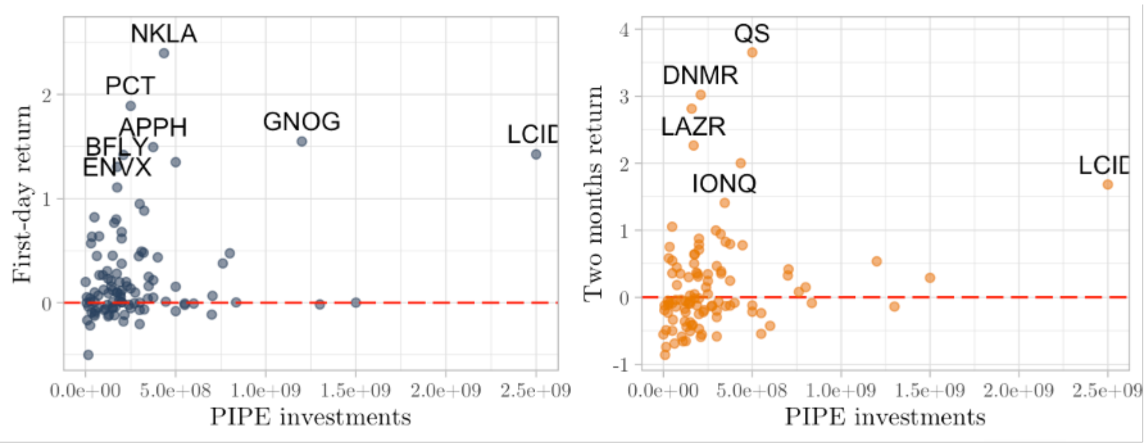

Figure 4.5: Plotting simple returns against PIPE Investments

4.2.5 Total Assets

Total assets raised (i.e., the IPO proceeds) are, according to previous research, positively correlated with the performance. However, when plotting total assets against the returns, it is difficult to identify any clear pattern – especially for the first-day returns. The investment firm Owl Rock Capital Group ($OWL) managed to raise the most capital ($7.5 billion) through a reverse merger with the SPAC Altimar Acquisition Corporation. By the end of the year, OWL had $27.1 billion under management. Despite raising the most capital, the De-SPAC saw modest returns and got beaten by others with less capital raised, such as $NKLA, $LCID, and $SKLZ.

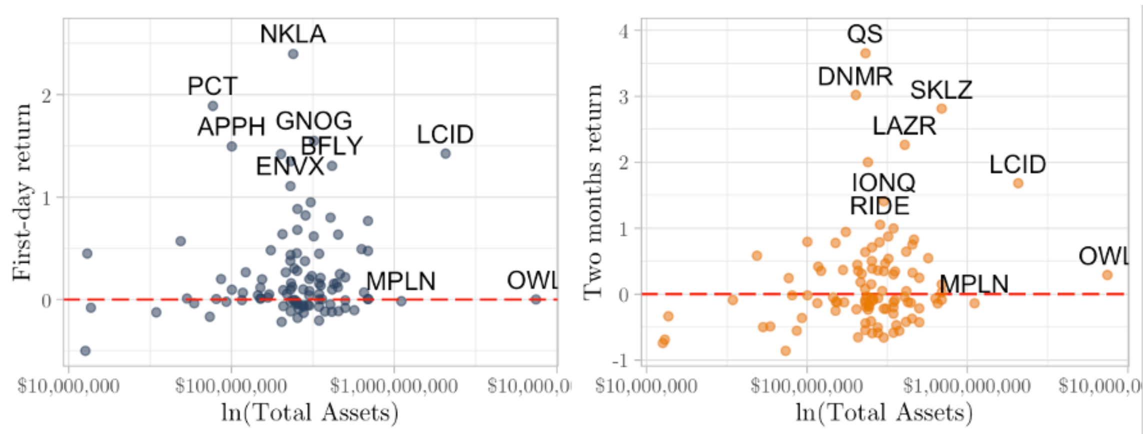

Figure 4.6: Plotting simple returns against Total Assets

4.2.6 Profitability

Profitability is defined as the target company's Return on Equity (ROE) one year prior to the merger. Below, we have plotted the returns against ROE. We define "unprofitable" and "profitable" companies as companies with negative and positive ROE, respectively. The plots suggest no clear relationship between profitability and returns. In our Merger Cohort, 29 target companies had positive ROE prior to their De-SPAC process. Their average first-day return is 27.63%, whereas their counterpart achieved 9.07%.

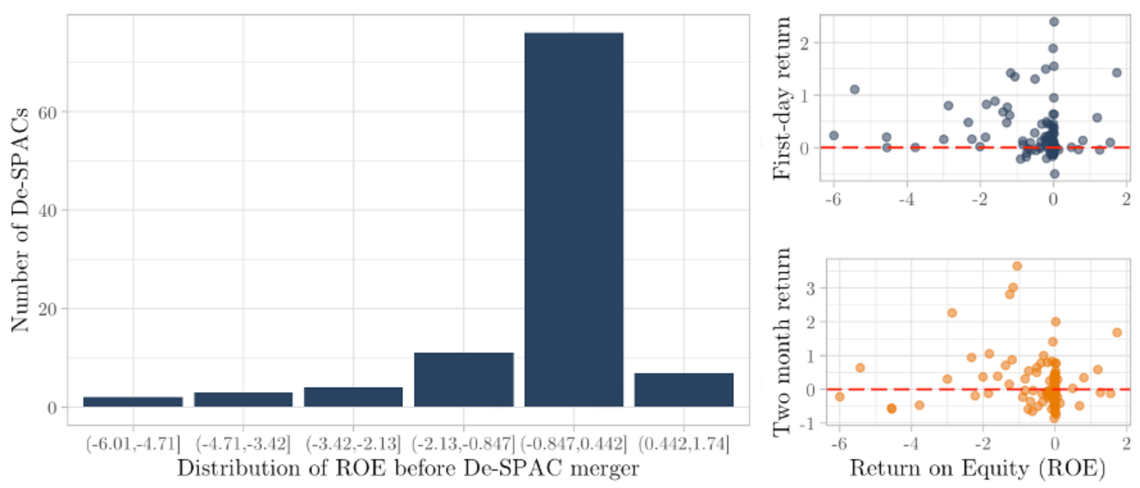

Figure 4.7: Plotting simple returns against Profitability

4.2.7 Lifespan

As shown previously in this paper, the median age for the companies of interest is close to eight years. Ritter et al. (2021) found that older target companies tend to outperform younger ones in the long run. Based on the graphs and the distribution plot 4.8 of our Merger Cohort, it is difficult to identify any clear pattern. Thus, suggesting a definite linear relationship between returns and company age at IPO is unlikely. Despite the non-linear pattern identified, some useful information can be extracted. We see that large returns are mainly coherent with age less than 25. Moreover, it seems like an increase in age over 25 decreases the return. The oldest target company in our dataset is Algoma Steel ($ASTL) at 119 years old, achieving 12.5% and -3.7% first-day and two months return, respectively. Whereas the automotive unicorn, Nikola ($NKLA) at only six years old, reached 240% on its first day of trading and 200% after two months.

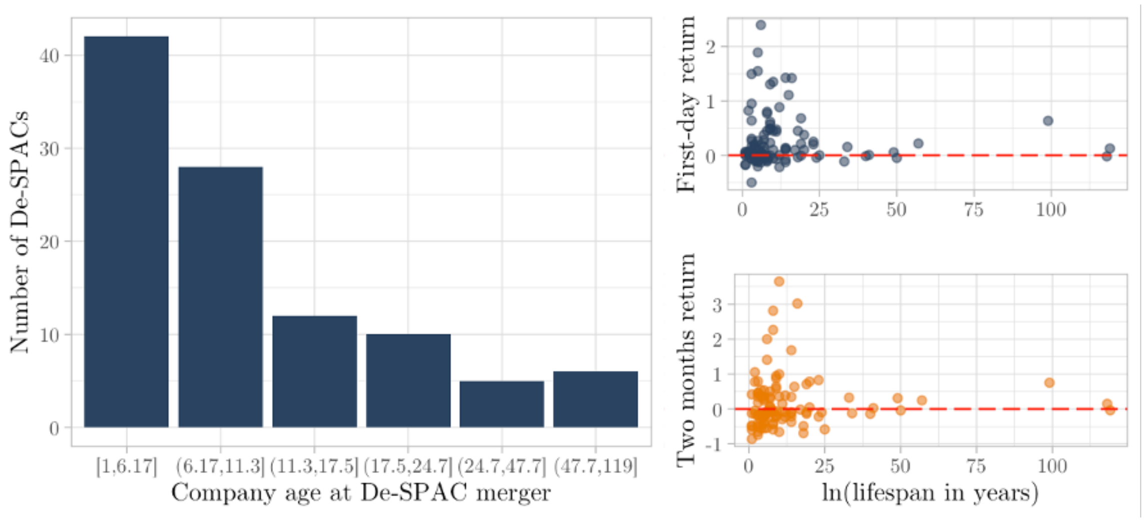

Figure 4.8: Plotting simple returns against Lifespan

4.2.8 Sector

The five sectors below are the five most common in our dataset. According to previous literature, "Healthcare" companies have tended to perform poorly in the short run compared to other industries. Contrary, we find that the sector has performed significantly worse on the longer time horizon (two months) than on its first day of trading. This suggests that the sector might be less important for the initial stock performance. The positive returns observed for the industries "Financial" and "Automotive" can largely be attributed to a few unicorns that have achieved more than 200% gain, such as $QS (365%), $LAZR (226%), and $NKLA (200%). Furthermore, the below chart 4.9 align with our expectations in hypotheses 1 and 2.

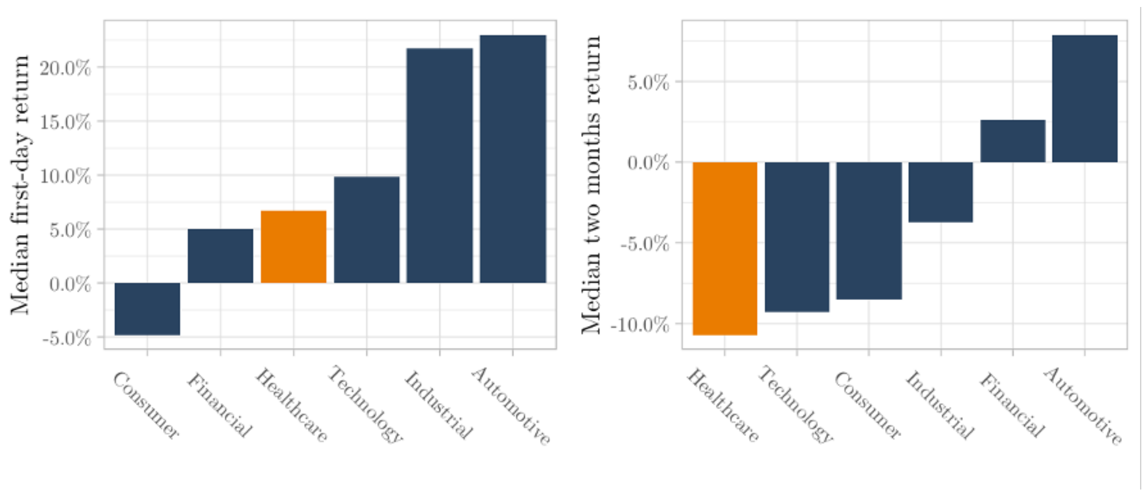

Figure 4.9: Plotting simple returns against Sector

5 Analyzing Individual Attributes of SPACs and Target Companies

We discussed the cleaned dataset in the preceding section. This chapter will provide a more in-depth examination of the individual characteristics related to our hypotheses. The goal is to find statistically significant and non-significant correlations between variables. To assess possible patterns, we will utilize regression analysis to see if there is any linear relationship between post-merger stock performance and the dataset predictors.

5.1 Method: Multiple Regression Analysis

In order to investigate significant relationships between post-merger performance and certain attributes, we use the multiple regression analysis and the ordinary least squares (OLS) method. The OLS method is easily applied to estimate the multiple regression model. Each slope estimates the partial effect of the corresponding independent variable on the dependent variable, holding all other independent variables fixed (Wooldrigde, 2012).

5.1.1 Model Formulations

Our regression models will take simple returns as the dependent variable, two-month, and first-day, respectively. In addition to the key attributes forming our hypothesis, we add control variables (two-month return for S&P 500 and Russell 2000) to establish a more casual relationship between our two-month dependent variable. All the below intuition follows Ritter (1991) and his paper on IPO performance.

Regression Model for the Two-Months Return

$$\begin{aligned} Return_{60} = \alpha &+ \beta_1 \ln \mathrm{Redeemed} + \beta_2 \ln \mathrm{PIPE} + \beta_3 \ln \textit{Total Assets} \\ &+ \beta_4 \mathrm{D}_{Profitable} + \beta_5 \mathrm{Lifespan} + \beta_6 \mathrm{D}_{Healthcare} \\ &+ \beta_7 \mathrm{S\&P500}_{60} + \beta_7 \mathrm{Russell2000}_{60} + \epsilon \end{aligned} \tag{5.1}$$

Regression Model for the First-Day Return

$$\begin{aligned} Return_1 = \alpha + \beta_1 \ln \text{Redeemed} + \beta_2 \ln \text{PIPE} + \beta_3 \ln \textit{Total Assets} \\ + \beta_4 \text{D}_{Profitable} + \beta_5 \text{Lifespan} + \beta_6 \text{D}_{Healthcare} + \epsilon \end{aligned} \tag{5.2}$$

5.1.1.1 Explanation for Model Variables

In the following section, we will briefly summarize each of the variables used in the regression models. Some of the variables are log-transformed in order to improve linearity and reduce the distance between the data points.

Dependent Variables

`Two-month simple return`: $Return_{60}$

The dependent variable is the raw simple return from the unit price to the first two-months after the closed merger. The intuition behind using raw returns (i.e., not log-transformed) follows Ritter (1991).

`First-day simple return`: $Return_1$

The raw simple return for the first trading day following the closed merger is the dependent variable return1. For the first-day return, the same intuition as previously is applied.

Independent Variables

`Shares Redeemed:` $\ln redeemed$

The shares redeemed prior to the merger is the redemption rate. This is a percentage number between 0% and 100%. The higher the percentage, the higher number of shares has been redeemed. A log transformation was applied for easier interpretation and improvement of linearity between variables (Wooldrigde, 2012). Additionally, since some of the observations are zero (i.e., all shares are retained upon merger), we set the value of these equal to 0.0001 to account for the fact that the log of zero is undefined.

`PIPE Investments:` $\ln pipe$

The variable PIPE investments is log-transformed due to skewness in the observations and for linearity improvements.

`Total Assets:` $\ln total\_assets$

The variable total assets is also log-transformed based on the same intuition as above.

`Dummy variable for profitable companies:` D_{prof1}

The dummy variable has a value of one when the company has a positive Return on Equity (ROE) the year prior to the merger. A value of zero is applied if the ROE is negative. As stated in previous chapters, we refer to these as “profitable” and “unprofitable” target companies.

$$D_{prof1} = \begin{cases} 1 & \text{if profitability} \equiv \text{Positive} \\ 0 & \text{if profitability} \equiv \text{Otherwise} \end{cases} \tag{5.3}$$

`Company age target company at merger:` $lifespan$

This variable describes the age of the target company at the merger. Following Ritter (1991), we dot not log-transform this variable.

`Dummy variable for healthcare companies:` D_{sector}

The dummy variable has a value of one when the company is classified as a healthcare company and a value of zero for non-healthcare companies. The classifications are matched up against Orbis and SPAC research.

$$D_{sector} = \begin{cases} 1 & \text{if sector} \equiv \text{Healtcare} \\ 0 & \text{if sector} \equiv \text{Otherwise} \end{cases} \tag{5.4}$$

5.1.2 OLS Assumptions

In order to meet the requirements of the OLS model, some key OLS assumptions must be satisfied in order to obtain convincing results (Wooldrigde, 2012). The key-OLS assumptions that must hold are listed below, in accordance with the Gauss-Markov Theorem (Theil, 1971):

- Linearity in the parameters
- Random Sampling
- No perfect Collinearity
- Zero Conditional Mean
- Homoscedasticity

For this thesis, the validity of our results must not violate these assumptions. In Appendix A2, we will go closer in detail, displaying the results of the various methods used to reveal potential violations. Below, we will shorty comment on the main conclusion drawn from various assumption testing methods.

To improve the linear relationship between the dependent and independent variables, we have log-transformed some of the predictors to account for the extreme values. Further, the residual plots in R. are used to check for linearity, homoscedasticity, and zero conditional means (Bartell, 2019). In order to account for potential multicollinearity, we have conducted a variance inflation factor test (VIF test) and analyzed the relationship between the variables in a correlation matrix.

The assumption of “Random Sampling” is somewhat more challenging to test. The residual plot Normal Q-Q shows the distribution of residuals across the model. If the residuals do not follow a normal distribution, this could be interpreted as a violation of this assumptions. However, the best way to ensure that this assumption is not violated is to follow good sampling techniques. In order to obtain a sample that is representative for the whole population, the data must be collected randomly from reliable sources. Thus, our data is collected from valid sources that are widely considered reliable. Consequently, the results from our testing are evident for the Gauss-Markow theorem being valid. Thus, according to Wooldrigde (2012), our OLS estimators are in line with being the Best Unbiased Estimators (BLUEs) for the model.

5.1.3 Regression Results

Table 5.1: OLS regression on the two-month return

	Dependent variable:						
	Two months return						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(redemption)	-0.102***						-0.089***
	(0.017)						(0.017)
log(pipe)		0.151***					0.075
		(0.056)					(0.053)
log(total_assets)			0.191**				0.055
			(0.085)				(0.082)
lifespan			-0.0001				-0.003
			(0.004)				(0.003)
sector				-0.352**			-0.212
				(0.171)			(0.145)
two_months_sp500					0.203		-0.686
					(2.796)		(2.472)
two_months_russell					4.012***		3.668***
					(1.309)		(1.163)
prof1						0.105	0.159
						(0.191)	(0.159)
Constant	-0.181**	-2.664**	-3.498**	0.270***	0.077	0.161*	-2.629*
	(0.090)	(1.052)	(1.641)	(0.086)	(0.088)	(0.084)	(1.442)
Observations	103	103	103	103	103	103	103
R^2	0.252	0.068	0.048	0.040	0.146	0.003	0.418
Adjusted R^2	0.244	0.059	0.029	0.031	0.129	-0.007	0.369
Residual Std. Error	0.665	0.743	0.754	0.753	0.714	0.768	0.608
F Statistic	33.953***	7.356***	2.545*	4.249**	8.558***	0.304	8.450***
Note:					*p<0.1; **p<0.05; ***p<0.01		

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with two-months return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant. The measurement of fit $R^2_{Adjusted}$ reports that our model (7) explain 36.9% of the variance in the two-months return.

Table 5.2: OLS regression on the first-day return

	Dependent variable:						
	First-day return						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(redemption)	-0.078***						-0.083***
	(0.010)						(0.010)
log(pipe)		0.087**					0.055*
		(0.036)					(0.032)
lifespan			-0.001				-0.003*
			(0.002)				(0.002)
sector				-0.160			-0.160*
				(0.110)			(0.087)
log(total_assets)					0.039		-0.083*
					(0.055)		(0.049)
prof1						0.155	0.280***
						(0.121)	(0.093)
Constant	-0.017	-1.383**	0.279***	0.303***	-0.492	0.232***	0.555
	(0.053)	(0.677)	(0.059)	(0.055)	(1.060)	(0.054)	(0.865)
Observations	103	103	103	103	103	103	103
R^2	0.365	0.056	0.002	0.020	0.005	0.016	0.469
Adjusted R^2	0.359	0.046	-0.008	0.011	-0.005	0.006	0.436
Residual Std. Error	0.392	0.478	0.491	0.487	0.490	0.488	0.367
F Statistic	58.149***	5.943**	0.230	2.090	0.508	1.636	14.159***
Note:					*p<0.1; **p<0.05; ***p<0.01		

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with first day return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant. The measurement of fit $R^2_{Adjusted}$ reports that our model (7) explain 43.6% of the variance in the first-day return.

5.2 Findings

In the next subsections, we will elaborate on the findings from the regression results by presenting each of our predictors and their associated effect on the dependent variables. To illustrate, we divide our Merger Cohort into two groups based on specified thresholds based on variable summary data. Furthermore, we comment on the effect of each predictor on the dependent variables in light of their statistical significance. Finally, we discuss how the various predictors act in relation to our hypotheses. Our findings is highlighted and discussed in section 5.3.

5.2.1 Redemption Rates

We uncovered a few signs of a linear relationship between the redemption rate and the initial return of De-SPACs in the descriptive analysis. In this segment, we broaden the scope and delve deeper into their relationship. To adequately shed light on the significance of redemption rate, we divided our Merger Cohort into two groups based on how much liquidity their SPAC managed to retain; (1) low rate of redemption, and (2) high rate of redemption, with the threshold set at 50%. The group with low rate is classified as "high quality" and the group with high rate is classified as "low quality". As a result, 62 companies out of 103 are categorized as high quality. The figure 5.1 below illustrates the performance of our merger cohort:

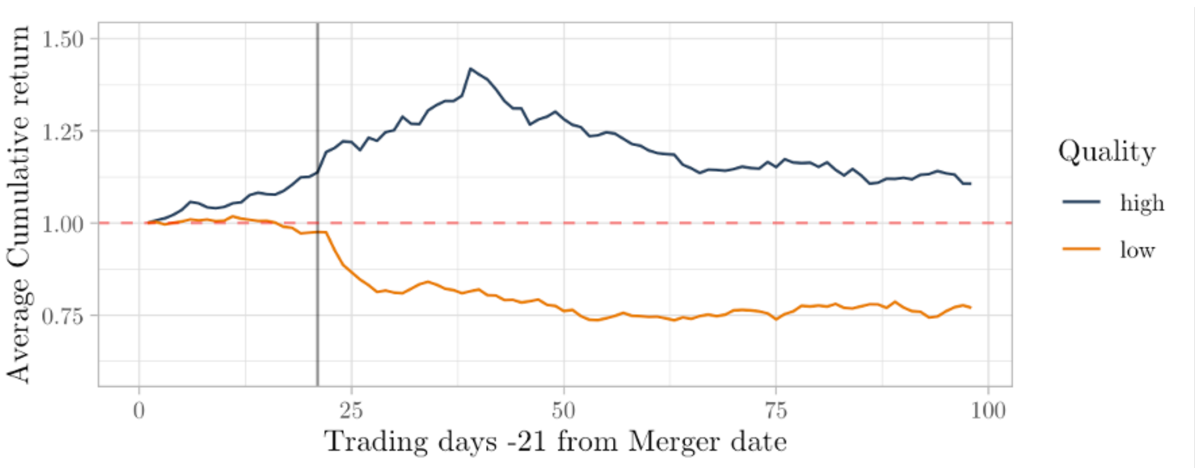

Figure 5.1: High versus Low redemption rate

The chart depicts the cumulative return indexed from 21 days prior to the merger date. This is to capture the price movement following the merger announcement and the market reaction to the redemption rate published in SEC filings. The market reacted favorably to high-quality firms both before and after their De-SPAC merger, outperforming their peers by a large margin. Stock price movements led the average cumulative return to a 40% gain over its first two trading months, whereas low-quality firms recorded a 24% loss during the same time frame. One possible reason could be that redemption rate depletes the target firm's cash holdings, which are required to fund positive cash flow initiatives. As a result, if the gap between promised and actual capital given is large, the target firm may not have enough capital to fund its planned profitable venture and generate excess value to the firm.

We expected a greater redemption rate to have a significantly negative effect on two months return based on hypothesis 1 (H1). The multiple regression results in table 5.1 support our hypothesis. At the 1% significance level, the independent variable $log(redeemed)$ is significantly negative in both uncontrolled and controlled conditions, indicating a strong negative linear relationship with the dependent variable two months return. The results further show that a 10% increase in redemption rate decreases the two months return by 0.0089 percentage points ($\frac{-0.089}{100} * 100$), all else equal.

Interestingly, concerning our hypothesis 2 (H2), we expected $log(redeemed)$ to be positively correlated with the first-day return. However, the multiple regression table 5.2 shows a significant negative relationship between the variable and first-day return. A 10% increase in redemption rate will, according to the model, decrease the first-day return by 0.0083 percentage points. The variable is statistically significant at a 1% level in both uncontrolled- and controlled-fitted models. Conclusively, the redemption rate has a significant negative linear relationship with both the two months and first-day return. The empirical result supports our H1, but invalidates our H2.

5.2.2 PIPE Investments

In order to conduct further analysis on PIPE investments, we divide our Merger Cohort into two groups based on whether they have raised PIPE investments above or below \$325mUSD. This represents the top quantile of our cohort. Figure 5.2 below displays the development in performance for the respective groups. It is evident that De-SPACs with PIPE investments above our threshold perform notably better than others. After two months of trading, the returns for more-PIPE-founded and less-PIPE-founded are 21% and -5.6%, respectively.

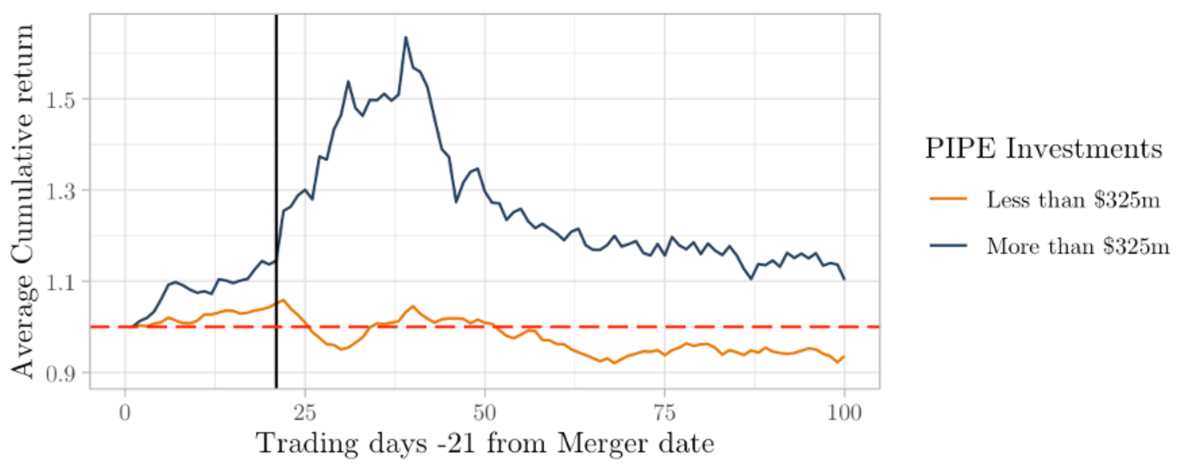

Figure 5.2: More funded versus less funded

Unsurprisingly, the more PIPE investments a company raises, the better the performance. Therefore, in accordance with hypothesis 1, we expect that low PIPE investments are associated with lower two-month returns. The results from our regression 5.1 follow this intuition. In an uncontrolled environment, the variable $log(pipe)$ is significant at a 1% level – implying a strong positive correlation with the two-month return. However, the results are not significant on any level when tested in a controlled environment (7).

Regarding hypothesis 2, we expect the opposite correlation – implying a negative relationship between PIPE investments and first-day return. Nevertheless, and similar to what we found on redemption rates; the PIPE investment acts in the same manner as in the two-month perspective. While this result is significant in an uncontrolled environment, the results are non-significant on all levels in a controlled environment. The regression results suggest that a 10% increase in PIPE investments increases both the two-month and first-day return by 0.0075 and 0.0055 percentage points, respectively.

5.2.3 Total Assets

As for our independent variable total assets, we divide our merger cohort into two groups based on whether the De-SPAC raised more than $255 million during the merger process or not. The exact amount is determined to differentiate the top quantile from the remainder of the cohort, rounded up to the closest $10 million. As a result, 26 companies managed to raise more than $255 million in their De-SPAC process. The figure below illustrates the group's performance

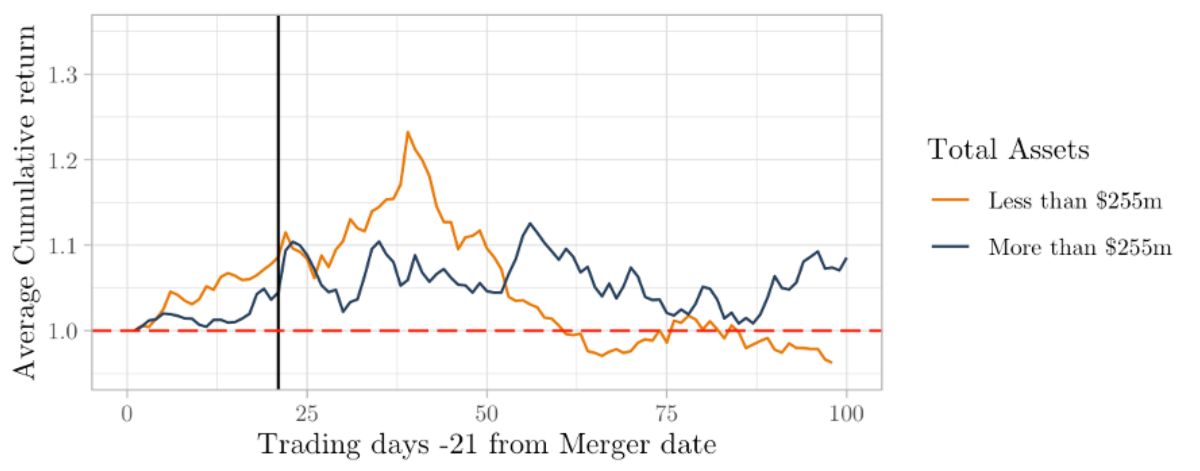

Figure 5.3: More versus less than $255m

Unlike previous independent variables, we cannot see a definite winner. Within the time horizons of interest, the group that raised less than $255 million outperformed the top quantiles. This group and the top quantile group achieved 11% and 9% average cumulative first-day return, respectively. Nevertheless, the top quantiles experienced less volatility in the longer time horizon and never approached the negative territory in cumulative terms. Ultimately, it is not evident from the figure that *total_assets* have a significant effect on the stock performance, as the groups yielded similar results.

In accordance with our hypothesis 1, we anticipated that the more total assets a SPACs raises, the better the two-month return. This is true and significant when tested in an uncontrolled environment, in table 5.1. In the controlled model (7) the coefficient suggests that an increase in the total asset of 10% will increase the two-month return by 0.0055 percentage points. The variable, however, is not significant on any level. A possible explanation for this result could be attributed to the absorption of explanatory power caused by including other variables. We suspect that the effect of total assets has been absorbed by the redemption rate, due to an overlooked causal relationship between the two variables. In order to investigate this further, we will present a complementary analysis in the summary and discussion section 5.3.1.

The regression result on the first-day return conveys a somewhat conflicting story. The variable total assets have a significant negative linear relationship with the first-day return in the controlled model (7) in table 5.2 - proposing that more capital is linked with a lower first-day return. The coefficient in the model can be interpreted as follows: if the De-SPAC raises 10% more capital, then the first-day return will decrease by 0.0083 percentage points. This is in line with our hypothesis 2. The opposite prefix applies when tested in an uncontrolled environment, though insignificant.

5.2.4 Profitability

We divide our Merger Cohort into two groups based on whether the companies had a positive or negative return on equity, the year before the merger. Consequently, we end up with a total of 20 profitable firms and 83 unprofitable firms. The results illustrate that the profitable firms tend to outperform the unprofitable firms, except for a short period of time.

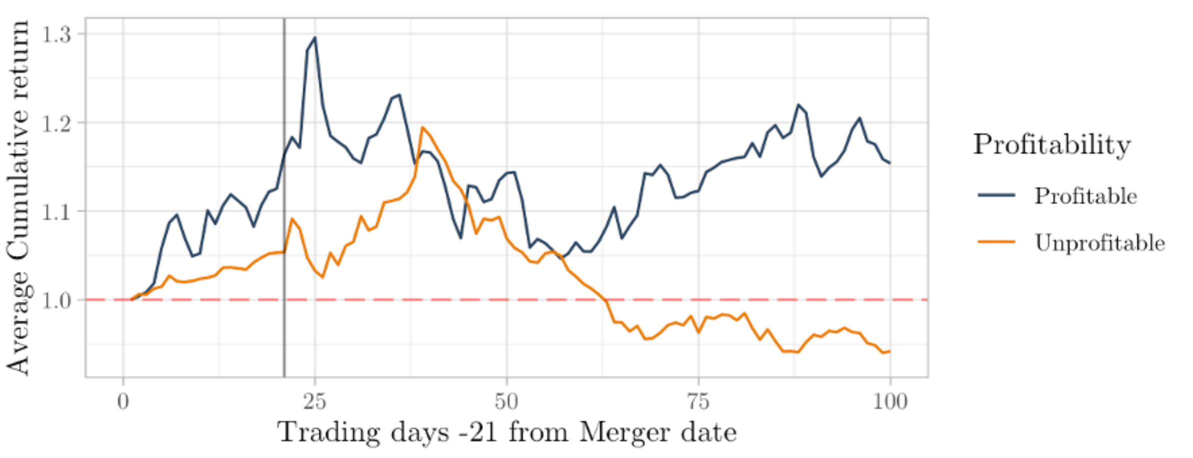

Figure 5.4: Profitable versus Unprofitable firms

In line with our hypothesis 1, we expect profitable firms to outperform unprofitable firms two months after merging. The results from our regression table 5.1 show that the dummy "profitability" has a positive prefix – implying that there indeed is a positive relationship between two-month returns and the profitability of the target company. Moreover, the coefficient in the controlled environment (7) can be interpreted as profitable target firms ($Dummy = 1$) will increase the two-month return by 0.159%. However, the relationship is not statistically significant on any level in any environment.

Table 5.2 displays the effect of profitability on first-day returns. According to our hypothesis 2, we expect profitable target companies to correlate negatively with first-day returns due to the underpricing discussed in the literature review. However, the results show the opposite. Our model finds a statistically significant positive relationship between profitable companies and initial return in the controlled model. The coefficient in the regression model suggests that if a company is profitable, the first-day return increases by 0.280%. Further, the results are significant on a 1% level.

5.2.5 Lifespan

Similar to the other variables, we separate the predictor “lifespan” into two groups: target companies with ages above eight years at the merger date and target companies below the age of eight years (i.e., above or below the median). Surprisingly, as figure 5.5 below illustrates, there seems to be a clear trend in De-SPACs merging with an older group yields poorer performance – both in terms of first-day and two-month returns. As a matter of fact, the average two-month returns for the groups are -9.3% for “older” and 11% for “young”.

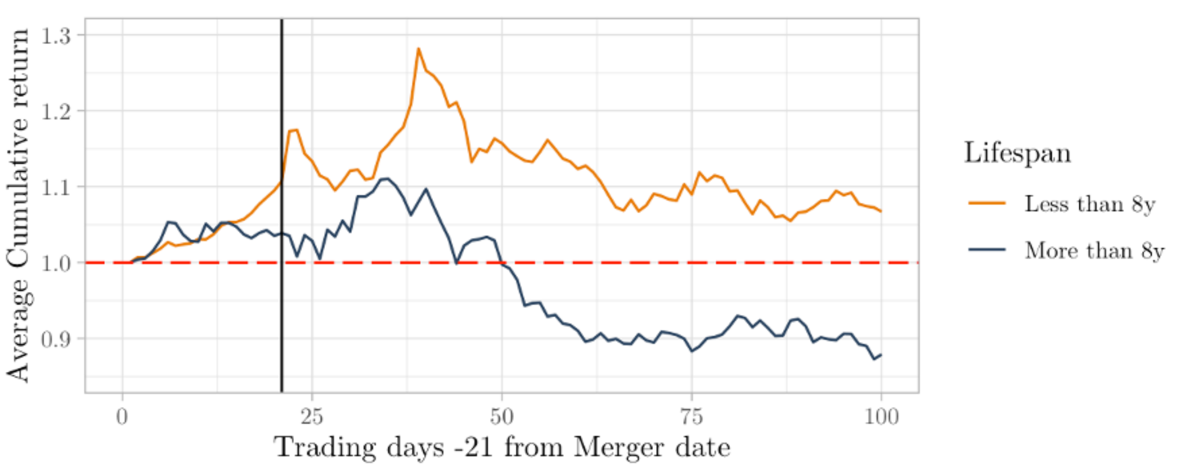

Figure 5.5: Older versus Younger firms

The above graph contradicts our hypothesis 1, where we assume that the younger the target company, the poorer its performance, as to Ritter (1991). This is based on our assumption that younger firms have a shorter financial history and little to no revenues – making the valuation of them more prone to error, since they are based on mere projections. As of the regression results in table 5.1, the prefix is negative when tested in both a controlled and an uncontrolled environment (implying: the younger, the better). Moreover, if the lifespan increases by one year, the two-month return decreases by 0.003% on average, all else equal. However, none of the relationships are statistically significant on any level.

The results from table 5.2 are similar to what we found in the first regression. In line with our hypothesis 2, in which we expect that a shorter lifespan has a significant positive effect on the first-day return, the results convey this sentiment. The fitted controlled model shows that first-day return decreases by 0.003% when lifespan increases by one year.

One likely explanation for the above findings could be that, at the time, the market favored young and innovative companies with high growth potential (Kunthara, 2021). Thus, older private companies that go public might not achieve the same potential, even though they manage to raise capital for their ventures. Additionally, we also suspect that the data might inherently be skewed, as most of the De-SPAC target companies are quite young.

5.2.6 Healthcare Sector

Our dataset consists of 26 healthcare companies and 77 non-healthcare companies. As figure 5.6 below shows, healthcare companies are remarkably outperformed by the other sectors. This is evident for both first-day and two-month returns, indexed from the unit price. Post-merger return after two months is 6.4% for healthcare companies and 8% for non-healthcare. In the longer run, however, healthcare companies achieve comparable returns to non-healthcare companies.

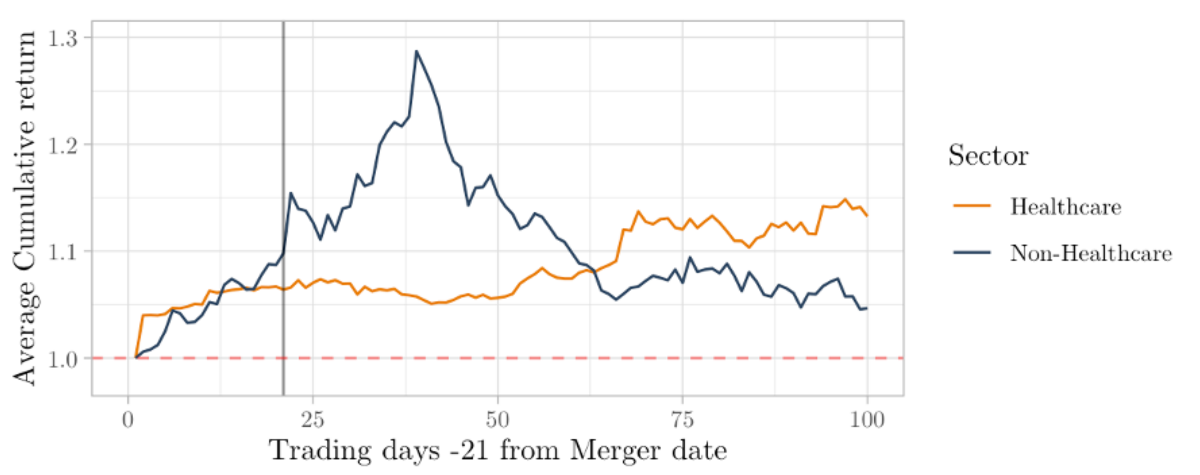

Figure 5.6: Healthcare versus Non-healthcare companies

The regression results from table 5.1 show that status as a healthcare company seems to have a significant negative impact on the two-month returns. However, in a controlled environment (7), the effect fads away, suggesting that the variable is insignificant when other predictors are included. The coefficient in the model suggests that if a company is categorized as a healthcare company (i.e., $Dummy = 1$), then the two-month return decreases by 0.212% compared to non-healthcare companies, on average.

In our Merger Cohort, 30% of the target companies operate within the healthcare sector. Exploring deeper into this group, we find that 65% (29 companies) of these are biotechnology companies with an average lifespan of seven years. Furthermore, during the Covid-19 pandemic, there was a strong demand for biotech solutions (DeFinance, 2021). Young and innovative Biotech companies wanted to swiftly capitalize on the opportunity of funding before it disappeared. SPACs allowed them to consummate an acquisition in a matter of months, as opposed to the year or more that standard IPOs typically take. Thus, one possible explanation for the relatively poorer performance could be revolved around the fact that these firms are pre-revenue companies (yet to generate sales) that rely on (for example) a patent to succeed. Potentially, this can make the due diligence of healthcare companies more difficult, possibly leading to poorer valuations from investors.

As of regression table 5.2, the results interpret that healthcare companies seem to harm first-day returns too. If the target company is classified as a healthcare company, the first-day return decrease by 0.160%. The results are statistically significant at a 10% level in a controlled environment. Contrary, in an uncontrolled environment, the results are non-significant. As for our hypotheses (1 & 2), we once again accept the first and reject the latter. Again, however, this can be interpreted as other in-the-model predictors absorbing the effect.

5.3 Summary and Discussion

In this chapter, we have analyzed the relationships between post-merger stock performance and our designated key attributes: shares redeemed, amount of PIPE investment, amount of total asset raised, and characteristics of the target company.

In line with hypothesis 1, the regression results stipulated that many of the key attributes indeed are associated with the two-months performance. The predictor with the highest statistical significance was the $log(redeemed)$-variable. The results propose the idea that high redemption rates contribute negatively to the two-month returns for De-SPACs. This is consistent with findings from Klausner et al. (2020). Other variables also showed intuitive prefixes, but these independent variables were not significant in a controlled environment. Despite the lack of statistical significance, most of the results align with our hypothesis 1.

As for hypothesis 2, the results were surprising. The hypothesis was based on the fact that De-SPACs obtain a pop similar to the "IPO-pop" on the first day of trading – implying that underpricing is present. Thus, we expected the key attributes to correlate opposite with the first-day return, as opposed to the two-month return. This expectation was motivated by Ritter (1991) and his findings on "[...] pattern in the other direction." However, the results from the regression contradict this expectation. Moreover, the predictors yield similar results to what we found in the two-months regression – some even more significant. The only independent variable that showed an inverse relationship was the total assets raised. Ultimately, this portrays that the attributes dramatically affect the initial performance of De-SPACs too, and that investors indeed emphasize them.

While most parts of our hypotheses are based on research on SPACs, some attributes are inspired by research on traditional IPOs. This applies to the characteristics of the target company: age at merger (*lifespan*), the profitability one year prior to the merger (D_{prof1}), and the sector (D_{sector})[7]. These predictors are inspired by Ritter (1991) and his acknowledgeable study on the long-run performance of IPOs. Interestingly, however, the results from our regressions deviate from what Ritter finds. The aforementioned predictors do not correlate differently in the two time periods of interest. Moreover, our results imply that younger target companies are associated with higher returns.

5.3.1 Analysing the Effect of the other Predictors on Redemption Rate

The last thing we want to touch upon is the fact that our results seems to be driven by the variable $log(redeemed)$. We suspect redemption rate to absorb the effect of

[7] We have included the independent variable D_{prof1} in our hypothesis based on research on the topic for both De-SPACs and IPOs.

other predictors, as it concurs concurrent with the merger, while others are already determined (ex-ante). The only unknown factor for shareholders deciding to redeem or retain their shares, is the PIPE investments (as of figure 2.1). Thus, the balance sheet fundamental (total assets) and the characteristics of the target company (sector, lifespan and profitability) are already known for the shareholders to determine whether to redeem their shares or not. Hence, it could be that the redemption rate is determined by the other predictors. For example, is it plausible that a rational shareholder will redeem her or his shares if the total assets raised are low, and if they observe poor characteristics with the target company? We will, in the following, perform different models seeking to answer these questions. Furthermore, our original models excluded redemption rate will be presented in Appendix A3.

5.3.1.1 Re-running the Model with Logit-Transformed Redemption Rate

In an effort to assess whether redemption rates absorb the effect of the other ex-ante predictors, we logit-transform the variable. Logit transformation is deemed helpful to stabilize the variance of proportionate and percentage variables that are not binomially distributed (Holland, 2017). Redemption rates in our dataset are transformed as such:

$$\text{Logit(y)} = log\left(\frac{\text{y}}{1-\text{y}}\right) \tag{5.5}$$

To investigate the effect, we regress our dependent variables on the transformed redemption variable. Moreover, we add the ex-ante predictors of the original models (excluding PIPE investments) to the new models (1) and (2). In addition, seeking to identify a pattern on whether shareholders will redeem or retain their shares, we add a model (3). This model uses linear regression with *logit_redemption* as the dependent variable and the aforementioned ex-ante variables as predictors.

The results from table 5.3 are entrancing, suggesting that redemption indeed absorbs some of the effects from the other independent variables. Compared to the regression models in tables 5.1 and 5.2, the below model (1) and (2) yield additional significant predictors. As of model 1 (two-month return), the independent variables *sector* and *prof1* are now significant. Model 2 (first-day return) now yields significant correlations for *lifespan*, *sector*, and *prof1*.

Model (3) also appears to yield interesting results. Total assets is statistically significant on a 5 percent level. Moreover, the negative prefix implies a negative correlation between redemption rates and total assets raised. Thus, the more capital a De-SPAC manages to raise, the lower the associated redemption rate. In an economic interpretation, a shareholder observing that the SPAC manages to raise considerable total assets (IPO proceeds) will be more likely to retain her or his shares upon merger.

Table 5.3: OLS regression with Logit trasformed Redemption Rate

	Dependent variable:		
	two_months_return_unit	first_day_return_unit	logit_redemption
	(1)	(2)	(3)
logit_redemption	−0.091***	−0.075***	
	(0.015)	(0.009)	
log(total_assets)	0.043	−0.067	−1.257**
	(0.078)	(0.044)	(0.510)
lifespan	−0.003	−0.004**	−0.027
	(0.003)	(0.002)	(0.022)
sector	−0.271*	−0.156*	−0.050
	(0.152)	(0.087)	(1.027)
prof1	0.277*	0.294***	1.753
	(0.163)	(0.093)	(1.085)
Constant	−0.848	1.387	21.530**
	(1.500)	(0.857)	(9.903)
Observations	103	103	103
R^2	0.330	0.465	0.107
Adjusted R^2	0.295	0.437	0.071
Residual Std. Error	0.642 (df = 97)	0.367 (df = 97)	4.343 (df = 98)
F Statistic	9.554*** (df = 5; 97)	16.835*** (df = 5; 97)	2.945** (df = 4; 98)
Note:			*p<0.1; **p<0.05; ***p<0.01

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with different dependent variable. F-statistics show that the joint effects of our variables are statistically significant. The measurement of fit $R^2_{Adjusted}$ reports that our model (3) explain 7.1% of the variance in the two months return.

5.3.2 Complementary Models

Based on the above results (regression table 5.3), we will in the following perform different supervised machine learning approaches where we predict the redemption rate on a trained dataset using the other variables (except PIPE investments) as predictors.

5.3.2.1 Tree-based Models

By using tree-based models, such as regression trees and Random Forest, we are better able to understand whether the aforementioned independent variables from our original regression can predict redemption rates. These models involve segmenting the predictor space into a number of simple regions. Each successive segmenting adds some complexity to the model, which can be used to make predictions (James et al., 2021). The decision tree illustrates the prediction of redemption rates and the associated probability:

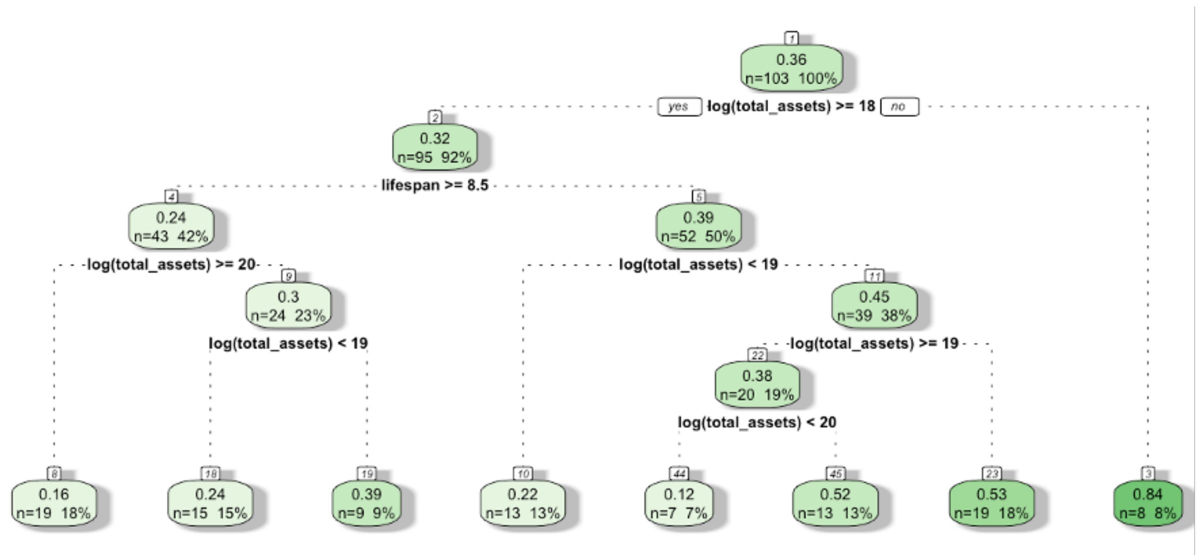

Figure 5.7: Decision Tree

As shown, the most important predictors of redemption rate are the $log(total\_assets)$ and the $lifespan$ variables. For instance, if the De-SPAC raises $log(total\_assets)$ below 18, then the associated redemption rate is 84%, with the probability of 8 percent. If, however, the De-SPAC manages to rise above 18, merges with a target company older than 8.5 years, and again has above 20 in $log(total\_assets)$, the associated redemption rate is 16% with a probability of 18%. Looking only at the two first branches, if the De-SPAC manages to rise above 18 in $log(total\_assets)$, there is, according to this model, a 92% chance that "only" $\frac{1}{3}$ shareholders will retain (not redeem) her or his shares.

In order to substantiate and validate the results from the above computations, we have performed a Random Forest and investigated the associated measure of variable importance based on the Gini impurity index used for calculating the splits in trees. The higher the value of the mean Gini score, the higher the importance of the variable is to our model (Tat, 2017). The associated Gini score of our model yields results interpreting that $log(total\_assets)$ and *lifespan* indeed are the variables of most importance, followed by *sector* and *prof1*. The Gini Index is presented in Appendix A4.1.

5.3.2.2 Ordinal Logistic Regression

Ordinal Logistic Regression (OLR) is an extension of the simple logistic regression model (R-bloggers, 2019). The model aims to classify an ordinal dependent variable (i.e., explicit ordering in the categories) based on sets of predictors. Thus, our redemption rates are split into three ordinal groups: (0) low, (1) modest, and (2) high redemption rate. The OLR summary is presented in Appendix A4.2. We deploy the model to quantify the log-odds of observing high redemption rates predicted by our independent variables. Illustratively, OLR obtains the probability of observing a redemption rate less than 50% by:

$$LogOdds_{Redemption<50\%} = log\left(\frac{p(Redemption = 0.5)}{p(Redemption > 0.5)}\right) \tag{5.6}$$

OLR yields similar results as the previous complementary models, where total assets, lifespan, and sector are the dominant predictors. The figure 5.8 below displays the joint effect of total assets and sector on redemption rate. As total assets increase on the X-axis, the probability of observing a high redemption rate (2) falls dramatically for both non-healthcare and healthcare companies. We observe the opposite for the modest redemption rate (1), where the probability increases as the total assets increases. From the regression result, we find that De-SPACs associated with the healthcare sector are 36.89% more likely to have a high redemption rate, relative to non-healthcare. Further, for every one-unit increase in log total assets, the likelihood of observing a firm in the high redemption rate group decreases by 21.78%. Our model is validated by a Brant test (Brant, 1990), which ensures that the coefficients do not differ across different cut points in the outcome variable, see Appendix A4.2.1.

In a predictive model using OLR, the misclassification error yields 27.49%. In other words, our model managed to classify 72.51% of the redemption rate groups correctly with the independent variables.

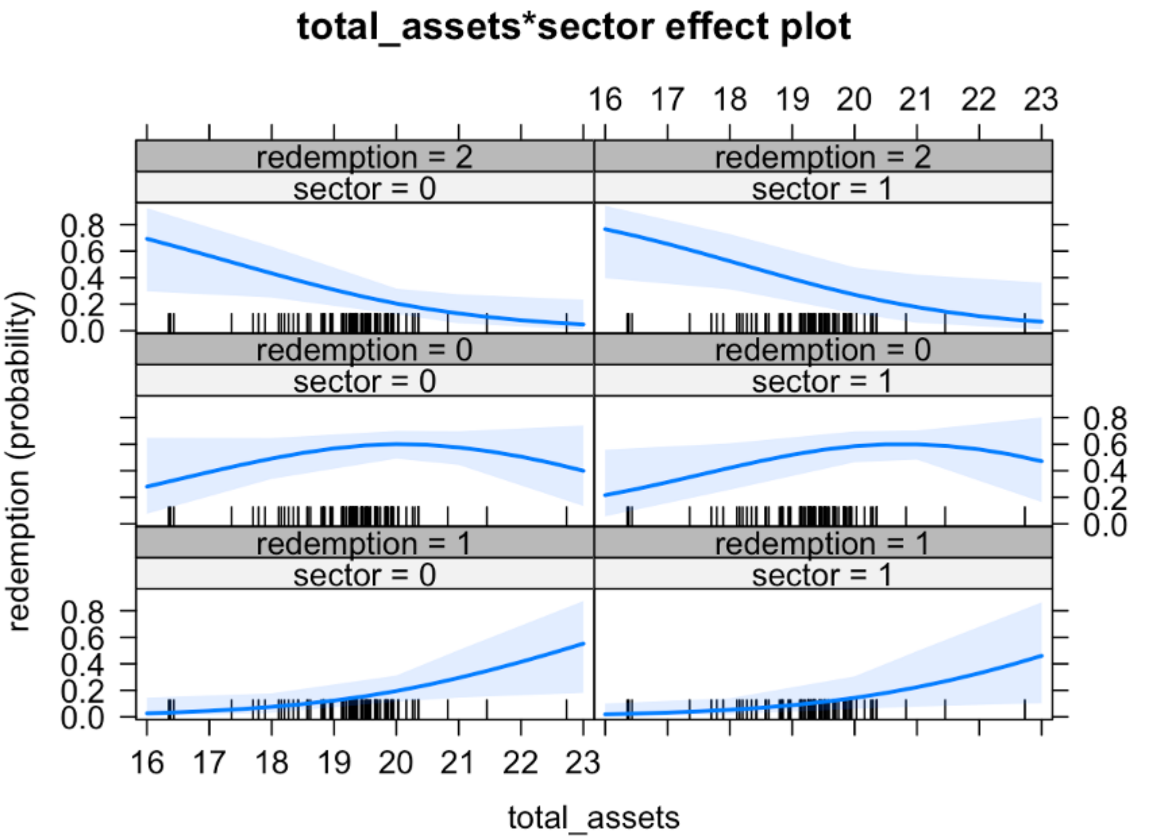

Figure 5.8: Ordinal Logistic Regression Effects

5.3.2.3 Summary and Discussion of Complementary Models

The results of our complementary models further substantiate the notion that independent variables may have a causal effect on the redemption rates. This suggests that certain financial indicators and characteristics of target companies affect investors' behavior in the run-up to the merger. According to our models, the independent variables of most importance are total assets raised, lifespan, and sector of the target company. The effect of profitability is less evident.

Total asset is a key financial metric made public by the SEC. An informed investor can, therefore, based on the amount the SPAC manage to raise, determine whether they will redeem their shares or not. According to our results, it seems like certain investors indeed are informed – redeeming their shares if the total assets raised are below some threshold.

From late-2020 to early-2021, the market favored early-stage companies with high growth potential. Statistically, most pre-revenue businesses fail to meet their growth projections. In aggregate, they may be attractive investments driven by a few eventual big winners, but there will be more disappointments than successes. In mid-2021, we saw a shift in sentiment, and redemption rates among early-stage companies rapidly rose (Ploeg, 2022). Investors may have developed a keen eye for identifying quality firms based on certain characteristics, favoring firms with better fundamentals and publicly available information. Our findings support that target companies' age, sector, and profitability all play a vital role in determining whether investors choose to redeem, though with varying degrees.

Profitability seems to be of less importance compared to the other predictors. This is intriguing, suggesting that shareholders tend to put less weight on the profitability of the target company. One likely explanation is that shareholders expect mergers would reverse unfavorable trends and that sponsors will provide future guidance for increased profitability. Nevertheless, the results from previous models on the performance show that SPACs merging with unprofitable targets are significantly outperformed by SPACs merging with profitable targets. Thus, the optimism described above can illustrate irrational investors and the endowment effect – an emotional bias that causes individuals to be overoptimistic about an owned object (Ganti, 2021).

The associated prediction accuracy of 72.51% is impressive at first glance. However, we must consider that some randomness could drive the models. Put differently, we cannot be entirely sure that observed shareholders are "skilled" based on the obtained results. Moreover, investors might have incentives to redeem their shares independent of the predictors above. In an effort to follow a common thread, we will implement the obtained predictions of redemption rates in the original regression models from tables 5.1 and 5.2.

According to the table 5.4, our predictive variables *prediction* is non-significant. However, the results are in line with the previous complementary models. "Prediction 2" and "prediction 3" represent predictions of "modest" and "high" redemption rates. The negative correlation in both models implies that both two-months and first-day return is lower for the dummy variables "modest" and "high", relative to the reference group "low." Intuitively, the *prediction3* variable's coefficient is of higher value than *prediction2* – implying that the return for the high redemption group is lower compared to its peers.

Table 5.4: OLS regression with predicted Redemption Rates

	Dependent variable:	
	two_months_return_unit	first_day_return_unit
	(1)	(2)
redemption	-1.139^{***}	-0.827^{***}
	(0.193)	(0.118)
prediction2	-0.148	-0.189
	(0.504)	(0.309)
prediction3	-0.219	-0.335
	(0.637)	(0.391)
log(pipe)	0.811	0.674
	(0.955)	(0.586)
log(total_assets)	-0.466	-2.466^{*}
	(2.069)	(1.269)
lifespan	-0.004	-0.005^{*}
	(0.005)	(0.003)
prof1	0.248	0.244^{**}
	(0.166)	(0.102)
sector1	-0.311^{*}	-0.161
	(0.169)	(0.103)
Constant	-0.165	6.149^{*}
	(5.941)	(3.642)
Observations	103	103
Adjusted R^2	0.295	0.351
F Statistic (df = 8; 94)	6.330^{***}	7.907^{***}
Note:		*p<0.1; **p<0.05; ***p<0.01

F-statistics show that the joint effects of our variables are statistically significant. The $R^2_{Adjusted}$ reports that our models (1) and (2) explain 29.5% and 35.1% of the variance in the two-months and first-day return, respectively.

6 Analyzing "Hype" Indicators

In this chapter, we will introduce a more targeted analysis of our hype indicators, Google Search Hits and Reddit Mentions, associated with our hypothesis 3 and the first-day return. First, a brief general descriptive analysis of our hype predictors is introduced. Then, we will present our regression method and the associated results and findings.

6.1 General Descriptive Analysis

6.1.1 Google Search Hits and Reddit Mentions

Using Google and Reddit APIs[8], we can see if there are any patterns between forum activity, internet search frequency, and our De-SPAC cohort stock prices. As previously stated, Google Hits show the relative popularity of specific keywords on the Google Search Engine. Reddit Mentions represent the frequency in which a term is mentioned on Reddit. The average Google Hits is calculated to be 52, with 64 De-SPACs over the 50/100 threshold. The most "hyped" De-SPAC in the run-up to their merger consists of Billtrust (99.5), Nerdy (98.5), and Embark Trucks (91). Furthermore, the average number of Reddit Mentions is 1872, with Hims (78,239), Shift (20,751), and Finance of America (15,764) leading the chart by large margins.

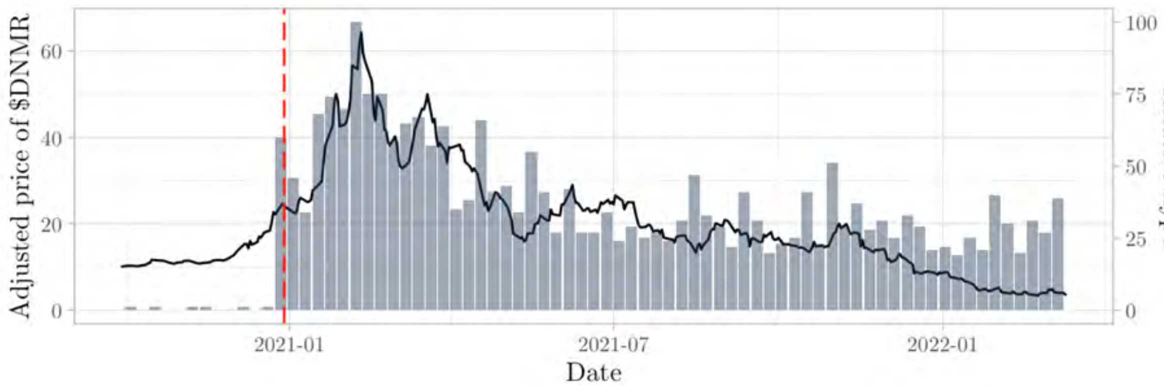

Figure 6.1: Google Hits and $DNMR

The graph 6.1 illustrates our intent regarding utilizing Google Hits as an indicator for stock prices, shown on ticker $DNMR. It is evident that Google Hits and stock prices

[8] API stands for application programming interface, which is a set of definitions and protocols for building and integrating application software.

are interconnected and correlated, suggesting that one might affect the other (i.e., search frequency drives up the price or the opposite). The red intercept line depicts their merger date.

The relationship between the variables and the first-day return is depicted in the figure 6.2 below. There are a few patterns in Google Hits that indicate a linear relationship between the first-day return and the independent variable. Intuitively, the relative popularity of a Google search query may explain both stock price changes leading up to the merger, and the first-day return. A similar tendency may seem present with Reddit Mentions.

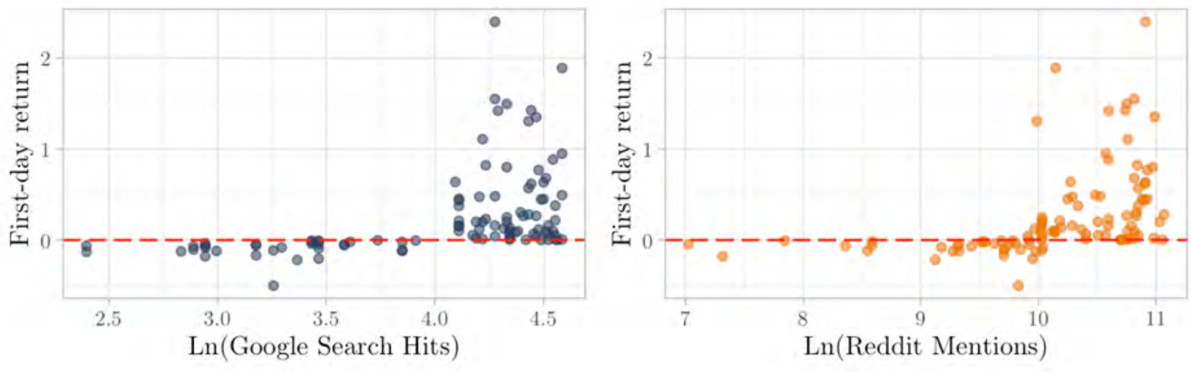

Figure 6.2: Plotting first-day return against Hype indicators

6.2 Method: Multiple Regression Analysis

To investigate significant relationships between first-day performance and our associated hype indicators, we will (based on the descriptive analysis) perform a multiple regression analysis.

6.2.1 Model formulations

Our regression models will take first-day return as the dependent variable, and Google Search Hits and Reddit mentions as independent variables.

$$\text{Return}_1 = \alpha + \beta_1 \text{log(google)} + \beta_2 \text{log(reddit)} + \epsilon \tag{6.1}$$

6.2.1.1 Explanation for Model Variables

`First-day simple return`: $first\_day$

The dependent variable is the simple return for the first trading day after the closed merger. It is indexed from the IPO offer price of $10.

`Google Search Hits`: $log(google)$

This independent variable is log-transformed to improve linearity.

`Reddit Mentions`: $log(reddit)$

This independent variable is log-transformed due to skewness in the observation and linearity improvement.

6.2.2 Regression Results

Table 6.1: OLS regression with Hype indicators

	Dependent variable:		
	first_day		
	(1)	(2)	(3)
log(google)	0.400***		0.264**
	(0.077)		(0.108)
log(reddit)		0.275***	0.139*
		(0.056)	(0.078)
Constant	-1.346***	-2.533***	-2.208***
	(0.311)	(0.575)	(0.577)
Observations	103	103	103
R^2	0.213	0.191	0.236
Adjusted R^2	0.205	0.183	0.221
Residual Std. Error	0.436 (df = 101)	0.442 (df = 101)	0.432 (df = 100)
F Statistic	27.271*** (df = 1; 101)	23.797*** (df = 1; 101)	15.482*** (df = 2; 100)
Note:			*p<0.1; **p<0.05; ***p<0.01

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with two months return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant and non-zero explanatory effect. The $R^2_{Adjusted}$ reports that our model (3) explains 22.1% of the variance in the two months return.

6.3 Findings

6.3.1 Google Hits

In the general descriptive analysis, we identified a rather evident and linear association between first-day return and Goolge Hits. We split our merger cohort into two unique groups, in the same fashion as in section 5.2. There are 52 De-SPACs (hereby the most hyped group) which have achieved average relative popularity of 72 on Google Trends. This value is the median number of relative popularity achieved by our merger cohort. These firms had an average first-day return of 46.7%, which was 41.2 percentage points above the other group.

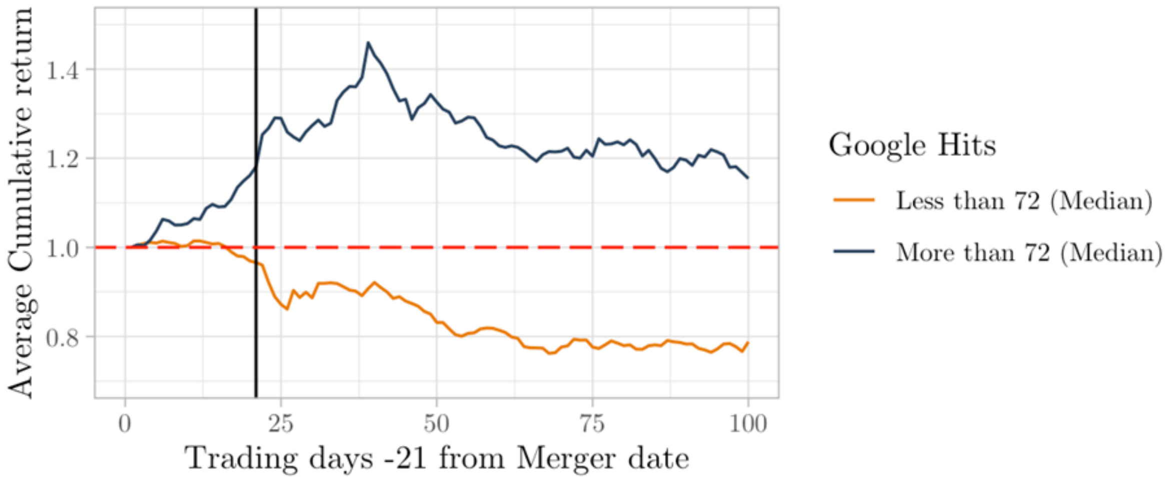

Figure 6.3: More versus less Google Hype

Regarding hypothesis 3, we expect that the hype indicator Google Hits would have a significant positive correlation with the first-day return. I.e., higher relative popularity of a search term on Google Trends is associated with greater first-day returns. In the regression table 6.1, the independent variable $log(google)$ has a significant positive impact on the dependent variable. This is consistent in both uncontrolled (1) and controlled (3) environments. The result of our level-log regression suggests that a 10% increase in the relative popularity in Google Trends increases the first-day return by 0.0264 percentage points. Statistically, the variable is significant at 1%, but when controlled for Reddit Mentions, the significance level is reduced to 5%. The uncontrolled and controlled fitted models obtained an adjusted R^2 of 20.5% and 22.1%, respectively.

A concern we raised upon conducting this analysis was the possibility of a circularity problem inherited between search term frequency and the stock price movements. Intuitively, the popularity of a search term may increase due to stock price volatility. This again could very well feedback to the stock price, circularly affecting each other as shown in the figure 6.4. Our goal is not to assess whether this is the case, but simply shed light on the linear relationship between Google Hits and the first-day return. Nonetheless, we took countermeasures to reduce the effects of the circularity problem. This is obtained by using lagged time series extracted from Google Trends two weeks before the merger date, instead of using simultaneous values.

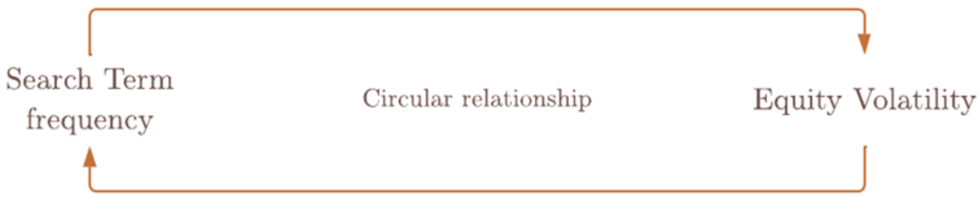

Figure 6.4: Circularity problem

6.3.2 Reddit Mentions

We divide the dataset into two categories based on whether firms are in the top 25% quantile, in terms of how many comments their relevant Reddit post received. The criteria is 47.000 comments, with 25 firms ranking in the top quantile. The graph 6.5 below depicts the performance disparities in each group:

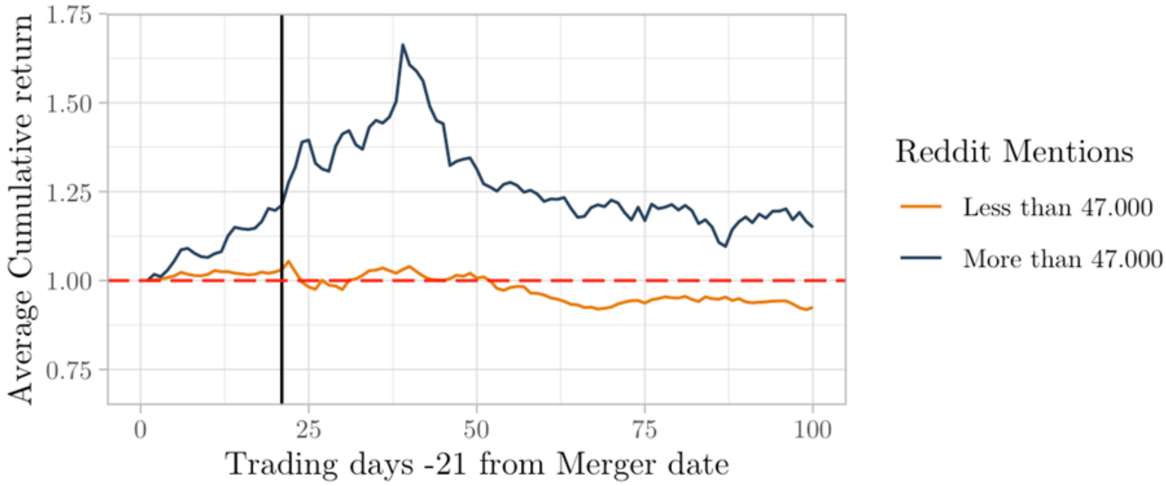

Figure 6.5: More versus less Reddit Hype

As expected, hyped De-SPACs performed well in the time horizons of interest and outperformed the less hyped group by large margins. The groups achieved 27.6% and 3.5% in average cumulative first-day returns, respectively. The regression table 6.2 clearly shows that Reddit Mentions is positively correlated with the first-day return, with a significance level of 1% and 10% in an uncontrolled (2) and a controlled (3) environment. The linear-log regression suggests that a 10% increase in Reddit Mentions two weeks before the merger, results in a 0.0139 percentage point increase in the first-day return. These findings are in accordance with our hypothesis 3.

While inspecting the dataset that we have extracted from Reddit, we saw that most of the posts and comments about our De-SPACs merger cohort stem from subreddits such as r/wallstreetbets, r/SPACs, r/stocks, r/investing, and r/Superstonk. The purpose of these groups is to empower retail investors with knowledge, share investing strategies, and function as a trading forum. The phenomenon dubbed "the rise of retail investors" (Reddel, 2022) in 2020 has its roots on Reddit, with GameStop's ($GME) short squeeze as their first big breakthrough. In short, the outlook for GameStop was dire and short-sellers such as Melvin Capital[9] bet heavily against the firm. At one point, 140 percent of the total float (issued shares) were sold short. Both retail and institutional investors then took the opportunity to squeeze short-sellers out of their position. Consequently, institutional and retail investors own 27% and 56% of GameStop, respectively (SimplyWallSt., 2022).

The battle between retail investors and Hedge funds gained large tractions on conventional media, and as a result, the market saw a massive inflow of both new investors and capital. It is estimated that in 2020, 15% of the market participants were first-timers (Reddel, 2022). Undoubtedly, Reddit is shown to have some level of market impact in the aggregate. This aligns with our findings that indicate a clear positive linear relationship between initial stock performance and forum activity.

[9]Following the battle between retail investors and Melvin Capital, the fund is shutting down in May of 2022. Since the GameStop frenzy, the fund has lost over 6 billion dollars under management. https://www.reuters.com/markets/us/hedge-fund-melvin-capital-tells-investors-it-plans-shut-down-letter-2022-05-18/

6.4 Summary and Discussion

In this chapter, we have analyzed the relationships between first-day stock performance and our "hype" indicators. As a starting point, we wanted to examine whether the recent hype for SPACs as "the poor-man's equity" could further explain the positive first-day returns.

The regression results imply a strong positive correlation between the dependent and both independent variables. The predictor of the highest statistical significance was *log(google)*. However, since data are weekly reported, this result could be prone to circular reasoning. This issue does not apply to the *log(reddit)* variable, which is also significant in both models. In order to account for any OLS violations, we have performed similar tests as with models 5.1 and 5.2. The results are in line with the Gauss-Markow Theorem (Theil, 1971).

The results are fascinating, and suggest that the recent hype for SPACs as an investment vehicle for retail investors is present and could further explain the positive first-day returns. While the interpretation of the results conveys modest returns, they are nonetheless significant. Moreover, when adding these predictors to the regression model associated with hypothesis 2, the adjusted R^2 increases. This model is presented in Appendix A3.3. While none of the hype predictors are significant in this new model, the increase in explanatory power is in line with our hypothesis.

7 Conclusion

We have examined the post-merger performance of 103 U.S. companies going public through a SPAC from 2020 to early 2022. We have investigated relationships between stock performance in the first two months of trading and key attributes of public available information. We hypothesized that: (1) elements of the net cash per share, (2) certain characteristics of the target company, and (3) indicators of hype, are all important for the performance. Moreover, based on initial trading gains found in other papers, we expected (1) and (2) to act contradictory on the first day of trading, compared to after two months of trading.

Our regression results revealed a statistically significant negative relationship between post-merger performance and the redemption rate. These relationships are evident in both models. PIPE investments yield a positive correlation with returns. However, the variable is only statistically significant in the model associated with first-day return. While total assets is shown to have a significant negative relationship with the initial performance, we uncovered the opposite correlation with the two-months return. Moreover, healthcare companies performed considerably worse than non-healthcare companies both on the first-day and two-month horizons. According to both models, profitable target companies outperformed non-profitable by large margins. Finally, we learned that young target companies tend to outperform more established target companies, statistically significant on the first day of trading.

The results from the original models described above suggested that the redemption rate absorbs the effect of the other independent variables. The variable was logit transformed and accounted for in a supplementary OLS. The model revealed that the redemption rate did, in fact, absorb the effect of other variables. Using machine learning models, we predicted if shareholders would redeem their shares upon the merger, based on the ex-ante predictors from the original model. Tree-based models suggested that total assets raised and the lifespan of a target company are of most importance to shareholders. Using Ordinal Logistic Regression, the model predicted the redemption rate with an accuracy of 72.51%. Based on these results, we added the predicted redemption rate in the original model to test for significance. While the predicted variables showed no statistical significance, the insight provided conforms with our previous discoveries.

In an effort to discover possible correlations between initial returns and “hype,” we deployed various data mining techniques incorporated with API interaction using R. Substantiated by the “Rise of Retail investors,” we sought to investigate the relationship between hype indicators and post-merger returns. Our hype indicators, two weeks lagged Google Hits (search frequency) and Reddit Mentions (forum activity), were proven to have a significant positive linear relationship with the first-day return. Despite the question of causality, the results suggest that retail investors do have an influence on the financial market in the aggregate.

The obtained results partially confirm our hypotheses. The results associated with hypothesis 1 are in line with most of our reasoning based on former studies on SPACs and IPOs. As of hypothesis 3, the results from the regression model suggest that “hype” arising from Google search frequency and Reddit mentions contributes to higher initial stock returns. The results associated with hypothesis 2, however, contradict what we expected to find — no inverse relationship is present. Thus, contrary to our initial beliefs, investors are shown to emphasize key attributes in their financial decision-making upon merger — disfavoring less attractive De-SPACs.

In a broader perspective, the great acceleration in the economy in the wake of the COVID-19 pandemic has pushed more companies into the public market. Given SPACs’ favorable advantages over traditional IPOs, more companies have used the “backdoor” into the market in order to swiftly capitalize on the opportunities available. In hindsight, however, the poor long-run performance of De-SPACs reveals that these advantages come with a bad aftertaste. While the hype foreshadows the lingering effects of the dilution, we find that the cost is greater than the benefits in the aggregate.

To summarize, pre-merger dilution in net cash per share appears destructive for both first-day and two-month returns. Moreover, certain characteristics of the target company are shown to have a prominent effect on the performance. Thus, to adjust for the declining trend in performance, De-SPACs ought to implement sustainable structural reformations if they are not to go entirely off track. Unless, of course, the hype saves them.

Suggestion for Further Research

Although our regression results yield some significant evidence for attributes affecting the post-merger performance of De-SPAC companies, further research in numerous fields can be conducted. First, it is possible to address the overall mediocre performance of the initial public market. Our results suggest that going public in recent years has been unfavorable for the target/operating company and its existing shareholders. Thus, a more in-depth examination of the IPO market, in general, could provide answers to whether going public is a wise strategy or not.

Second, due to regulatory differences between traditional IPOs and SPACs, it is harder for SPAC investors to follow up on their investments. We believe a more transparent disclosure of information on detailed data, such as SPAC cost, would benefit investors. This would make it easier to track where money is being diverted and to which sources (for instance, how much net cash per share is left to invest in the target firm). For our thesis, we stumbled upon such hurdle when collecting data on PIPE Investments and redemption rates. A more comprehensive disclosure would have yielded a greater sample size.

Third, our observation period of two months could have been expanded. As already argued, the reason for choosing such a short period of time is due to the fact that we have mergers from February 2022 – implying that any longer time period is impossible as for data collection. However, previous literature finds a continuous decline in performance. Thus, it is possible that a longer observation period may have yielded clearer and more significant results.

Fourth, we could have expanded our analysis on "hype" by including more hype indicators. For instance, we thought of including predictors based on hype on both Twitter and Tik Tok. However, we were not able to acquire access to the Twitter API or the Tik Tok API necessary to conduct this analysis. Moreover, a textual analysis of the sentiment in the different social media could have yielded more accurate results. Nevertheless, due to informal and rather vulgar language on the forums, the available "sentiment dictionaries" in R yielded poor results. Thus, an alternative would be to compute tailored dictionaries catered to each platform. However, this is out of the scope of our proposed thesis.

Finally, a more extensive dataset might have provided more robust results. Due to constraints caused by the availability of accounting data from SEC, some completed De-SPACs are missing. In addition, the limited time period looking at years when SPAC IPOs ruled over traditional IPOs has restricted the number of observations. It is possible that a longer time period, focusing on for example, third-generation SPACs, would have given us more clear results.

References

Agarwal, A. (2021). *A Primer on SPACs and PIPEs: How they work.* https://www.menabytes.com/spac-pipe/ [Accessed: 2022].

Bartell, E. (2019). *Code Through - Checking Assumptions for OLS.* https://www.rpubs.com/elliottb90/olsassumptions [Accessed: 2022].

Brant, R. (1990). Assessing proportionality in the proportional odds model for ordinal logistic regression. *JSTOR.*

Breiman, L. and Cutler, A. (2022). Random forests. https://www.stat.berkeley.edu/~breiman/RandomForests/ [Accessed: 2022].

Bryant, C. (2022). Spacs, the 'poor man's private equity,' made investors poorer. https://www.bloombergquint.com/gadfly/-spce-virgin-galactic-reveals-how-spacs-have-made-retail-investors-poorer [Accessed: 2022].

Chen, J. (2021). Excess returns. https://www.investopedia.com/terms/e/excessreturn.asp [Accessed: 2022].

Choueiry, G. (2021). What is an acceptable value for vif? https://quantifyinghealth.com/vif-threshold/ [Accessed: 2022].

Coates, J. and Hunter, P. (2021). Three thoughts on the rise of retail investors. https://www.sec.gov/news/public-statement/accounting-reporting-warrants-issued-spacs [Accessed: 2022].

DeFinance (2021). Definance etfs. https://www.defianceetfs.com/will-2021-be-the-year-of-the-biotech-spac/ [Accessed: 2022].

Fernando, J. (2021). Return on equity (roe). https://www.investopedia.com/terms/r/returnonequity.asp [Accessed: 2022].

Fink, A. (2022). Handbook of regression modeling in people analytics. https://peopleanalytics-regression-book.org/gitbook/ [Accessed: 2022].

Ganti, A. (2021). Endowment effect. https://www.investopedia.com/terms/e/endowment-effect.asp[Accessed: 2022].

Gigante, G. and Notarnicola, D. (2021). Performance analysis of healthcare-focused special purpose acquisition companies. *Investment Management and Financial Innovations.*

GoogleTrends (2022). Google trends. https://trends.google.com/trends/?geo=NO [Accessed: 2022].

Holland, S. (2017). Data analysis in geosciences. http://strata.uga.edu/8370/rtips/proportions.html [Accessed: 2022].

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021). *An Introduction to Statistical Learning with Applications in R.* Springer.

Kim, T. and Crystal, K. (2021). Spacs were hot in 2020 and are hotter now. here's why. https://www.bloomberg.com/professional/blog/spacs-were-hot-in-2020-and-are-hotter-now-heres-why/ [Accessed: 2022].

Klausner, M., Ohlrogge, M., and Ruan, E. (2020). A sober look at spacs. *Yale Journal*, 39.

Kunthara, S. (2021). Why are spac targets so young? https://news.crunchbase.com/news/vc-backed-spac-targets-vs-ipo/ [Publised: 9 December 2021][Accessed: 2022].

Lambert, J. (2021). Why so many companies are choosing spacs over ipos. https://advisory.kpmg.us/articles/2021/why-choosing-spac-over-ipo.html [Accessed: 2022].

Longoni, G. (2021). *SPAC Invasion and the advantages of "backdoor" listing*. https://www.bscapitalmarkets.com/spac-invasion-and-the-advantages-of-ldquobackdoorrdquo-listing.html [Accessed: 2022].

Orbis (2022). Overview. https://www.bvdinfo.com/en-gb/our-products/data/international/orbis [Accessed: 2022].

Panton, D. (2020). What is a spac and why are they suddenly so popular? *Interviewer: Brian Adams.*

PitchBook (2021). *A guide to every step in the IPO process.* https://pitchbook.com/blog/ipo-process-explained [Accessed: 2022].

Ploeg, E. V. (2022). Ver ploeg. https://everploeg.medium.com/tech-spac-redemption-rates-e8ff7d32633f[Accessed: 2022].

R-bloggers (2019). *Perform Ordinal Logistic Regression in R.* https://www.r-bloggers.com/2019/06/how-to-perform-ordinal-logistic-regression-in-r/ [Accessed: 2022].

Reddel, S. (2022). Three thoughts on the rise of retail investors. https://capitalmarketsblog.accenture.com/three-thoughts-rise-retail-investors [Accessed: 2022].

Ritter, J. R. (1991). The long-run performance of initial public offerings. *The Journal of Finance.*

Ritter, J. R. (2021). Initial public offerings: Median age of ipos through 2021. *Yale Journal.*

Ritter, J. R. (2022). Initial public offerings: Updated statistics. https://site.warrington.ufl.edu/ritter/files/IPO-Statistics.pdf[Accessed: 2022].

Ritter, J. R., Gahng, M., and Zhang, D. (2021). Spacs.

Russell, F. (2021). Russell us indexes. https://www.ftserussell.com/products/indices/russell-us [Accessed: 2022].

Røttingsnes, K. and Gjærum, J. L. (2019). Digging for fool's gold. *NHH Brage.*

Securities, U. and Commission, E. (2021). What you need to know about spacs – updated investor bulletin. Technical report, U.S. Securities and Exchange Commission.

SimplyWallSt. (2022). Institutions own 27% of gamestop corp. (nyse:gme) shares but individual investors control 56% of the company. https://simplywall.st/stocks/us/retail/nyse-gme/gamestop/news/institutions-own-27-of-gamestop-corp-nysegme-shares-but-indi [Accessed: 2022].

Tat, M. J. (2017). *Towards Data Science*. https://towardsdatascience.com/seeing-the-random-forest-from-the-decision-trees-an-intuitive-explanation-of-random-forest-beaa2d6a0d80 [Accessed: 2022].

Telford, T. and Siegel, R. (2021). 2021 goes down as a year of high risk and high reward for markets. https://www.washingtonpost.com/business/2021/12/31/stocks-markets-2021-omicron-pandemic/ [Accessed: 2022].

Theil, H. (1971). *Principles of Econometrics*. John Wiley Sons, Inc.

Westenberg, D. A., Hilts, C., and Murray, S. (2022). Ipo report 2022.

Wooldrigde, J. M. (2012). *Introductory Econometrics - A Modern Approach*. South-Western Cengage Learning.

Appendix

A1 Pre-merger Net Cash per Share Equation

While only some of the elements of the Net Cash per Share equation are used as predictors in our analysis, the remaining elements are of importance to understand the bigger picture. Thus, the following section will describe the equation in more detail. In addition, we will shortly comment on how we collected the necessary data.

The complete equation, as of Klausner et al. (2020) is displayed under. It is important to emphasize the fact that the equation below is excluded redemption rates – presented in a contingent form:

$$\mathrm{NC} = \left[\frac{\begin{array}{r}\text{Total Cash(From SPAC Public shareholders + PIPE)} - \text{Cash expenses} \\ + \text{ Value of Warrants} + \text{Value of other Equity Dervatives}\end{array}}{\text{Public shares} + \text{Founder shares} + \text{PIPE shares} + \text{Other} + \text{Shares w/rights}} \right] \tag{.1}$$

The already described elements of interest can be found in section 3.1. Below follows a short explanation of the other elements, including how the data was collected.

A1.1 The numerator: Cash

A1.1.1 Cash Expenses

The Cash Expenses is mainly consisting of two sub-elements. First, the *Deferred Underwriting Fee*, which is the fees the underwriter demands on completion of the merger. This is paid out of cash in the trust at the time of the merger. Second, the Additional Merger-Related Fees is an expense for the SPAC. This relates to financial advisory and other fees in connection with the merger (Klausner et al., 2020).

The Cash Expenses are calculated using different proxy-numbers found. This is due to the fact that the disclosure of such numbers are missing in filings. Hence, we have calculated the total cash expenses of being 15% of total assets, whereas approximately 5.5% is the initial deferred underwriting fee. The remaining additional fees equals approximately 10%. This is in accordance with (Klausner et al., 2020) that finds that the median Underwriting fee is 16%.

A1.1.2 Value of Warrants

SPAC typically issue warrants to sponsors in exchange for the sponsor's initial investment. Moreover, SPAC also often issue similar warrants to underwriters and PIPE investors. The value of these warrants must be subtracted from the equation in order for it to be precise. The valuation of warrants should be conducted looking at the market value of them as of the day before the merger (Klausner et al., 2020).

The market value of warrants is somewhat difficult to extract, and they are not available via SEC. However, after calculating the average convention ratio of 0.38 (shares per warrant) for our Merger Cohort, the total shares outstanding and the intrinsic value[10] \$1.5, we are able to find the market value. For example, if company A has 1.000.000 shares outstanding, the market value of warrants for company A equals: $1.000.000 * 0.38 * \$1.5 = \$570.000$. While this is not an exact value for all companies respectively, we are confident that the mean works as a good proxy of the true value. Also, after performing a sensitivity analysis with different convention ratios (ranging from 0.2-0.8), the results do not significantly differ.

Moreover, on warrants, it is important to comment on the fact that SEC recently changed the guidance on warrants, no longer treating them as a liability (Coates and Hunter, 2021). Thus, unlike many other papers, the Net Cash Per Share in this thesis will be somewhat higher than what other finds. This is due to the fact that we now subtract this part of the numerator, instead of addition it.

A1.1.3 Value of other Equity Derivates

These are elements such as convertible debt or convertible preferred stock in private placement concurrently with a merger. Thus, the embedded option that allows the holder to convert the debt (or preferred stock) into shares should be extracted (Klausner et al., 2020).

From what we can find, convertibles are not present in our Merger Cohort. Thus, we set this equal to zero. *(This, however, does not affect any metrics in our analysis. As a matter of fact, no elements described in this appendix does. The purpose of this appendix*

[10]The intrinsic value of the warrant is equal to the uniformly set warrant price of \$11.50 subtracted the \$10 share price at merger

is to give the reader a better understanding and intuition of the importance behind Net Cash per Share equation)

A1.2 The Denominator: Total Shares

A1.2.1 Public Shares

The public shares is "traditional" shares issued to the public in order to raise capital. These are uniformly referred to as "Class A common Stock" in SEC filings. Hence, we can pull this from 10-Q/K filings prior to the merger using R.

A1.2.2 Founder shares

The founder shares is the shares issued to sponsors as "promotes". This is near-uniform in totalling 20% of the post-IPO equity. In addition, some SPACs provide "Sponsor Earnouts" which is an agreement that under which some of their shares (typically 30-40%) will be cancelled if the share price of the company does not meet specific thresholds (typically $12.50, $15.00 and $17.50) (Klausner et al., 2020).

The founder shares was collected through the latest SEC filing 10-Q/K prior to the merger. These shares are typically referred to as "Class B shares". Thus, using R., we scrape the SEC website and pull this data.

A1.2.3 PIPE shares

This is shares issued to PIPE investors in compensation for the equity infusion concurrent with the merger. Some PIPE investors demand a discount, while some do not. All the shares issued to PIPE investors must be accounted for in the denominator in order to obtain accuracy for the net cash per share

The shares to PIPE investors was collected through the SEC 8-K filing after the merger (item 2.01: "Completion of Acquisition or Disposition of Assets"). PIPE shares are extracted manually from the respective filing, often under "PIPE Shares".

A1.2.4 Other shares and shares issuable under Rights

Other shares is any other shares issued up to the point of the merger. Shares issuable under rights is typically a right to acquire 1/10 of a share and occasionally a right to acquire 1/20 of a share. These are exercisable after a merger and require no payment from the holder. The total number of rights multiplied with the fraction of a share should be added to share count in the denominator (Klausner et al., 2020).

Since this part of the denominator only applies to some SPACs, we do not account for this in our calculations of the Net cash per share.

A2 OLS Assumptions

A2.1 Residual Plots

To reveal aspects of our different models which are not discovered in the summary statistics, we run a simple diagnostic in R. by plotting our controlled models to obtain different residual plots. Hopefully, this will give a better picture on how well the models fits our data (Bartell, 2019). Moreover, this will help checking whether the OLS Assumptions are violated or not.

The first plot (Residuals vs Fitted) are making sure that the variable coefficients (beta) are linear in relation to their corresponding independent variables *(Assumption 1: linearity in the parameters)*. A horizontal line, without distinct patterns, indicates a linear relationship (Bartell, 2019).

The second plot (Normal Q-Q) ensures that the sample used in the regression comes from a truly random sample of the study population, with a greater number of observations than parameters, and fixed independent variables which do not impact the dependent variable (Bartell, 2019). Moreover, this plot inspect that there is no relationship between the independent variable and the error terms, which should have a mean of zero *(Assumption 2 & 4: Random sampling and Zero conditional means)*. In an ideal model, the residuals points follow the straight dashed line.

The third plot (Scale-Location) validates that the variances of the error terms do not depend on the independent variables, and are ideally all equal and not correlated with each other *(Assumption 5: Homoscedasticity)* (Bartell, 2019). The models does not violate such assumption if the red line is straight and horizontal.

Finally, the fourth plot (Residuals vs Leverage) is used to determine the influence of extreme values produced by the model (Bartell, 2019). While these plots does not account for any violations in the Gauss-Markov Theorem (Theil, 1971), it is, however, helpful in identifying points that are influential to the regression results. These points are identified outside the "Cook's distance" (past the red dotter hyperbolic lines) in the plot.

(As of assumption 3: No perfect collinearity, this is elaborated on further down in this section under A.2.2: Multicollinearity)

A2.2 Controlled model of two months return

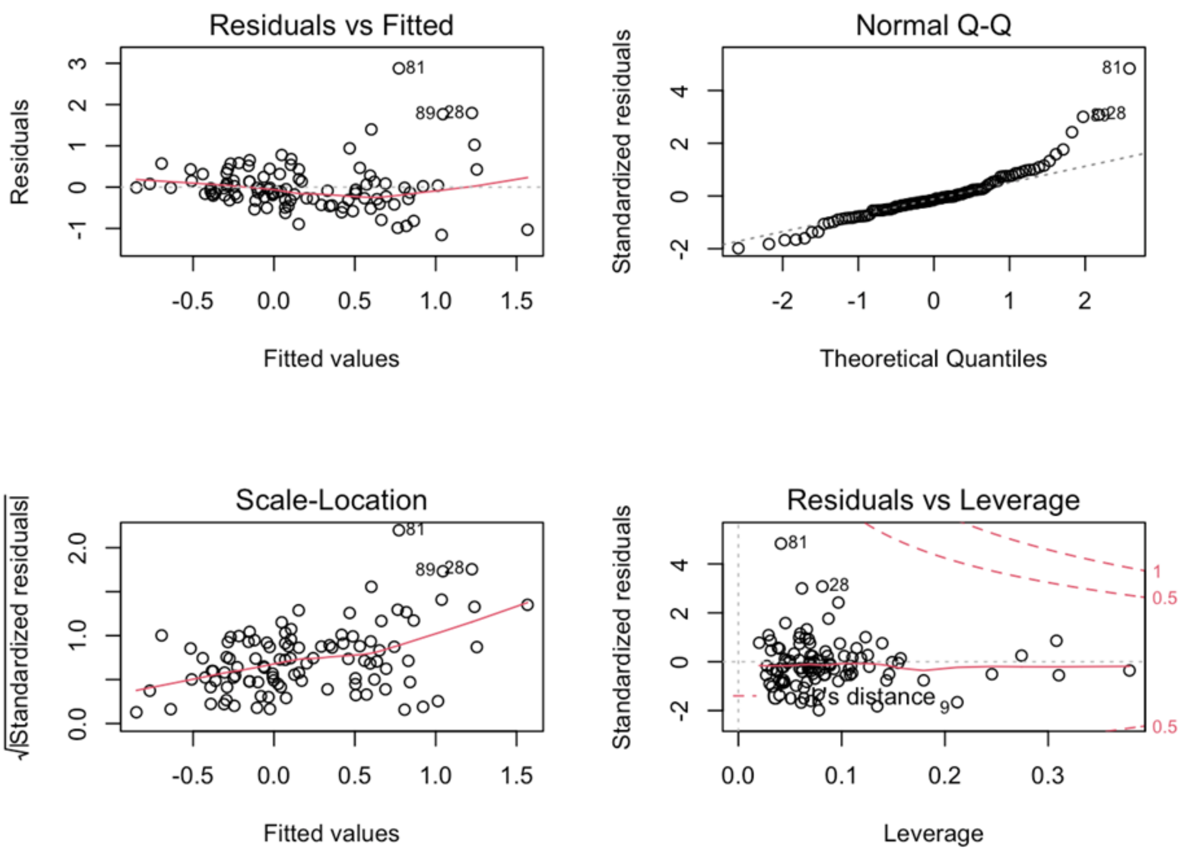

Figure A2.1: OLS Assumptions plots for two months regression model

The result from the plots when plotting the controlled model (7) in 5.1 indicates that all the OLS assumptions holds. Some of our observations (81, 89, 28) can be viewed as outliers. However, the overall fit of the model is good and the aforementioned outliers are in line with the thresholds of Bartell (2019). Thus, we decide not to remove these in order to keep our original number of observations in our dataset.

A2.3 Controlled model of first-day return

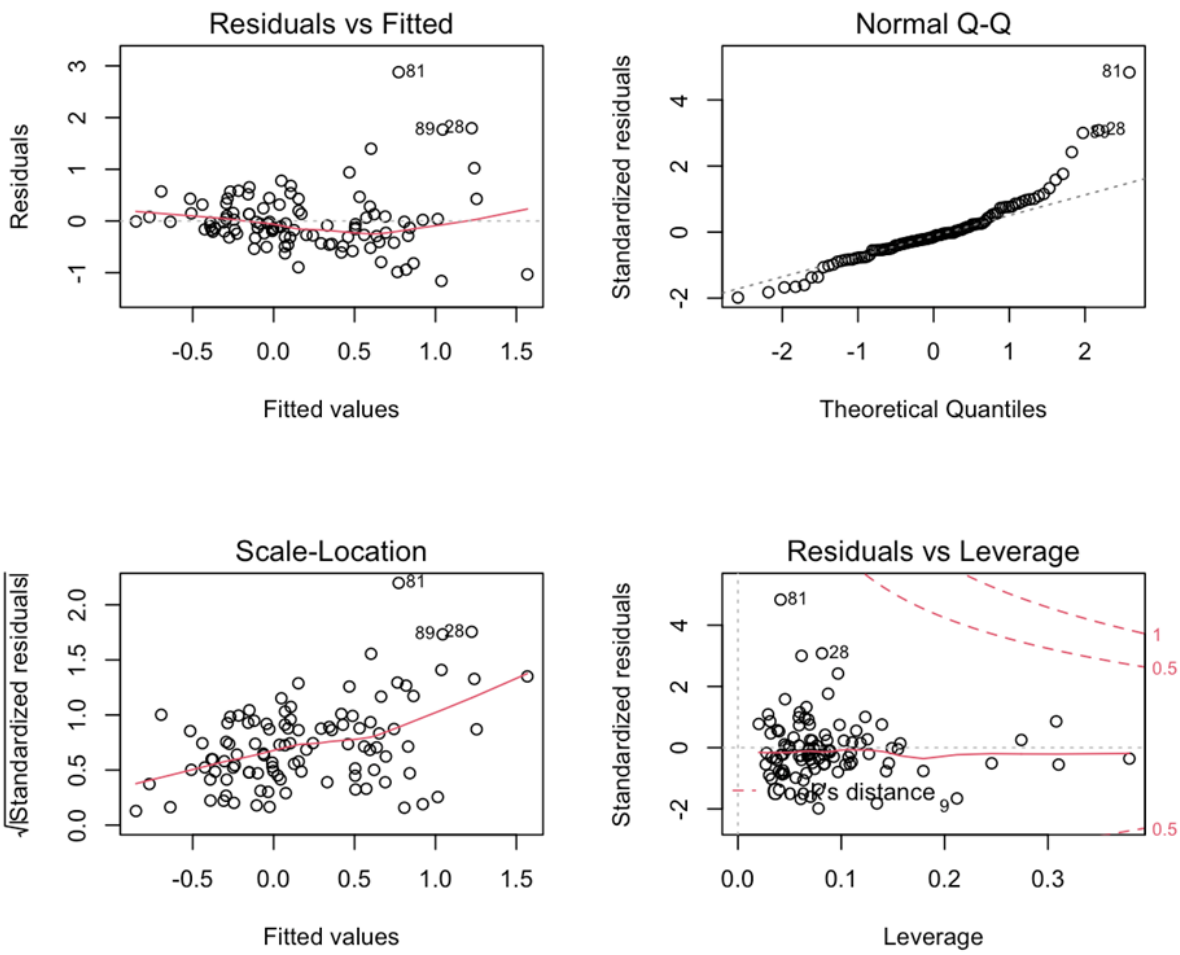

Figure A2.2: OLS Assumptions plots for first-day regression model

The result from these plots in the controlled model (7) from regression table 5.2 yield the same results as the plotted model above. This is not surprising given the fact that these models are similar. The only difference between them (the predictors) are that this model does not have market returns as predictors. None of the assumptions is violated, and the overall fit of the model is good. The same outliers are present, but none interfere with the thresholds. Thus, we decide not to remove these in order to keep our original number of observations in our dataset.

A2.4 Hype indicators on first-day return

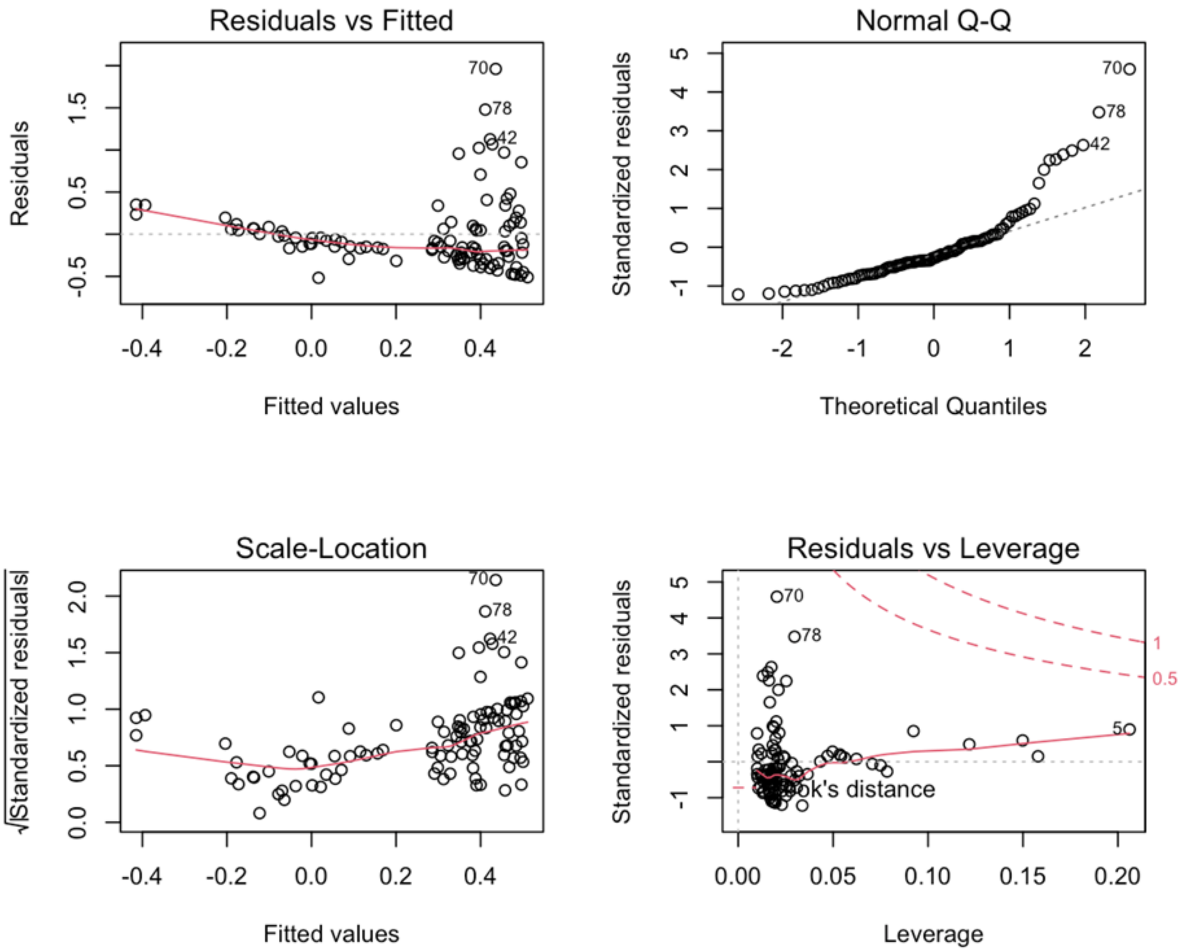

Figure A2.3: OLS Assumptions plots for Hype regression model

In the model on hype indicators **??**, the above plots again indicate that the OLS assumptions holds. The model fits the data well, as of (Bartell, 2019). As of with the previous plotted models, we do identify possible outliers in this model (70, 78, 42). However, the same intuition on keeping the observations applies here, and we decide not to "winzorize" or remove them.

A2.5 Multicollinearity

OLS assumptions 3 (no perfect collinearity) states that if an independent variable is an exact linear combination of the other independent variables, then the model suffers from perfect collinearity, and it cannot be estimated by OLS (Wooldrigde, 2012). However, some level of multicollinearity in a model is warranted and preferred as it produces better OLS estimates. Reduction in levels of correlation can be achieved by simply removing one of the correlated variables. In doing so, the data is subjected to omitted variable bias. Hence, there is a tradeoff between accuracy and biases when conducting such a study, but in general, one would rather avoid omitted variable bias in favor of accuracy.

Measurement of multicollinearity can be done in multiple ways. We have provided both the calculated variance inflation indicator (VIF) and the correlation matrix in the graphs (1) and table (2), respectively. The VIF estimates how much of the variance of a regression coefficient is inflated due to multicollinearity in the model. VIF is a continuous number, with a higher value linked to greater levels of correlation. A commonly used threshold in scientific research papers is > 10, which is considered a clear indication of multicollinearity (Choueiry, 2021). However, a recent research on VIF and collinearity, Johnston R. (2018) in his paper argues that a VIF of >2.5 already indicates considerable collinearity. Our values observed in the graph below show no value above the aforementioned threshold, and ultimately indicate no signs of high correlation between our independent variables.

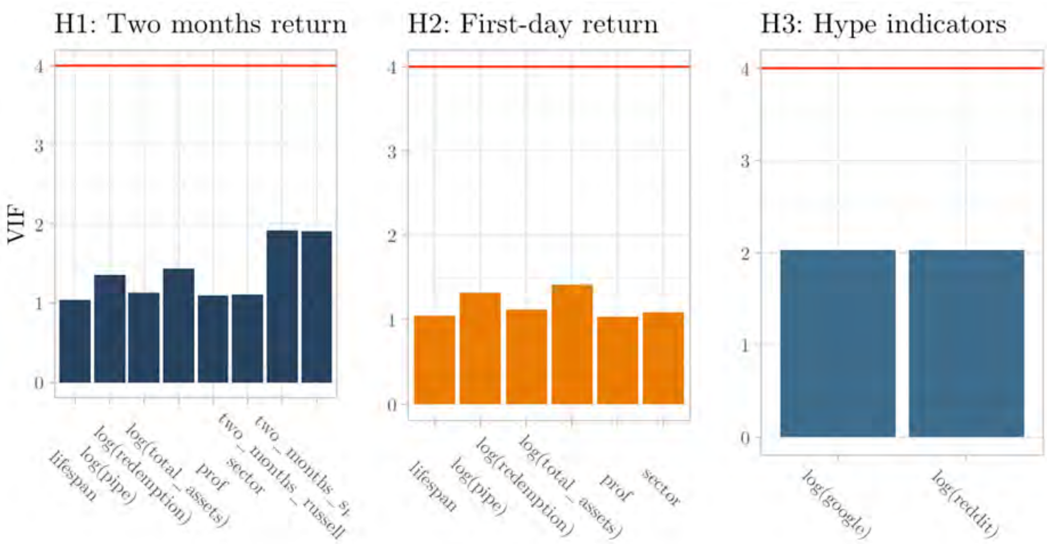

Figure A2.4: Variance Inflation Indicator

In order to further investigate the assumption of no perfect collinearity, we have computed a correlation matrix below in table A2.1. The result from the matrix confirms our conclusions from the VIF. According to Røttingsnes and Gjærum (2019) "correlations higher than 0.8 yields severe problems with multicollinearity commonly appear." None of the independent variables below exceeds this threshold. The highest correlated predictors are the hype indicators, Google search hits and Reddit mentions, with a correlation of 0.7049. One could argue that this is high, as "less correlation is better". To account for the high correlation, one could either increase the observations or drop predictors. Due to the recent data we are looking at, we cannot include more observations as these have traded for less than two months during our data collection. Moreover, as of the latter, this could possibly lead to a higher bias (Wooldrigde, 2012). Thus, since no multicollinearity between our predictors in the 0.8 threshold is present, we keep both predictors.

Table A2.1: Correlation Matrix

	log_redemption	log_pipe	log_total_assets	lifespan	sector	log_google	log_reddit
log_redemption	1						
log_pipe	-0.2362	1					
log_total_assets	-0.2162	0.4634	1				
lifespan	-0.1249	0.0369	0.1436	1			
sector	0.0673	-0.0685	-0.2764	-0.0947	1		
log_google	-0.4534	0.1768	0.1487	0.0173	0.0178	1	
log_reddit	-0.4056	0.0337	-0.0451	-0.0721	0.0145	0.7120	1

A2.6 Disclosure

All above diagnostics and test indicates that our data does not violate the OLS assumptions as of (Wooldrigde, 2012). Consequently, the results are evidential for the Gauss-Markow theorem being valid. Thus, our OLS estimators are in line with the conditions for being the Best Unbiased Estimators (BLUEs) for the model.

A3 Alternative regression models based on results

A3.1 Two-month Return excluded Redemption Rate

Table A3.1: OLS regression with the two-months return excluded Redemption Rate

	Dependent variable:					
	two_months_return_unit					
	(1)	(2)	(3)	(4)	(5)	(6)
log(pipe)	0.151***					0.119**
	(0.056)					(0.059)
log(lifespan)		0.094				
		(0.075)				
sector			−0.352**			−0.213
			(0.171)			(0.164)
two_months_sp500						−0.688
						(2.794)
two_months_russell						4.266***
						(1.308)
log(total_assets)				0.190**		0.101
				(0.084)		(0.092)
lifespan						−0.001
						(0.003)
prof1					0.105	0.016
					(0.191)	(0.177)
Constant	−2.664**	−0.013	0.270***	−3.493**	0.161*	−4.059**
	(1.052)	(0.172)	(0.086)	(1.622)	(0.084)	(1.600)
Observations	103	103	103	103	103	103
R^2	0.068	0.015	0.040	0.048	0.003	0.249
Adjusted R^2	0.059	0.006	0.031	0.039	−0.007	0.194
Residual Std. Error	0.743	0.763	0.753	0.750	0.768	0.687
F Statistic	7.356***	1.570	4.249**	5.141**	0.304	4.499***
Note:						*p<0.1; **p<0.05; ***p<0.01

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with first day return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant and non-zero explanatory effect. The measurement of fit $R^2_{Adjusted}$ reports that our model explain only 4.7% of the variance in the first-day return. This implies that removing the redemption rate significantly reduces the explanatory power of the model. From an alternative perspective, interpreting this as the variables are of less importance for the initial return, this aligns better with previous research on initial IPO performance.

A3.2 First-day Return excluded Redemption Rate

Table A3.2: OLS regression with first-day return excluded redemption rate

	Dependent variable:					
	first_day_return_unit					
	(1)	(2)	(3)	(4)	(5)	(6)
log(pipe)	0.087**					0.098**
	(0.036)					(0.040)
log(lifespan)		0.017				
		(0.048)				
sector			−0.160			−0.171
			(0.110)			(0.113)
log(total_assets)				0.039		−0.045
				(0.055)		(0.063)
lifespan						−0.002
						(0.002)
prof1					0.155	0.159
					(0.121)	(0.119)
Constant	−1.383**	0.228**	0.303***	−0.492	0.232***	−0.696
	(0.677)	(0.111)	(0.055)	(1.060)	(0.054)	(1.104)
Observations	103	103	103	103	103	103
R^2	0.056	0.001	0.020	0.005	0.016	0.098
Adjusted R^2	0.046	−0.009	0.011	−0.005	0.006	0.052
Residual Std. Error	0.478	0.491	0.487	0.490	0.488	0.476
F Statistic	5.943**	0.122	2.090 (df = 1; 101)	0.508	1.636	2.114*
Note:				*p<0.1; **p<0.05; ***p<0.01		

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with first day return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant and non-zero explanatory effect. The measurement of fit $R^2_{Adjusted}$ reports that our model explain only 4.7% of the variance in the first-day return. This implies that removing the redemption rate significantly reduces the explanatory power of the model. From an alternative perspective, interpreting this as the variables are of less importance for the initial return, this aligns better with previous research on initial IPO performance.

A3.3 First-day Return Regression Model with Hype Indicators

Table A3.3: OLS regression with original independent varibles and hype indicators

	Dependent variable:							
	first_day_return_unit							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(redemption)	−0.078***							−0.068***
	(0.010)							(0.011)
log(lifespan)		0.017						−0.034
		(0.048)						(0.037)
sector			−0.159					−0.163*
			(0.110)					(0.086)
prof1			0.155					0.244**
			(0.121)					(0.094)
log(total_assets)				0.039				−0.084*
				(0.055)				(0.050)
log(pipe)					0.087**			0.056*
					(0.036)			(0.031)
log(google)						0.400***		0.137
						(0.077)		(0.097)
log(reddit)							0.275***	0.051
							(0.056)	(0.069)
Constant	−0.017	0.228**	0.273***	−0.492	−1.383**	−1.346***	−2.533***	−0.423
	(0.053)	(0.111)	(0.060)	(1.060)	(0.677)	(0.311)	(0.575)	(1.100)
Observations	103	103	103	103	103	103	103	103
R^2	0.365	0.001	0.036	0.005	0.056	0.213	0.191	0.495
Adjusted R^2	0.359	−0.009	0.017	−0.005	0.046	0.205	0.183	0.453
Residual Std. Error	0.392)	0.491	0.485	0.490	0.478	0.436	0.442	0.362
F Statistic	58.149***	0.122	1.873	0.508	5.943**	27.271***	23.797***	11.540***
Note:						*p<0.1; **p<0.05; ***p<0.01		

The regression table reports the coefficients, standard errors (in parentheses) and significance level (*) from the regressions run with first day return as a dependent variable. F-statistics show that the joint effects of our variables are statistically significant and non-zero explanatory effect. The measurement of fit $R^2_{Adjusted}$ reports that our model explain 46.3% of the variance in the first-day return. The obtained

$R^2_{Adjusted}$ increases when adding the "hype" predictors. Put differently, the explanatory power of the model increases when log(google) and log(reddit) are included.

A4 Complementary Models

A4.1 Gini Impurity Index

The Gini Impurity index measures the difference in the residual sum of squares (RSS) before and after the split of the variable. In general, the below plot illustrates the predicting power of our Random Forest Model. The top variables (total assets and lifespan) are of the most importance. Thus, removing one of these will greatly reduce the prediction power of redemption rates (Breiman and Cutler, 2022).

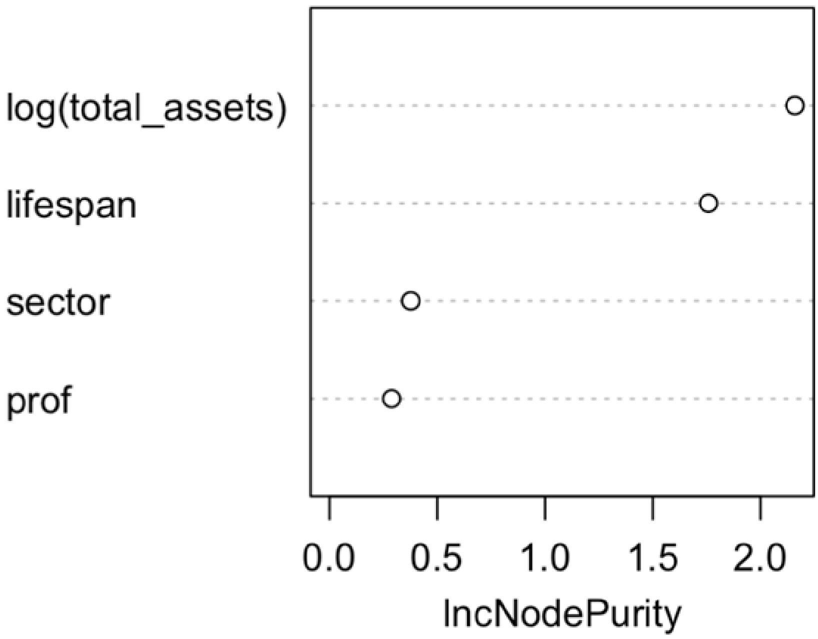

Figure A4.1: Gini Impurity Index

A4.2 Ordinal Logistic Regression

```
Call:
MASS::polr(formula = fct_redemption ~ total_assets + lifespan +
    sector + prof, data = df_logit)

Coefficients:
                Value Std. Error t value
total_assets -0.54293   0.242334 -2.2404
lifespan     -0.02023   0.009938 -2.0358
sector1       0.36898   0.453487  0.8137
prof1        -0.43435   0.516439 -0.8410

Intercepts:
    Value    Std. Error t value
1|0 -12.6428   4.7520    -2.6605
0|2  -9.8667   4.6881    -2.1046

Residual Deviance: 188.8763
AIC: 200.8763
```

Figure A4.2: Ordinal Logistic Regression Table

A4.2.1 Brant Test

In short, Brant Test assesses whether the observed deviations from our Ordinal Logistic Regression model are larger than what could be attributed to chance alone. A p-value of less than 0.05 on this test, particularly on the Omnibus plus at least one of the variables should be interpreted as a failure of the proportional odds assumption (Fink, 2022).

Table A4.1: Brant Test

	X2	df	probability
Omnibus	8.267	4	0.082
total_assets	2.023	1	0.155
lifespan	2.110	1	0.146
sector1	2.075	1	0.150
prof1	2.521	1	0.112