UC Santa Barbara
Econ 196 Honors Thesis

Title
Does Rhetoric Help Stock Returns? A Tex tAnalysis of Unicorn S-1 Documents

Permalink
https://escholarship.org/uc/item/389560ds

Author
Curl, Ben

Publication Date
2020-03-16

Undergraduate

eScholarship.org Powered by the California Digital Library
University of California

Does Rhetoric Help Stock Returns? A Text Analysis of Unicorn S-1 Documents

Ben Curl

Faculty Advisor: Erik Eyster

Course Advisor: Shelly Lundberg

University of California, Santa Barbara

March 16, 2020

Abstract

This paper seeks to develop and analyze a relationship between venture capital investment, rhetorical corporate strategies, and public stock performance. Examining every firm since 2000 that went public at a market cap of 1 billion USD or above, I analyze the rhetoric of unicorns and its relation to risk. Using SEC archives of S-1 financial documents and two sentiment dictionaries, I attempt to capture levels of positive language in firms' business summaries and negative language in its risk factors. Using this data, I test the correlation between a firm's venture capital investment and its S-1 language, as well as the relationship between a firm's S-1 rhetoric and its ensuing stock performance as a public company. A significant positive correlation is established between venture capital investment and a firm's levels of positive language in their business summaries, as well as a significant positive correlation between venture capital investment and a firm's levels of negative language in their risk factors. Impacts of business summary language on daily, weekly, and monthly returns after a firm's IPO are negligible.

1. Introduction

This paper seeks to develop and analyze a connection between venture capital investment, rhetorical corporate strategies, and public stock performance. I examine every firm since 2000 that went public on the New York Stock Exchange or NASDAQ at a market cap of 1 billion USD or above. These firms are commonly referred to as “unicorns”, defined as privately-held companies with a valuation at or above 1 billion USD.

Using SEC archives of S-1 financial documents, I collect and analyze the business summary and risk factor section of each firm’s S-1. The business summary section provides an overview of the firm and includes details such as its mission statement, clients, supply chain, future strategy, and revenues. I use this section of the S-1 as an analog to a unicorn firm’s “pitch”, the presentation startup companies make to prospective venture capital firms. The risk factor section of the S-1 is intended to disclose significant factors that would make a firm’s IPO especially risky or uncertain. I use this section of the S-1 as a proxy for a firm’s perceived risk.

Text analysis is conducted using the Linguistic Inquiry and Word Count (LIWC) module, a proprietary text analysis software. I use LIWC’s own textual sentiment word bank, which contains over fifty different text analysis categories, as well as a word bank compiled by Notre Dame researchers Tim Loughran & Bill McDonald. Loughran & McDonald applied linguistic analysis to financial documents to develop word lists of negative financial words, as well as banks for positive, uncertain, litigious, strong, and weak language. Text analysis in a purely financial context will help capture an accurate assessment of a firm’s risk, as measured by a text analysis of its risk factors.

Using this data, I test the correlation between a firm’s venture capital investment and its S-1 language. Given the nature of the venture capital investment process and the importance it places on firms’ pitches, I hypothesize that firms with venture capital investment will use more positive and rhetorical language in their business summaries relative to firms without VC investment. I also hypothesize that firms with venture capital investment will be riskier than their non-VC counterparts, as quantified by negative financial text sentiment in each observed firm's S-1 risk factor section.

To examine the relationship between S-1 rhetorical strategies and security performance, I collect observed firms' daily, weekly, and monthly stock returns following the IPO. I also collect the market return for each firm over the same observed period to determine each stock's abnormal return. I hypothesize that firms using higher levels of rhetorical and emotional language will be correlated with lower relative returns.

2. Historical Background

The preponderance of "unicorns", privately-held companies with a valuation at or above 1 billion USD, has grown exponentially in the past ten years. The term unicorn is a 21st century invention. Venture capitalist Aileen Lee dubbed the word in 2013, when only 39 firms qualified for unicorn status (Lee, 2013). Today, 575 companies fit the label of a unicorn, and an additional 179 former unicorns have successfully been acquired or existed into an IPO (Crunchbase, accessed 2020). Unicorns are now so common that, ironically, the metaphor is no longer apt.

Such a sharp uptake in the number of unicorns is correlated with the growth of the venture capital firms which fund them. Venture capital (abbreviated VC) is a form of private equity financing for early-stage startup companies. The focus of VC investment has shifted in recent years. Once concentrated in seed and early-stage investment for early-stage firms, the majority of VC funding has now flowed to deals with late-stage, established startups. The implications of such a change are significant. So-called "super-giant" deals, equity funding rounds totaling over 100 million USD, accounted for 56 percent of VC dollar volume in 2018. As evidenced by Figure 1, these types of deals were a rarity as recently as 2013. Super-giant deals have become so common that analysts have dubbed a new term, “hyper-giant” rounds, to describe equity funding rounds totaling over 250 million USD. Figure 2 captures a sharp uptick in the volume of VC funding for late-stage firms over the past ten years.

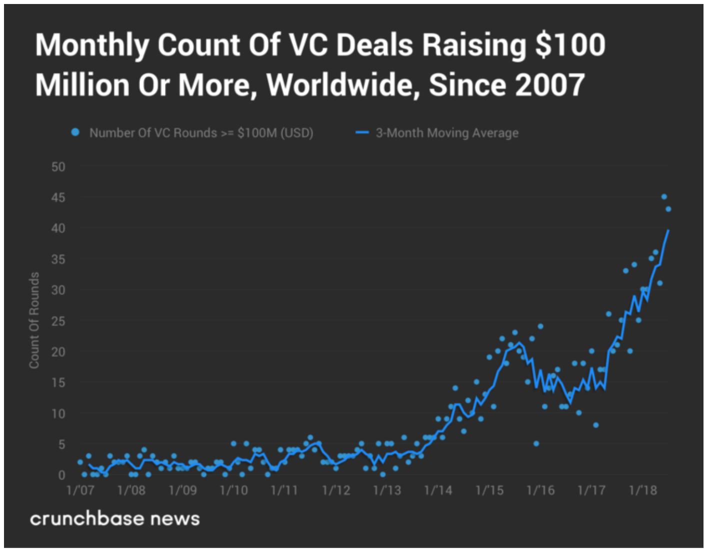

Figure 1: Credit: Crunchbase News, 2018

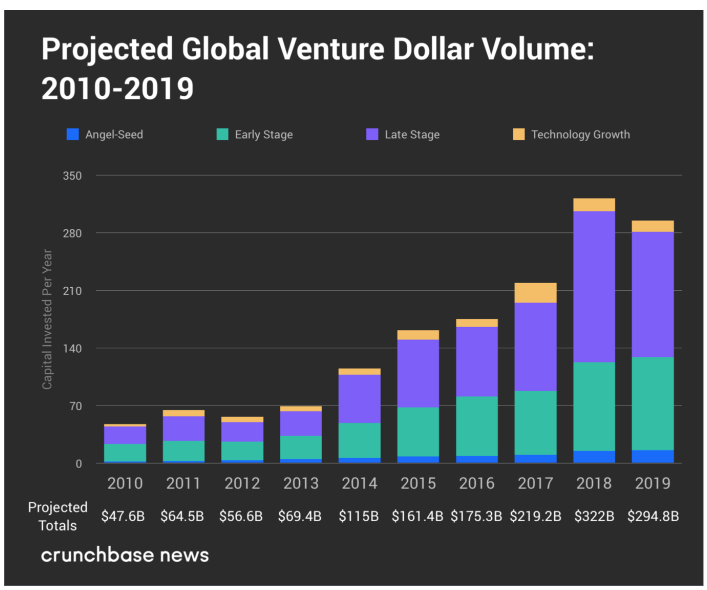

Figure 2: Credit: Crunchbase News, 2020

The implications of these drastic changes have consequences for other financial markets. Delaying exits allow companies to avoid releasing financial data required by the SEC, and prevents retail investors with no access to private markets from realizing any future gains. Conventional wisdom dictates that the clearest way for private companies to raise money is to launch an initial public offering (IPO). Google, for example, raised just $31.6 million in venture funding before it went public at a market cap of $21.3 billion in August 2004 (Rowley 2018). Nowadays, firms are increasingly able to raise gargantuan sums of capital through private investors, bypassing traditional growth trajectories for unicorn startups. Google went public six years after its founding in September 2008; it took Facebook eight years before its IPO in 2012. Nowadays established, large-cap unicorns have little need to go public as they raise billions of dollars on the private market. It took Uber ten years and approximately 24.2 billion USD in funding before going public, and even well-known large-cap companies such as Airbnb are still private 12 years after its founding.

Drastic changes in venture capital markets may have reached a tipping point in the past 18 months, as some of the world's largest unicorns filed documents to go public. Uber, the most valuable unicorn in U.S. history before its IPO, raised $24.2 billion in funding as a private company before going public in 2019. The staggering amounts of capital on hand has allowed Uber to continue its core strategy of "disruption", undercutting fares of established taxi services and rideshare competitors. This strategy has led to rapid yearly growth, but in the process, Uber has lost money on every transaction they make. The company continues to bleed cash, has promised investors it will reach profitability in 2021 as it continues to widen its market share (Conger 2019). This appeal was successful—to VCs, at least. When Uber released its S-1 in April of last year, analysts outside of the VC bubble were able to peek under the ridesharer's hood for the first time. In an attempt to spin their weak profitability metrics, Uber made grandiose proclamations about their purpose. The company's stated mission was to "ignite opportunity by setting the world in motion", it claimed that it was "fueling the future of independent work" through their controversial labor tactics, and claimed to employ organizational synergy between platforms without actually outlining how that synergy occurs (Uber Technologies 2019). When addressing the negative margins it earns per ride, Uber employed a form of doublespeak: "We can choose to use incentives, such as promotions for Drivers and consumers, to attract platform users on both sides of our network, which can result in a negative margin until we reach sufficient scale to reduce incentives". Uber's language may have convinced VC firms to invest billions in the private market, but retail investors were not as easily convinced. Uber's stock dropped 10.75% on its first day of trading, and the stock has underperformed the market by nearly 20% since its IPO at the time of this writing.

An even larger IPO calamity occurred several months later with WeWork, a real estate firm that provides shared workspaces for startups and entrepreneurs. Like Uber, WeWork was able to achieve rapid revenue growth by spending billions of dollars of VC-funded money. As of this writing, WeWork has raised 47 billion USD in venture capital deals. The majority of capital raised comes from one source: Japanese VC firm SoftBank. WeWork was the cornerstone investment in SoftBank's Vision Fund, a growth stage venture fund with $100 billion of cash on hand. By the end of 2018, WeWork was a behemoth; the company had become the largest occupier of office space in both London and Manhattan and operated 400 locations in 99 cities around the world (Landy 2018). Such staggering growth led the company to plan an IPO for Fall 2019. That IPO ultimately never came to fruition—WeWork released its offering after investors had been burned by the IPOs of cash-burning unicorns like Uber. The nonsensical language employed in their S-1 didn't help either. WeWork's mission statement? "Elevate the world's consciousness" (The We Company 2019). Attempting to latch onto the clout of high-performing software-as-a-service (SaaS) companies, WeWork labeled itself as the pioneer "space-as-a-service" company. In addition to WeWork's exaggerated rhetoric, the firm botched several key financial details in the report, further dooming its hopes of going public at their desired price (Eaglesham & Brown 2019). The backlash ultimately forced WeWork founder Adam Neumann to resign, and WeWork retracted all plans to go public. Ultimately, WeWork was an example of the market correcting VC overenthusiasm; most analysts and investors saw the preposterous language WeWork used as overly emotional and a cover for their poor financials.

Market corrections do not always play out like the cases of Uber and WeWork. Consider the drastic fall of Theranos, the most catastrophic example of overeager VC funding. Theranos was a healthcare company that claimed to have the technology to conduct over 200 blood tests using just a finger prick. Such a claim threatened to revolutionize the field of blood testing. The founder of Theranos, Elizabeth Holmes, quickly became a mini-celebrity in her own right. The CEO became noteworthy for her unnaturally baritone voice and frequent use of black turtlenecks, ala Apple co-founder Steve Jobs. At the annual TEDMED healthcare conference in 2014, Holmes laid out her grand vision using emotional pleas. The CEO noted her grandfather's sudden death from a form of skin cancer that quickly led to brain cancer. Holmes's grandfather died before she had a chance to say goodbye. Building off of her personal tragedy, Holmes proposed an ambitious vision for the future of healthcare:

> *"If I had one wish, standing here with all of you, it would be that today, just for a minute, you think about the fact that we have this right, a human right, to engage with information about ourselves, about our bodies, and for those that we love to engage with information about themselves. And when we do that, we will change our lives, and the lives of those we love will change. And we'll begin to change our healthcare system and our world."*

Holmes's emotional appeals had almost no relation to her actual business. Its purpose was to deeply resonate with potential investors, and it worked. Theranos quickly became one of the most valuable unicorns in tech, garnering investments from high-profile figures in all sectors of the economy. Walmart's Walton family invested roughly 150 million USD in the company, current Education Secretary Betsy DeVos invested $100 million, and business mogul Rupert Murdoch invested around $125 million (Carreyrou 2018). By 2015, Theranos's Board of Directors was an all-star team of political personalities; the board included former Secretaries of State Henry Kissinger and George Shultz as well as former defense secretaries and senators. These politicians had one thing in common: they had no experience in the field of healthcare. Leveraging political capital and celebrity clout, Theranos was able to attract people with large pocketbooks to invest, typically though emotional pleas like the pitch Holmes gave in 2014. Ultimately, Theranos's bold vision came crashing down. A series of investigations from the Wall Street Journal alleged Theranos defrauded investors by lying about the capabilities of its blood-testing technology. Eventual investigations by the Security and Exchange Commission charged Holmes and Theranos with "massive fraud". Theranos had completely made up their success, lying about its advanced technology and overinflating its revenue by 1000 times its true value to investors (Aiello 2018). The company's assets were liquidated in 2018; Holmes was given a 500,000 USD fine and was barred from serving as an officer or director of a public company for ten years (Thomas & Abelson 2018).

The cases of Uber, WeWork, and Theranos are emblematic of a growing trend among unicorn firms. All three companies used elaborate rhetoric and emotional signals to craft a narrative around its brand. These pitches, no matter how divorced from reality, proved to attract billions of VC dollars and turned these firms into household names. Bombastic language only goes so far, however. Once exposed to public scrutiny, each firm's weak financials caused its astronomical valuations to dwindle. In September 2019, New York University marketing professor Scott Galloway dubbed the term "yogababble" to describe the phenomena of firms using spiritual language to make their brand more attractive to investors and consumers. "Overpromise and underdeliver has become a means for access to cheap capital," Galloway wrote in his blog. "The lines between charm, vision, bullsh*t, and fraud have become so narrow as to be one line" (Galloway 2019). But how far does "yogababble" reach? Is it an occurrence that only applies to just a few VC-backed darlings, or does the phenomena apply to all unicorns? In this paper, I will attempt to quantify the rhetoric used by these unicorns, and establish relationships between this rhetoric and future stock returns.

3. Literature Review

The history of venture capital dates back to the dawn of the internet era. Hellman and Puri (2000) note that in the late 20th century, investors played an active role in the governance of the ventures they invested in. Serving beyond the role of typical financial intermediaries, venture capital firms helped startups build their internal organization, specifically their employee base. Some VC firms went as far as helping startups recruit an outsider to assume the role of CEO. The obligations of a VC firm to its investments kept deal activity low in the 1990s. Venture capital firms invested in just a few ventures because of the ample time and resources it took to assume a central role in the invested firms.

Venture capital has undergone a paradigm shift in the last two decades. Supply shocks in the technology sector have lowered the costs of starting new businesses, introducing new investment opportunities that were previously not viable. Ewens, Nanda, & Rhodes-Kropf (2018) remark that venture capital firms began to adopt a "spray and pray" strategy around the mid-2000s, where an increased number of startup firms would receive funding without the levels of governance that the VC firms previously employed. The authors attribute this change to recent innovations such as cloud computing services, which allowed investors to bypass purchasing expensive hardware for startups while the probability of the startup's success was still low. This change caused investments per year made by VCs in relevant sectors to nearly double. Startups in sectors where the "spray-and-pray" approach was used had a higher likelihood of failure, but if the startup were to receive another round of funding, it had almost a 20% greater increase in valuation than startups in untreated sectors.

Technological advancements have made it easier than ever to launch a startup. To properly analyze the potential of these startups, VC firms use a multi-stage selection process to weed through opportunities. Gompers et al. (2016) outline the so-called deal "pipeline"; startups are initially evaluated by a member of the VC firm, a VC member then meets with management from the startup, then the startup is brought to other members of the VC firm for additional review. If the startup gets past further evaluation, the VC firm then conducts a formal process of due diligence. A term sheet is then presented to the startup, and the deal is formally agreed to. In a given year, VC firms conduct initial evaluations in roughly 250 early-stage firms. Of those 250 startups, only 60 firms receive a visit from a VC member. Only a third of those 60 firms make it to partner review, and only five firms receive an offer sheet for a deal. This remains the primary process in which VCs evaluate talent because it puts a large emphasis on the meeting with management. Gompers et al. found that 95% of VC respondents mentioned management teams as an important factor in investing, with 47% of VCs listing management as the most important factor when evaluating a startup. In such a large market of startups, venture capital firms often turn to their perceptions of strong leadership to make investments.

The nature of venture capital dictates that not every investment needs to be successful, but available literature suggests that most unicorns are severely overvalued. Gornall & Strebulaev (2017) develop a model to fairly value VC-backed securities. It is challenging to properly value private unicorns due to their often extreme growth and illiquidity, so firms typically reach a new valuation every time it offers a new series of equity funding. VC firms typically mark up the value of their investments to the price of the most recent funding round, assuming that all of the company's shares have the same price as the most recently issued shares. Since each round of funding has different cash flow and control rights, this assumption is false. For example, financial service provider Square went public at a share price of 9 USD, 42 percent below the price Series E investors had paid for Square's equity when the company was private. However, Series E investors were contractually protected from risk and received extra shares until the stock hit a price of $18.56 per share. Since Series E shares paid out more than other shares in downside scenarios and at least as much in upside scenarios, it must be more worth more than the common stock shares. Faulty accounting in markups was not unique to Square; Gornall & Strebulaev found that 53 of the 116 firms studied lost their unicorn status once accounting for different cash flows in funding rounds. The paper calls into question the manner in which these unicorns are valued; abnormally high valuations behoof both the unicorn and the investor due to positive attention firms receive once they reach unicorn status.

Knowing that venture capital firms prioritize management as a basis for future success, it is important to assess the methods in which management can persuade VC firms into investing. Oral presentation skills have been proven to be an important factor in angel investors' initial screening investment decisions (Clark 2008). Using questionnaire data from a UK investor forum, Clark established a significant relationship between investors' evaluations of the content quality of entrepreneurs' presentations and the likelihood that the investors would be interested in pursuing an investment opportunity with the entrepreneurs. Presentational factors tended to have the strongest influence on investors' evaluations, although the investors' stated reasons for their evaluations were entirely based on non-presentational criteria, such as specific information about the company and its market. Anglin et al. (2018) demonstrated the power of positive language in public campaigns through a study of nearly two thousand Kickstarter campaigns, finding a strong correlation between language that evoked positive psychological capital and campaign success. A 10% increase in a campaign's use of positive psychological capital was associated with a 3% increase in the probability the campaign succeeded. Parhankangas & Renko (2017) also use Kickstarter campaign outcomes to analyze the rhetoric of social entrepreneurs. Their results indicate that specific and precise language along with interactivity is a strong predictor of campaign success. Crowdfunding platforms like Kickstarter provide a strong analog for startup pitches to VC firms. As with startup pitches, Kickstarter campaigns present a company's best portrayal of itself. Risk factors and poor financial data are absent from these pitch desks; instead, these companies use persuasion and exaggeration tactics to attract investors. When VC firms meet with management from potential portfolio companies, the meeting typically comes before the startup's financials are analyzed with due diligence. As a result, VCs are blind to any fundamental problems with the company before it meets with investors.

The extent to which linguistic tone influences venture capital funding has not been well studied; after all, pitches are closed-door and not publically accessible. Once a firm is public, however, there is a vast trove of publically accessible financial documents available to be analyzed. Loughran & McDonald (2011) applied linguistic analysis to financial documents by creating word banks to properly assess positive and negative sentiment in a financial context. Their research found that word lists that had been used in prior financial research such as the Harvard Psychological Dictionary did not accurately categorize words in a financial context. In a text analysis of more than 50 thousand corporate 10-K reports, almost 75% of negative word counts reported by the Harvard list were attributable to words that are typically not negative in financial contexts. Words such as CAPITAL, BOARD, and VICE were featured in Harvard's negative word bank, words that appear frequently in financial documents but not in a negative context. Loughran & McDonald developed their own word list of negative financial words, as well as banks for positive, uncertain, litigious, strong, and weak language.

Baginski et al. (2016) used financial word lists developed by Loughran & McDonald to analyze forward-looking earnings forecasts, voluntary disclosures from management regarding their firm's future financial performance. To determine if a statement's linguistic tones impacted future security performance, Baginski sought to establish a relationship between the management's emotional sentiment in the forecast and its ensuing adjusted return. The study finds that a one-standard deviation increase in a forecast's positive tone is associated with a 3.36% increase in its adjusted stock return. The correlation strengthens when the quantitative predictions of the forecast agree with its linguistic tone, specifically when a quantitative forecast predicting good news confirms the positive linguistic tone of the forecast.

Loughran & McDonald (2013) apply their own financial word lists to perform text analysis on S-1 IPO documents. Building on Beatty & Ritter's (1986) findings of a positive correlation between investor uncertainty about an IPO's value and its initial expected return, the paper tests the relationship between textual sentiment in firms' S-1 forms and their first-day security returns. Uncertain, negative, and weak language was found to have a significant correlation with first-day returns; a one-standard deviation increase in the proportion of uncertain and negative language in the S-1 was associated with a 3% and 4% increase in first-day returns, respectively. In the 60-day period following the offering, firms with more uncertain language in their S-1 were associated with higher volatilities in its stock returns, implying that firms with higher levels of uncertainty prove more difficult to properly value.

The role of venture capital in IPO filings is examined by de Carvalho et al. (2020) through the dynamics of earnings management. Earnings management is a term used to describe deceptive practices and techniques in the production of a firm's financial statements. This practice is often used to mask poor financial performance and typically occurs when management sets a predetermined target for earnings. The paper finds that firms with venture capital investment engage in less earning management than firms without VC investment. However, VC-sponsored firms engage in more earnings management than firms without VC investment in periods leading up to the firm's IPO. While non-VC sponsored firms tend to inflate earnings during the IPO period and deflate earnings during the lock-up and post-lockup period, VC-sponsored firms inflate earnings before its IPO and maintain inflation until the lock-up period.

4. Methodology

4a. Hypothesis

This thesis studies the rhetoric and performance of successfully-exited unicorns in an attempt to establish a relationship between venture capital investment, narrative, and security performance. I present three hypotheses in an attempt to answer my overarching research question: Can the language used by unicorn startup companies in their S-1 SEC filings predict its eventual stock performance once public?

Existing literature on the subjects suggests relevant details to a potential hypothesis. Analysis of public funding campaigns on Kickstarter by Anglin et al. as well as Parhankangas & Renko suggests that positive and confident language in a company's pitch is a costless and effective strategy in the effort to gain funding. Additionally, surveys of VC firms conducted by Gompers et al. indicate that a startup's management team is the most important component when evaluating future success. Considering these two findings, startup managers benefit vastly from using rhetorical strategies in their pitches to VC firms. In fact, the research suggests that VC firms may overlook a startup's red flags if the firm's management team is perceived to be persuasive. Uber, WeWork, and Theranos were all founded by strong personalities with grand visions for their company, and all three firms were able to amass billions of dollars in VC funding while masking poor fundamentals. This association implies that positive, persuasive language helps startups increase their likelihood of receiving VC funding, and thus contributes to a difference between the narrative that a company creates and its actual business strategy.

As an analog for a unicorn startup's pitch to investors, I will use the business summary section in S-1 IPO filings. This section of the S-1 provides an overview of the firm and includes details such as its mission statement, clients, supply chain, future strategy, and revenues. Given the nature of pre-IPO investment rounds, as well as the incentives startups have to use positive rhetorical signals in their pitches, I hypothesize that unicorn firms with high levels of venture capital investment will use more positive language in their business summaries relative to firms with lower levels of venture capital investment.

Within the text of the S-1 form, I look to establish a relationship between a firm's positive language, as demonstrated in its business summary, and its risk, as determined by negative and uncertain language in the risk factors section of the S-1. I hypothesize that firms with relatively high levels of negative language in its risk factors will have relatively high levels of positive language in its business summary; firms with higher disclosed risk will attempt to compensate by using positive signals in its pitch.

I will use the results of the S-1 text analysis to establish relationships between textual sentiment in financial disclosures and ensuing security returns. Drawing from the results found by Loughran & McDonald (2013), I hypothesize that firms with abnormal amounts of positive language in their business summary will have a negative association with ensuing security performance.

4b. Data

This paper studies every company that completed an initial public offering (IPO) on the New York Stock Exchange or NASDAQ at a market cap valuation of 1 billion USD or higher since 2000. I use Crunchbase, a data-as-a-service platform for public and private companies, to collect every requirement-meeting firm and its corresponding stock ticker. Using the SEC EDGAR database, I convert each firm's S-1 business summaries and risk factors into separate text files. Certain firms that met the requirement but did not file an S-1, F-1, or S-11 document were not included in the sample. A total of 233 firms were used.

Venture capital investment was determined using available data from Crunchbase, which collects private venture funding round data for all firms. Crunchbase reports each firm's top five pre-IPO investors, and I collected each observed firm's investors from the database.

Additionally, I collected each venture capital firm that was featured on CBInsights' 2019 Top 100 Venture Capitalists Ranking. CBInsights is an angel investing research platform that, in conjunction with the New York Times, releases yearly rankings of the 100 top venture capitalists in the world. 100 people from 64 firms were featured in the 2019 ranking. To determine a variable indicating venture capital investment in an observed firm, I created a binary variable which took a value of 1 if one of the firm's top five investors was included in the CBInsights ranking. 90 of the 233 observed firms qualified for this distinction.

Text analysis for both business summaries and risk factors are done through the Linguistic Inquiry and Word Count (LIWC) module, a proprietary software program that counts and processes words into psychologically-relevant categories. The analysis is conducted using two different word banks.

The first analysis is done using LIWC-created psychological categories. Table 1 displays the LIWC categories relevant to this thesis, its frequencies in different forms of text, as well as its frequencies within the sample of business summaries and risk factors. Outputs for business summaries and risk factors were taken as an average of all firms recorded in the sample, while output for other forms of text was taken from data reported in the 2015 LIWC language manual (Pennebaker et. al, 2015). Except for the four summary variables (Analytic, Clout, Authentic, and Tone), all mean values in Table 1 are expressed as a percentage of total words used in the sampled text. The four summary variables are percentiles based on standardized scores from large comparison samples. LIWC is not transparent about what types of language are factored into the four summary variables, but the variables are based on previous research conducted by the developers of the software. The categories are defined as such:

- **Clout**: category of words indicative of certainty, dominance, and confidence (Kacewicz et al., 2012)
- **Authenticity**: to what extent the language used is personal and self-revealing, rather

than detached and guarded (Newman et al., 2003)

- **Analytic**: the degree of analytical, logical and consistent thinking, as opposed to more intuitive, narrative writing (Pennebaker et al., 2014)
- **Tone**: the degree of positive emotional tone, as measured by the difference between LIWC scores for negative language and positive language (Cohn et al., 2004)
- **Emotion**: levels of positive and negative emotional words

	Prospectus Summaries	Risk Factors	Blogs	Nov-els	Natural Speech	New York Times
Clout	89.11	83.6	47.87	75.37	56.27	68.17
Authenticity	21.59	15.19	60.93	21.56	61.32	24.84
Analytic	93.07	86.31	49.89	70.33	18.43	92.57
Tone	65.09	45.82	54.5	37.06	79.29	43.61
Positive Emotion	2.88	2.95	3.66	2.67	5.31	2.32
Negative Emotion	0.73	1.87	2.06	2.08	1.19	1.45

Table 1: Prospectus Summary and Risk Factor LIWC textual sentiment as compared to other forms of text, in percentages

The second text analysis is done in the LIWC module using Loughran and McDonald's (2011) financial sentiment word lists. These word lists have been taken from a dictionary of words and word counts from all 10-Ks filed from 1994 to 2008. Average outputs for both business summary texts, as well as risk factors, are shown in Table 2. The categories for the word list is defined as such:

- **Fin-Neg**: negative sentiment in a financial context. The category includes 2,337 words.
- **Fin-Pos**: positive sentiment in a financial context. The category includes only 353 words, substantially lower than the number of words in the Fin-Neg category. Loughran

& McDonald attempted to remove any qualifying positive words, focusing only on words unilateral in a positive tone.

- **Fin-Unc**: words denoting uncertainty, specifically imprecision. The category includes 285 words.
- **MW-Strong** and **MW-Weak**: strong and weak modal words as developed by Jordan (1999). There are 19 MW-Strong words and 27 MW-Weak words.

	Prospectus Summaries	Risk Factors
Fin-Neg	0.72	3.3
Fin-Pos	1.81	0.96
Fin-Unc	0.88	0.96
MW-Strong	0.36	0.58
MW-Weak	0.32	2.12

Table 2: Prospectus Summary and Risk Factor textual sentiment in Loughran & McDonald's word banks, in percentages

To compare text analysis outputs to successive security returns, I used Koyfin, a financial research database, to calculate post-IPO stock returns. Returns were found for each firm the day, week, month, and year after each firm's stock went public. Daily returns were calculated by finding the percentage change in security price from the beginning of the day's trading to the close. Each percentage return was then subtracted by the market's return for the given period of time to get the abnormal return of the security. Since every return observed was following the stock's IPO, the beta value for each stock is assumed to be 1. The market return was calculated using the S&P 500, a market index that measures the stock performance of 500 large-cap stocks. A company's assigned industry was determined by Financial Visualizations, another financial research database.

4c. Empirical Strategy

To determine the effect of venture capital investment on a firm's S-1 language, I estimate the following OLS regression models:

$Y_{ib} \;=\; \beta_0 \;+\; \beta_1 VC_i \;+\; \gamma_i \;+\; \epsilon_i$

$Y_{ir} \;=\; \beta_0 \;+\; \beta_1 VC_i \;+\; \gamma_i \;+\; \epsilon_i$

Where:

- $\mathbf{Y_{ib}}$ measures language output in the business summary section of each observation's S-1 document
- $\mathbf{Y_{ir}}$ measures language output in the risk factor section of each observation's S-1 document
- $\mathbf{VC_i}$ is a dummy variable indicating whether a top venture capital firm is one of the observed firm's five largest investors.
- $\boldsymbol{\gamma_i}$ is a dummy variable controlling for a firm's industry. Sectors included are technology, finance, services, consumer goods, basic materials, healthcare, industrial goods, and utilities.
- $\mathbf{e_i}$ is an error variable.

This model addresses my hypothesis that firms with investment from top VC firms will use more positive language in their business summaries relative to firms without venture capital investment.

To determine the effect of a firm's risk on the types of language it uses when summarizing its business, I estimate the following OLS regression model:

$$Y_{ib} \;=\; \beta_0 \;+\; \beta_1 X_{ir} \;+\; \gamma_i \;+\; \epsilon_i$$

Where:

- $\mathbf{Y_{ib}}$ measures language output in the business summary section of each observation's

S-1 document

- $\mathbf{X_{ir}}$ measures language output in the risk factor section of each observation's S-1 document
- $\boldsymbol{\gamma_i}$ is a dummy variable controlling for a firm's industry. Sectors included are technology, finance, services, consumer goods, basic materials, healthcare, industrial goods, and utilities.
- $\mathbf{e_i}$ is an error variable.

This model addresses my hypothesis that firms with relatively high levels of negative language in their risk factors will have relatively high levels of positive language in its business summary.

To determine the effect of a firm's S-1 textual sentiment on its stock performance once public, I estimate the following regression model:

$$AR_{it} \; = \; \beta_0 \; + \; \beta_1 X_{ip} \; + \; \gamma_i \; + \; \epsilon_i$$

Where:

- $\mathbf{AR_{it}}$ measures the return of the firm's stock price in percentages when adjusted by the market's performance in the same time, $\mathrm{AR_{it}} = \mathrm{R_{it}} - \mathrm{R_{mt}}$
- $\mathbf{X_{ib}}$ measures language output in the business summary section of each observation's S-1 document
- $\boldsymbol{\gamma_i}$ is a dummy variable controlling for a firm's industry
- $\mathbf{e_i}$ is an error variable

This model addresses my hypothesis that abnormal amounts of positive language in their business summary will have a negative association with its security performance after going public.

5. Results

Table 3: Impact of Venture Capital Investment on S-1 LIWC Language

	(1)	(2)	(3)	(4)
	posemo_ps	posemo_ps	negemo_ps	negemo_ps
top_5_vc	0.261**	0.219*	0.00617	0.0332
	(0.0920)	(0.0963)	(0.0526)	(0.0477)
technology		0.529***		0.0210
		(0.0887)		(0.0596)
financial		0.799***		0.189*
		(0.0988)		(0.0902)
services		0.524***		
		(0.0879)		
consumer_goods		0.623***		-0.191**
		(0.153)		(0.0620)
basic_materials		-0.218**		0.310**
		(0.0786)		(0.108)
healthcare		0.00119		0.289***
		(0.0875)		(0.0840)
industrial_goods		-0.0886		0.0600
		(0.202)		(0.0998)
utilities		0.685		0.00633
		(0.420)		(0.123)
_cons	2.850***	2.400	0.615***	0.543***
	(0.0529)	(.)	(0.0297)	(0.0551)
N	232	232	232	232
adj. R^2	0.032	0.153	-0.004	0.066

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 displays the regression output testing the hypothesis that venture capital investment affects a firm's use of emotional language when summarizing their business, using the Linguistic Inquiry and Word Count module. The variables tested are defined as follows:

- **top_5_vc** is a binary variable that indicates whether the observed firm has received investment from a top venture capital firm, namely whether a top VC firm was one of the observed firm's top five investors.

- **posemo_bs** is a variable totaling the percentage of positive emotional language in a firm's business summary section, as determined by the Linguistic Inquiry and Word Count word bank.
- **negemo_bs** is a variable totaling the percentage of negative emotional language in a firm's business summary section, as determined by the Linguistic Inquiry and Word Count word bank.

The regression establishes a significant positive correlation between venture capital investment in a firm and the firm's levels of positive language in its business summaries. Firms that receive investment from an elite venture capital firm are associated with a 0.219% increase in levels of positive emotional language in its S-1 business summary, controlling for the firm's industry. This effect is significant at the 1% level without industry controls, and at the 5% level with industry controls. Firms that receive investment from an elite venture capital firm are associated with a 0.0332% increase in levels of negative emotional language in its S-1 business summary, controlling for the firm's industry. This effect is not significant at the 5% level and contains standard errors higher than coefficients when controlling for industry.

When using Loughran & McDonald's financial sentiment word lists, I found both a positive and negative impact of venture capital investment on textual sentiment in business summaries. The coefficients for both positive and negative language are very weak, as the standard errors of the regression were larger than the coefficients when industry controls are applied. The LIWC word bank is a more useful variable for this analysis than Loughran & McDonald's financial sentiment dictionary because the LIWC list acts as a more effective gauge on positive psychological capital. Loughran & McDonald's dictionary is not a relevant variable for business summaries because financial context is not necessary for business summaries. Summaries are written in a manner to catch the reader's attention and retain engagement, so a general text analysis that captures emotional sentiment serves as a more effective gauge for positive language than a word bank that only captures positive and negative financial sentiment.

Table 4: Impact of Venture Capital Investment on S-1 LIWC Summary Variables

	(1)	(2)	(3)	(4)
	log_analytic_bs	log_authentic_bs	log_clout_bs	log_tone_bs
top_5_vc	-0.0143***	-0.0451	0.00952	0.0383
	(0.00352)	(0.0452)	(0.0111)	(0.0292)
technology	-0.0492***	0.474***	0.140***	0.0374
	(0.00302)	(0.0440)	(0.00902)	(0.0274)
financial	-0.0330***	0.427***	0.118***	0.0830**
	(0.00374)	(0.0661)	(0.0203)	(0.0268)
services	-0.0479***	0.646***	0.150***	0.0428
	(0.00357)	(0.0429)	(0.0148)	(0.0319)
consumer_goods	-0.0595***	0.482***	0.149***	0.129***
	(0.00775)	(0.104)	(0.0280)	(0.0360)
basic_materials	-0.0385***	0.595***	-0.00370	-0.264***
	(0.00848)	(0.0838)	(0.0454)	(0.0447)
healthcare	-0.0103*	0.381***	-0.119***	-0.155***
	(0.00484)	(0.0560)	(0.0348)	(0.0312)
industrial_goods	-0.0449***	0.695***	0.107***	-0.135
	(0.00803)	(0.106)	(0.0202)	(0.0688)
utilities	-0.0128*	0.428***	0.0250	0.0728
	(0.00532)	(0.0545)	(0.0138)	(0.108)
_cons	4.584***	2.366***	4.337	4.175
	(2.39e-09)	(2.66e-08)	(.)	(.)
N	232	232	232	232
adj. R^2	0.291	0.077	0.400	0.163

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4 displays the impact of venture capital investment on LIWC summary category variables. I have taken the log of each variable to prevent a non-linear relationship. The variables are defined as follows:

- **log_analytic** is a variable denoting the degree of analytical, logical and consistent thinking, as opposed to more intuitive, narrative writing (Pennebaker et al., 2014)
- **log_authenticity** is a variable denoting to what extent the language used is personal and self-revealing, rather than detached and guarded (Newman et al., 2003)
- **log_clout** is a variable denoting words indicative of certainty, dominance, and confidence (Kacewicz et al., 2012)
- **log_tone** is a variable denoting the degree of positive emotional tone, as measured by the difference between LIWC scores for negative language and positive language (Cohn et al., 2004)

The regression finds a significant negative correlation between venture capital investment and the degree of a firm's analytic writing in its S-1 business summary. Firms that receive investment from an elite venture capital firm are associated with a 1.64% decrease in levels of analytical language in its S-1 business summary, as established by Pennebaker et al (2014). Pennebaker used the Categorical-Dynamic Index to develop a binary metric contrasting cognitive complexity (greater article and preposition use) with time-based narrative writing, which uses more pronouns and auxiliary verbs. A negative coefficient for levels of analytical language implies that firms with venture capital investment are more likely to use a dynamic, narrative style in their business summary, controlling for industry. This correlation is significant at the 0.1% level. There is no discernable connection between venture capital investment and business summary language for the three other summary variables. Most regressions have a very low R^2 and contain high standard errors.

Table 5 displays the impact of venture capital investment on textual sentiment in the risk factor section of S-1s, as measured by both the LIWC and Loughran & McDonald word bank. The dependent variables are defined as follows:

- **posemo_rf** is a variable totaling the percentage of positive emotional language in a

Table 5: Impact of VC Investment on S-1 Risk Factor Language

	(1)	(2)	(3)	(4)	(5)	(6)
	posemo_rf	negemo_rf	liwc_rf_diff	all_positive_rf	all_negative_rf	lm_rf_diff
top_5_vc	-0.0964	0.0741	-0.171*	0.0179	0.220*	-0.161*
	(0.0507)	(0.0428)	(0.0748)	(0.0310)	(0.104)	(0.0648)
technology	0.547***	0.313***	0.234**	0.0677	0.0904	-0.0535
	(0.0494)	(0.0410)	(0.0709)	(0.0384)	(0.124)	(0.0828)
financial	0.809***	0.294***	0.515***	0.0201	-0.116	0.0436
	(0.0657)	(0.0620)	(0.102)	(0.0600)	(0.185)	(0.134)
services	0.651***	0.221***	0.430***			
	(0.0536)	(0.0392)	(0.0675)			
consumer_goods	0.453***	0.341***	0.112	0.111	0.0742	-0.0833
	(0.0991)	(0.0948)	(0.155)	(0.0667)	(0.220)	(0.164)
basic_materials	0.518***	0.287***	0.230	0.00169	0.112	0.0130
	(0.0875)	(0.0847)	(0.128)	(0.0549)	(0.236)	(0.159)
healthcare	0.801***	0.0946	0.706***	0.288***	0.214	0.0914
	(0.0869)	(0.0513)	(0.125)	(0.0648)	(0.141)	(0.102)
industrial_goods	0.697***	0.466***	0.231	0.00598	0.705*	-0.351
	(0.131)	(0.0771)	(0.180)	(0.0640)	(0.299)	(0.196)
utilities	0.744***	0.124	0.620*	-0.0453	-0.0750	0.0517
	(0.153)	(0.0979)	(0.250)	(0.0878)	(0.249)	(0.136)
_cons	2.370***	1.570***	0.800***	1.468***	7.105***	-2.266***
	(5.32e-08)	(6.79e-08)	(6.66e-08)	(0.0331)	(0.0992)	(0.0744)
N	232	232	232	232	232	232
adj. R^2	0.098	0.050	0.101	0.090	0.024	0.018

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

firm's risk factor section, as determined by the Linguistic Inquiry and Word Count word bank.

- **negemo_rf** is a variable totaling the percentage of negative emotional language in a firm's risk factor section, as determined by the Linguistic Inquiry and Word Count word bank.
- **liwc_rf_diff** is a variable calculating the difference between the percentage of positive emotional language and the percentage of negative emotional language in risk factors, as determined by the Linguistic Inquiry and Word Count word bank. A positive output for this variable indicates that the risk factor text contains more positive than negative language.
- **all_positive_rf** is a variable totaling the percentage of positive financial language in a firm's risk factor section, as determined by Loughran and McDonald's financial sentiment word bank. This includes Laughran and McDonald's positive and strong modal word lists.
- **all_negative_ rf** is a variable totaling the percentage of negative financial language in a firm's risk factor section, as determined by Loughran and McDonald's financial sentiment word bank. This includes Laughran and McDonald's negative, uncertain, litigious, and weak modal word lists.
- **lm_rf_diff** is a variable calculating the difference between the percentage of positive financial language and the percentage of negative financial language in risk factors, as determined by Loughran and McDonald's financial sentiment word bank.

The risk factor section of the S-1 is intended to disclose significant factors that make a firm's IPO especially risky or uncertain. Companies are legally required by Item 503(c) of Regulation S-K to disclose all potential risks that the company faces. In this paper, I use the risk factor section as a proxy for a firm's perceived risk. The regression output indicates that for both the LIWC and L&M word bank, there is a significant positive correlation between venture capital investment and negative language in S-1 risk factors. Given that Laughran and McDonald's financial dictionary is specifically equipped to analyze financial texts, it is important that significance was found at the 5% level for both negative financial language and the difference between positive and negative language. These findings imply that unicorn firms with venture capital investment contain higher levels of risk than firms without investment.

Table 6: Impact of Negative Language in Risk Factors on Positive Language in Business Summaries

	(1)	(2)	(3)	(4)
	liwc_bs_emodiff	lm_bs_emodiff	liwc_bs_emodiff	lm_bs_emodiff
all_negative_rf	-0.126	-0.0684	-0.0929	-0.0726
	(0.0690)	(0.0683)	(0.0671)	(0.0681)
technology			0.0805	0.0163
			(0.141)	(0.144)
financial			0.0697	-0.404*
			(0.155)	(0.172)
consumer_goods			0.312	0.362*
			(0.198)	(0.172)
basic_materials			-1.077***	-0.698*
			(0.163)	(0.283)
healthcare			-0.726***	-0.857***
			(0.135)	(0.190)
industrial_goods			-0.643*	-0.0303
			(0.251)	(0.362)
utilities			0.163	-0.639**
			(0.432)	(0.213)
_cons	3.248***	1.133*	3.078***	1.298**
	(0.505)	(0.485)	(0.474)	(0.462)
N	233	233	233	233
adj. R^2	0.009	-0.000	0.162	0.125

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6 displays the relationship between language within the S-1. I look to establish a relationship between positive language in business summaries and negative language in risk factors to see if firms that use higher levels of positive language in their business summaries are associated with higher levels of negative language in their risk factors. I hypothesize that firms with higher levels of negative sentiment in their risk factors will be associated with higher levels of positive language in their business summaries, as riskier firms may be incentivized to use rhetorical strategies to make their prospects more appealing to investors. The regression shows no significant connection between the two variables for either dictionary. Both word banks have negative coefficients, implying that firms with higher levels of negative risk factor language also have higher levels of negative language in business summaries. These findings stand in stark contrast to my hypothesis, which suggested that the opposite would occur. The standard errors for these regressions are also very high, suggesting that there is not a significant connection between the two parts of the S-1.

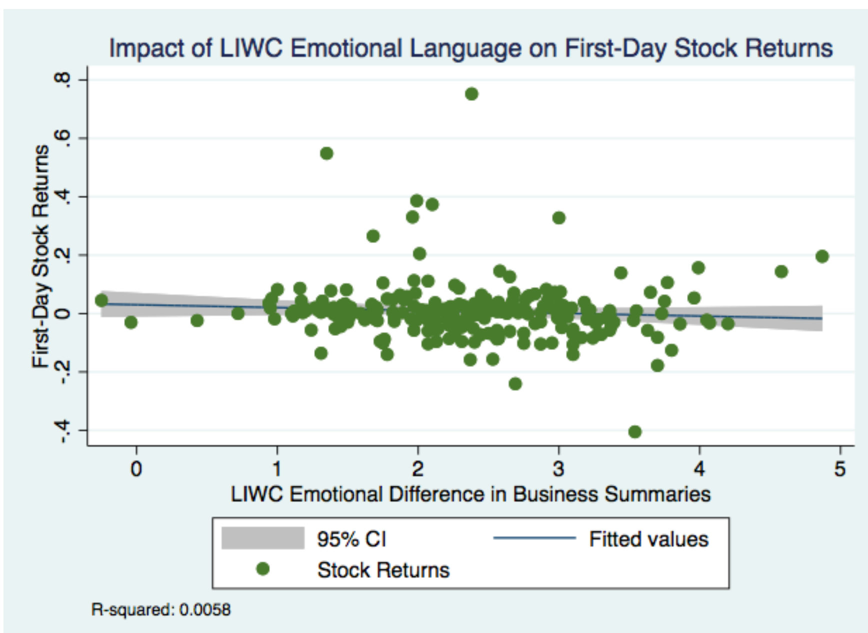

Figure 3: Impact of LIWC Emotional Language on First-Day Stock Returns

Figure 3 displays the relationship between a firm's net positive emotional language in its business summary (measured in the percentage of the entire text) and its first-day stock returns (measured in decimals) once public. The graph shows a weak negative correlation between the two variables; a one percent increase in net positive emotional language in business summaries results in a 0.9% decrease in the stock's abnormal return on its first day of public trading. This regression has high standard errors, which increase when industry controls are applied, and an extremely low R^2 output.

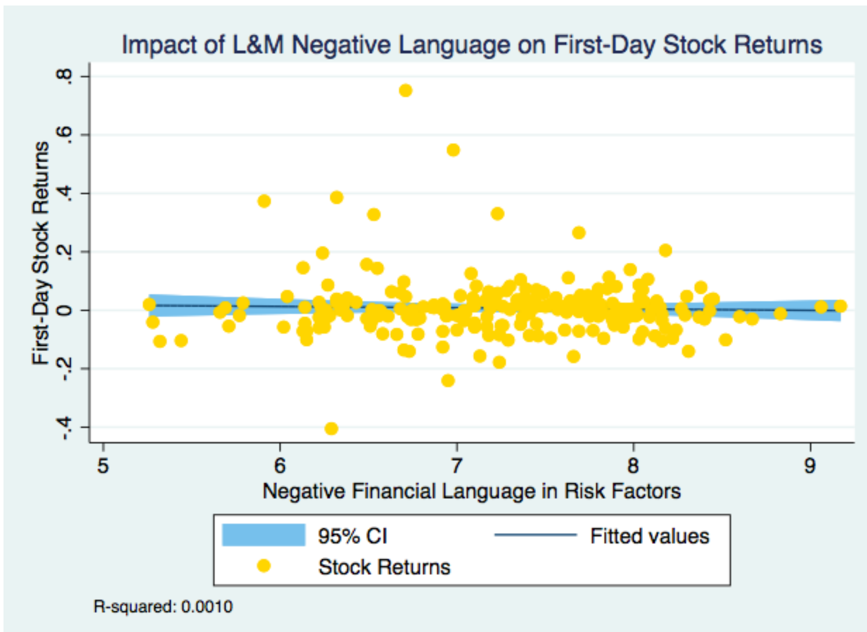

Figure 4: Impact of L&M Negative Financial Language on First-Day Stock Returns

Figure 4 shows the relationship between a firm's negative financial language in its risk factors (measured in the percentage of the entire text) and its first-day stock returns (measured in decimals) once public. There is virtually no correlation between the two variables, as the standard errors of this regression are twice as large as the coefficient. The regressions in Figure 3 and Figure 4 are both adversely affected by large outliers. However, even when observations with stock returns exceeding 30% and lower than 20% were dropped from the regression, there were no significant correlations established between S-1 language and daily returns.

Table 7: Impact of Language and Investment on Weekly and Monthly Stock Returns

	(1) mar_week	(2) mar_week	(3) mar_week	(4) mar_month	(5) mar_month	(6) mar_month
liwc_bs_emodiff	-0.0320**			-0.000915		
	(0.0111)			(0.0268)		
technology	-0.0118	-0.0149	-0.0127	-0.0421	-0.0480	-0.0316
	(0.0203)	(0.0205)	(0.0223)	(0.0391)	(0.0426)	(0.0464)
financial	-0.00291	-0.00673	-0.00841	-0.0259	-0.0327	-0.0324
	(0.0276)	(0.0274)	(0.0267)	(0.0435)	(0.0472)	(0.0407)
consumer_goods	0.101*	0.0908	0.0897	0.168*	0.162*	0.165*
	(0.0491)	(0.0481)	(0.0491)	(0.0785)	(0.0753)	(0.0741)
basic_materials	-0.0226	0.0119	0.00784	0.0377	0.0380	0.0269
	(0.0278)	(0.0241)	(0.0238)	(0.0365)	(0.0419)	(0.0377)
healthcare	0.0154	0.0378	0.0396	0.0773	0.0692	0.0848
	(0.0369)	(0.0379)	(0.0374)	(0.0695)	(0.0769)	(0.0777)
industrial_goods	0.0306	0.0537*	0.0490*	0.0643	0.0676	0.0531
	(0.0238)	(0.0231)	(0.0225)	(0.0501)	(0.0531)	(0.0521)
utilities	-0.00531	-0.0120	-0.0127	-0.00532	-0.0122	-0.00801
	(0.0234)	(0.0208)	(0.0186)	(0.0527)	(0.0554)	(0.0569)
log_authentic_bs		-0.00632			-0.0338	
		(0.0262)			(0.0505)	
top_5_vc			-0.00888			-0.0373
			(0.0212)			(0.0384)
_cons	0.0838**	0.0254	0.0108	0.0199	0.119	0.0295
	(0.0279)	(0.0810)	(0.0137)	(0.0494)	(0.171)	(0.0280)
N	233	233	232	233	233	232
adj. R^2	0.045	0.011	0.011	0.035	0.037	0.042

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7 shows the relationship between relevant variables and both weekly and monthly stock returns. There is a significant negative correlation between positive emotional language and weekly stock returns; a one percent increase in net levels of positive emotional language in a firm's business summary results in a 3.2% decrease in weekly returns. The effect remains negative for monthly returns but is very small and not significant. The impact of dynamic language on stock returns is weakly negative for both weekly and monthly returns. There is no discernable correlation between venture capital investment and short-term stock returns post-IPO.

6. Conclusion

Presentation and rhetoric are both vital components of a company's pitch to investors. This paper collects and analyzes venture capital interest in every firm since 2000 that went public on the New York Stock Exchange or NASDAQ at a market cap of 1 billion USD or above. Using text analysis software, an OLS regression model was estimated to evaluate the relationship between venture capital investment in unicorns and unicorns' rhetorical strategies in their business summary. The regression found that venture-funded firms use higher levels of positive emotional language when summarizing their business relative to firms without venture funding. Venture-funded firms are also more likely to use dynamic, time-based narrative writing in their summaries. Firms that receive funding are correlated with higher levels of negative financial language in the risk factor section of their S-1, an analog to a firm's risk. Despite the significant relationship between VC investment and rhetorical text strategy, there is no observable connection between VC investment and post-IPO stock returns. Additionally, it does not appear that the rhetorical strategy in the business summary section of the S-1 has a significant influence on stock returns once the firm is public.

Existing literature supports the relationship between venture capital investment and rhetorical strategies. Previous research suggests that venture capital firms prioritize management when selecting startups to invest in (Gompers et al. 2016). Due to the nature of the VC process, a startup's initial pitch to investors has large weight to the likelihood of eventual investment. This stresses the significance of an effective pitch; Clark (2008) noted that investors are more likely to show interest in investing in a startup that had a strong presentation, regardless of the strength of the company or its market fit. The implication that venture capital firms are more likely to invest in firms with rhetorical strength suggests that my findings could suffer from simultaneity bias. In this case, it is unclear whether venture capital investment causes startup companies (that later become unicorns) to build a rhetorical narrative about its firm, or if venture capitalists tend to invest in firms that already have strong rhetoric. It is also worth considering that out of the hundreds of startup companies that elite venture capital firms invest in, the unicorns studied in this paper are already considered extreme successes. The eventual goal for venture capital investments is to lead startups to a successful exit in the form of an IPO, as this is one of the only ways for investors to reap the monetary rewards of early investments. As a result, my data set may suffer from sampling bias.

Further research could move beyond the context of exited unicorns and develop a much larger sample size to analyze. This would likely allow for a more accurate assessment of the relationship between VC investment and stock returns. Additional time would help grow a larger sample size for exited unicorns, as there are currently more unicorns that are privately-held than publically-held. The rapid, recent growth of unicorns calls for further investigation regarding these special firms, which will help specify academic research about this subject.

References

Aiello, Chloe. "Theranos President Exaggerated the Company's Revenue by 1,000 Times to Investors, Says SEC." CNBC. CNBC, March 14, 2018. https://www.cnbc.com/2018/03/14/theranos-president-exaggerated-revenue-by-1000x-says-sec.html.

Anglin, Aaron H., Jeremy C. Short, Will Drover, Regan M. Stevenson, Aaron F. Mckenny, and Thomas H. Allison. "The Power of Positivity? The Influence of Positive Psychological Capital Language on Crowdfunding Performance." *Journal of Business Venturing* 33, no. 4 (2018): 470–92. https://doi.org/10.1016/j.jbusvent.2018.03.003.

Baginski, Stephen, Elizabeth Demers, Chong Wang, and Julia Yu. "Contemporaneous Verification of Language: Evidence from Management Earnings Forecasts." *Review of Accounting Studies* 21, no. 1 (September 2016): 165–97. https://doi.org/10.1007/s11142-015-9347-6.

Beatty, Randolph P., and Jay R. Ritter. "Investment Banking, Reputation, and the Underpricing of Initial Public Offerings." *Journal of Financial Economics* 15, no. 1-2 (1986): 213–32. https://doi.org/10.1016/0304-405x(86)90055-3.

Carreyrou, John. "Theranos Cost Business and Government Leaders More Than $600 Million." The Wall Street Journal. Dow Jones & Company, May 4, 2018. https://www.wsj.com/articles/theranos-cost-business-and-government-leaders-more-than-600-million-1525392082.

Clark, Colin. "The Impact of Entrepreneurs Oral 'Pitch' Presentation Skills on Business Angels Initial Screening Investment Decisions." *Venture Capital* 10, no. 3 (2008): 257–79. https://doi.org/10.1080/13691060802151945.

Cohn, M. A., M. R. Mehl, and J. W. Pennebaker. "Linguistic Markers of Psychological Change Surrounding September 11, 2001." *Psychological Science* 15, no. 10 (January 2004): 687–93. https://doi.org/10.1111/j.0956-7976.2004.00741.x.

Conger, Kate. "Uber Posts $1.2 Billion Loss as Growth Improves." The New York Times. The New York Times, November 4, 2019. https://www.nytimes.com/2019/11/04/technology/uber-earnings-loss.html.

de Carvalho, Antonio Gledson & Pinheiro, Roberto B. & Sampaio, Joelson Oliveira, 2020. "The dynamics of earnings management in IPOs and the role of venture capital," *Research in International Business and Finance*, Elsevier, vol. 51(C).

Eaglesham, Jean, and Eliot Brown. “WeWork Investors Turned Off by 'Sloppy' IPO Filings.” The Wall Street Journal. Dow Jones & Company, October 7, 2019. https://www.wsj.com/articles/wework-investors-turned-off-by-sloppy-ipo-filings-11570440674?mod=hp_lead_pos1.

Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf. “Cost of Experimentation and the Evolution of Venture Capital.” *Journal of Financial Economics* 128, no. 3 (2018): 422–42. https://doi.org/10.1016/j.jfineco.2018.03.001.

Galloway, Scott. “Yogababble.” No Mercy / No Malice, September 27, 2019. https://www.profgalloway.com/yogababble.

Gompers, Paul, William Gornall, Steven Kaplan, and Ilya Strebulaev. “How Do Venture Capitalists Make Decisions?” *European Corporate Governance Institute*, August 2016. https://doi.org/10.3386/w22587.

Gornall, William, and Ilya Strebulaev. “Squaring Venture Capital Valuations with Reality.” *Journal of Financial Economics*, July 11, 2017. https://doi.org/10.3386/w23895.

Healthcare is the Leading Cause of Bankruptcy: Elizabeth Holmes @ TEDMED 2014. TED, 2014.

Hellmann, Thomas F., and Manju Puri. “Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence.” *SSRN Electronic Journal*, September 2000. https://doi.org/10.2139/ssrn.243149.

Jordan, R.R., 1999, *Academic Writing Course* (Longman, London).

Kacewicz, Ewa, James W. Pennebaker, Matthew Davis, Moongee Jeon, and Arthur C. Graesser. “Pronoun Use Reflects Standings in Social Hierarchies.” *Journal of Language and Social Psychology* 33, no. 2 (2013): 125–43. https://doi.org/10.1177/0261927x13502654.

Landy, Heather. “WeWork by the Numbers.” Quartz at Work. Quartz, December 17, 2018. https://qz.com/work/1498781/wework-by-the-numbers/.

Lee, Aileen. "Welcome To The Unicorn Club: Learning From Billion-Dollar Startups." TechCrunch. Verizon Media, November 2, 2013. https://techcrunch.com/2013/11/02/welcome-to-the-unicorn-club/.

Loughran, Tim, and Bill Mcdonald. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance* 66, no. 1 (June 2011): 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x.

Loughran, Tim, and Bill Mcdonald. "IPO First-Day Returns, Offer Price Revisions, Volatility, and Form S-1 Language." *Journal of Financial Economics* 109, no. 2 (2013): 307–26. https://doi.org/10.1016/j.jfineco.2013.02.017.

Newman, Matthew L., James W. Pennebaker, Diane S. Berry, and Jane M. Richards. "Lying Words: Predicting Deception from Linguistic Styles." *Personality and Social Psychology Bulletin* 29, no. 5 (2003): 665–75. https://doi.org/10.1177/0146167203029005010.

Parhankangas, Annaleena, and Maija Renko. "Linguistic Style and Crowdfunding Success among Social and Commercial Entrepreneurs." *Journal of Business Venturing* 32, no. 2 (2017): 215–36. https://doi.org/10.1016/j.jbusvent.2016.11.001.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. "The development and psychometric properties of LIWC2015." (2015). Austin, TX: University of Texas at Austin.

Pennebaker, James W., Cindy K. Chung, Joey Frazee, Gary M. Lavergne, and David I. Beaver. "When Small Words Foretell Academic Success: The Case of College Admissions Essays." *PLoS ONE* 9, no. 12 (2014). https://doi.org/10.1371/journal.pone.0115844.

Rowley, Jason D. "Here's When $100M Venture Rounds Took Over." Crunchbase News. Verizon Media, August 21, 2018. https://news.crunchbase.com/news/heres-when-100m-venture-rounds-took-over/.

Rowley, Jason D. "The Q4/EOY 2019 Global VC Report: A Strong End To A Good, But Not Fantastic, Year." Crunchbase News. Verizon Media, January 10, 2020. https://news.crunchbase.com/news/the-q4-eoy-2019-global-vc-report-a-strong-end-to-a-good-but-not-fantastic-year/.

"The Crunchbase Unicorn Leaderboard." TechCrunch. Verizon Media. Accessed March 16, 2020. https://techcrunch.com/unicorn-leaderboard/.

"The Top 100 Venture Capital Partners & Firms." CB Insights Research, December 13, 2019. https://www.cbinsights.com/research/top-venture-capital-partners/#2019.

Thomas, Katie, and Reed Abelson. "Elizabeth Holmes, Theranos C.E.O. and Silicon Valley Star, Accused of Fraud." The New York Times. The New York Times, March 14, 2018. https://www.nytimes.com/2018/03/14/health/theranos-elizabeth-holmes-fraud.html.

Uber Technologies, Inc. (2019). *Form S-1 Registration Statement*. Retrieved from SEC EDGAR website http://www.sec.gov/edgar.shtml

The We Company (2019). *Form S-1 Registration Statement*. Retrieved from SEC EDGAR website http://www.sec.gov/edgar.shtml