How Does S-1 Filing Language Affect IPO Failure?

Drew Woodfolk

Washington and Lee University

Accounting Honors Thesis

ABSTRACT: This paper considers how the language used in the filings associated with an initial public offering (IPO) might predict the ultimate failure of an IPO. In part one of the study, I used qualitative methods to examine the language used in the S-1 business section for a set of failed IPO firms. This examination resulted in the creation of three language themes relevant to the set of failed firms: Innovation Language, Opportunity Language, and Stewardship Language. In part two of the study, I used quantitative methods to explore the effects of each identified language theme on subsequent IPO failure as measured by total return and likelihood of class action litigation. I found that Innovation language did not have a significant effect on either measure of IPO failure. Opportunity Language did not have a significant effect on return failure, but this language was associated with significant changes in litigation failure rates. I also found that increased use of Stewardship Language was associated with higher IPO returns but had no significant effect on a firm's risk of litigation. This paper contributes to the existing literature by offering new insights into the relationship between language usage in corporate filings and firm performance.

INTRODUCTION

An initial public offering, or IPO, marks the first time a previously private company offers shares of ownership to the public market. This event is one of the most significant activities in a firm's history and is often viewed as the finish line for up-and-coming private businesses. The funds raised through the IPO process can be deployed for a variety of corporate purposes, including capital expenditures and operational expansion, just to name a few. In recent years, the number of IPOs launched to market has reached all-time highs. The 407 completed IPOs in 2020 is only rivaled by the IPO market of the late 90's and early 00's which saw 486 and 429 IPOs respectively[1]. Furthermore, the size of these public offerings appears to be increasing as well. In 2020, IPOs successfully raised $76.3 billion in the United States alone. This figure is only rivaled by the $108.1 billion raised by IPOs in 2000[2]. An onlooker might assume that with IPO volume and size approaching all-time highs, the number of high-quality businesses going public must be approaching all-time highs as well. This assumption could not be further from reality. The percentage of profitable IPO firms has decreased from a high of 81% in 2009 to a low of 22% in 2020[3].

-- INSERT FIGURE 1 HERE –

This decline suggests that both institutional and retail investors are lowering their standards for the quality of firm that is ready to be traded publicly. Some in the business press have even commented that the current "blistering IPO market is rekindling dot-com era froth fears".[4] As investor standards for IPO firms appear to have decreased, several high-profile IPOs

---

[1] https://www.statista.com/topics/1272/ipos/#dossierKeyfigures
[2] https://www.statista.com/statistics/264607/ipo-volume-in-the-us/
[3] https://www.statista.com/statistics/914724/profitable-companies-after-ipo-usa/
[4] https://www.bloomberg.com/news/articles/2020-09-30/blistering-ipo-market-is-rekindling-dot-com-era-froth-fears

have destroyed vast amounts of shareholder value or failed to launch all together. Luckin Coffee, a Chinese coffee retailer, is a prime example of the drastic shareholder value destruction that can occur during an IPO. The firm, which was founded in 2017, quickly established itself as a dominant player in the rapidly growing Chinese coffee market. It subsequently went public on the Nasdaq in 2019, just two years after its founding. The firm was poised to become a formidable competitor to Starbucks in China and gained the attention of Western investors seeking exposure to world's second largest economy. A major part of the company's appeal to investors was its willingness to adopt technology as a major part of its business model. All customer transactions were facilitated using the firm's proprietary mobile app and Luckin claimed to have a competitive advantage over its competitors by using artificial intelligence (AI) to locate optimal store locations and target discounts to specific customers. It was revealed in early 2020 that Luckin had engaged in a major accounting fraud, and its shares were subsequently delisted from the Nasdaq.

Another high-profile example of IPO failure could be seen at WeWork. Before announcing its intentions to go public, the firm commanded a valuation of $47 billion, was backed by the Japanese conglomerate SoftBank, and was already a darling of Wall Street[5]. The firm's high-profile CEO was touted as a visionary who had the audacity to disrupt the legacy real estate industry with an innovative office arbitrage business model. A major part of the firm's narrative was the importance of technology to its core operations. The firm maintained a full team of data scientists and developers on its payroll and even went so far as to mention "technology" 110 times in its prospectus.[6] While the firm seemed poised to go public at a tech-type valuation, the IPO was pulled at the last minute as investor interest seemed to evaporate

[5] https://markets.businessinsider.com/news/stocks/peloton-joins-uber-lyft-wework-ipo-flops-and-failures-2019-9
[6] https://hbr.org/2019/08/no-wework-isnt-a-tech-company-heres-why-that-matters

overnight. Upon release of the firm's S-1, investors were taken aback by the firm's poor financial performance and the questionable actions taken by its CEO Adam Neumann.

The common thread between these two high profile cases of failed IPOs is that both firms placed emphasis on their capabilities as technology companies despite operating in industries that were not particularly technical in nature (the Beverage and Real Estate industries, respectively). Despite these high-profile cases of IPO failure, companies operating in legacy industries continue to market themselves to investors as technology companies at heart. One example of this trend is the fast casual salad chain Sweetgreen, which announced its intent to go public in October, 2021[7]. To market itself as a technology company, the firm released the following statement in early 2016:

> "We've always acted more like a tech company than a food one. Like many tech companies, we want to disrupt. Like many tech companies, we raised venture capital (from the same people who invest in actual tech companies). And now, like all tech companies, we have an in-house tech team. And we just released a sleek new app (more on that in a minute)."[8]

These examples suggest that IPO firms today may be misrepresenting themselves as technology-focused companies to make their offerings stand out and to command higher valuations. Inspired by this observation, this research project aims to explore the question of how the language employed by firms undergoing the IPO process may be used to predict the potential for IPO failure in the future.

Based on existing literature, we know that the language used in corporate disclosures can be a useful predictor of firm performance. The words, tone, and level of obfuscation employed by company leadership in public filings predicts share price performance (Benson 2015; Feldman et al., 2008), future cash flows (Wyatt 2014), and even accounting fraud (Larcker &

---

[7] https://www.eater.com/2021/10/28/22750957/sweetgreen-ipo-profit-losses-salad-chain-as-a-tech-company
[8] https://medium.com/sweetgreen/why-a-salad-company-has-a-tech-team-71c131f9aad2

Zakolyukina, 2012). Taken together, this body of research suggests that there is tangible information to be found in the qualitative portion if the S-1 document that can have a material effect on firm performance. Nevertheless, there is still room for additional research in this area. While some researchers have analyzed language usage in qualitative portions of the S-1, my review did not identify any studies focused on the *Business* section. This S-1 section provides unique insights into management's analysis of a firm's business operations, and as such, is worth studying in more detail. The research on corporate filing language usage is also in need of updating. Most research analyzing corporate filing language uses dated dictionaries from the 2000's which have not been updated to account for the increased use of technology focused language by companies going public today. To fill this gap in the literature, this research project includes a content analysis of sub-sample of failed IPO firms to create a new set of data dictionaries to identify the language associated with IPO failures. I then complement this qualitative analysis with a quantitative analysis that uses logistic regression models to assess the significance of the relationship between IPO failures and this language.

The remainder of this paper is organized as follows: First, I summarize the findings of previous research that has studied the effects of obfuscation and word choice in corporate filings. I then discuss how I used qualitative methods (content analysis) to build upon these findings and create a comprehensive list of the language used in the S-1 filings of "failed" IPO companies. The next section of the paper explains how I used this language to then analyze whether there is a correlation between the language of failed IPOs and subsequent IPO failure among a larger sample of firms that issued equity for the first time between 2019 and 2021. Finally, I conclude by discussing the results and implications of these findings and providing directions for future research.

LITERATURE REVIEW

In the past, scholars have taken an interest in studying the factors that seem to predict the success or failure of an IPO and the role of linguistics in corporate filings more generally as predictors of firm performance. This literature review aims to summarize the existing literature on these two areas of focus. It will also point out the gaps in the current literature and provide a further explanation as to why the study of language usage in the *Business* section of the S-1 filing specifically will provide an updated view to current scholarly works. Lastly, this review will use theory to develop a pair of hypotheses that will guide my inquiry into this understudied area.

Linguistics Studies of Corporate Disclosures

Scholars in the past have noted the importance of tone and word choice employed by firm management in a company's corporate filings and earnings calls. While linguistics and accounting may only appear to be tangential areas of research at first glance, many researchers have been able to prove that word choice and the tone of corporate filings matter and, in some cases, provide significant insights into management sentiment and future share price performance. For example, Benson et al. (2015) used a variety of textual analysis tools to measure the complexity, tone, and readability of a firm's initial filing documents. These metrics were combined to create a new metric to measure the overall level of camouflage used as a single variable. After a review of 1,665 corporate charters and bylaws for IPOs conducted between 1995 and 2011, these scholars found that firms who used more "camouflage" in their filing documents also had less investor-friendly governance structures. Interestingly, this research also found that firms with more camouflage language suffered less underpricing at IPO and were able to raise more capital than their peers who opted against camouflage language. This

finding suggests that even sophisticated investors can be fooled by management use of opaque disclosures and explanations of a company's performance.

Adding to the literature on readability and obfuscation, Courtis (2004) used readability scores to determine obfuscation used by publicly traded firms on the Hong Kong Stock exchange. This paper added to the existing literature around obfuscation by defining it in a slightly new way. Their new approach defined obfuscation as the simultaneous use of writing with low ease and high readability variability. The paper also provides interesting explanations as to why management would use obfuscating language in its corporate filings. The researcher posits that managers endorse obfuscated writing to lessen investor anxiety about organizational changes. They also suggest that management teams use obfuscation language to misrepresent or conceal unpleasant facts about a firm's financial health.

Adding to the previously discussed literature, Bushee et al. (2018) also explored the topic of linguistic complexity in firm disclosures. Bushee noted that researchers generally perceive complex language in corporate filings as a trailing indicator of managerial obfuscation. Rather than take this assumption at face value, Bushee et al. chose to categorize complex language into two components: obfuscation and information. This approach allowed Bushee and his team to separate complex language used for informative technical disclosures from the more general obfuscation language that is often perceived as a sign of deliberate management obfuscation. The scholars successfully developed an empirical approach to estimate the two components of complex language in quarterly earnings conference calls and found that the information component of complex language is negatively associated with information asymmetry. Conversely, the group found that conference calls with more of the obfuscation component were positively associated with information asymmetry. These findings suggest that not all complex

language is created equally and that the use of complex language when it is not warranted can be an indicator of deliberate managerial obfuscation.

Scholars have explored the information value of the use of deceptive language in other corporate disclosure areas, as well. For instance, Larcker and Zakolyukina (2012) used existing psychological and linguistic research to classify earnings calls as being either truthful or deceptive and created a predictive model that performed 6-16% better than random at guessing which companies subsequently restated their financial statements. Specifically, they found that the language of deceptive executive included more references to general knowledge, fewer non-extreme emotions, and fewer references to shareholder value. Perhaps most notably, the group created a portfolio of firms with the highest deception scores which produces an annualized alpha of between -4% and -11%. This groundbreaking paper illustrates the predictive value of language in corporate disclosures.

While the aforementioned linguistic studies focus on how the use of deceptive, complex, and obfuscating language in corporate disclosures can be used to predict firm failures, it is important to note that many scholars have also explored how positive word choices can be used to predict success, as well. Scholars have been studying the relationship between corporate ethics and financial performance since at least the 1970's. A recent meta-analysis of this stream of research (Friede, Busch, & Bassen, 2015) shows that across 2,200 different studies, 90 percent find a significant positive relationship between environmental, social, and governance (ESG) efforts and financial performance. Several factors may explain why firms that "do good" also seem to "do well." Some research shows that customers are more loyal to companies with an ESG focus (Bhattacharya & Sen, 2004; Flammer, 2015). Employee quality, loyalty, productivity, and worker satisfaction appear to be higher at socially responsible companies (Flammer, 2015;

Roberts & Dowling, 2002; Turban & Greening, 1997). Scholars have also shown that socially responsible firms are less likely to engage in earnings management or to be the subject of a Securities Exchange Commission investigation (Kim, Park, & Wier, 2012). An ESG focus may even enhance operational processes in ways that ultimately translate into improved financial performance. Orlitzky et al. (2003) in their meta-analysis claim that "through CSP (Corporate Social Performance) processes, firms develop competencies in scanning the external environment and dealing with external changes, turbulence and crises, and develop a forward-thinking managerial style" (p. 407).

Especially relevant to the current study, socially responsible firms also receive more optimistic analyst recommendations (Loannou & Serafein, 2010), attract a larger analyst following, have a lower level of analyst forecast error (Dhaliwal et al., 2011, 2012), attract more institutional investors, and have greater liquidity (Heinkel et al., 2001; Hong & Kacperczyk, 2009). Consistent with this evidence, better ESG performance is shown to be associated with a lower cost of equity (Dhaliwal et al., 2011; El Ghoul et al., 2010; Sharfman & Fernando, 2008), a lower cost of debt (Bauer & Hann, 2010), and higher credit ratings (Bauer & Hann, 2010; Standard & Poor's Governance Services, 2004).

Importantly, not all firms may experience superior financial performance simply by adopting an ESG focus or relying on ethics-based language. Firms whose value proposition is dependent on intangible resources (innovation, human capital, reputation, and culture) are more likely to benefit from following a socially responsible strategy (Surroca, Tribó, & Waddock, 2010). Moreover, ESG efforts cannot be empty signals without concrete action, what scholars variously call the trap of "greenwashing" or "all talk, no walk" (Hawn & Ioannou, 2016).

Other Studies of IPO Performance

In addition to the wide array of academic literature describing the relationships between linguistic traits and firm performance more generally, the factors underlying the relative success (or failure) of IPO firms as a specific category have been studied, as well. Starting at the broadest level, Bhattacharya et al. (2010) tried to determine if fundamental accounting variables are informative in predicting IPO returns. In their study, the group examined the S-1 filings of IPOs specifically in the technology sector. This sector was identified as one of particular interest following the crash of the internet-led tech bubble of the late 1990's. During this period, some investors seemed willing to ignore fundamental accounting variables for the fear of missing out on the next big competitor in the technology space. Additionally, these scholars wanted to explore criticisms raised by some that accounting metrics failed to capture the financial health of technology companies and therefore played a major role on fueling the tech bubble itself. After analyzing the IPO filings and included financial data for IPO firms between '82 and '00, Bhattacharya et al. found that fundamental accounting variables offered significant explanatory power in predicting technology IPO failures. These finding seem to suggest that even in times of investor exuberance and when evaluating emerging industries, the traditional financial reporting system still offers significant insights into the financial health of firms.

Bhabra and Pettway (2003) also chose to focus on the financial metrics of IPO firms and tried to determine if these metrics offered any insights into subsequent performance. Specifically, these scholars researched if there are financial or operating characteristics of IPO firms that have predictive value in determining IPO success or failure. Bhabra et al. analyzed this question using an industry agnostic sample of 242 IPOs between '87 and '91. Staying consistent with the previously discussed study, this study found that traditional financial metrics like pre-IPO profitability and R&D expense both offered significant value in predicting a firm's subsequent

survival or failure. Moving away from traditional financial metrics, the group also determined

that a firm's IPO offer size, firm size, and the number of risk factors listed in the prospectus also

offered significant predictive value in determining a firm's survival rate. This study adds

significantly to the existing literature and suggests that traditional financial metrics in tandem

with a thorough analysis of the S-1 content can predict subsequent IPO success. This study also

set the stage for future in-depth analyses of the S-1's content beyond financial accounting

metrics.

Other existing literature around IPOs considers the informational content of the S-1 filing

rather than focusing on financial metrics. In fact, many of these studies draw on the existing

linguistics literature as the backbone of their analysis. The first of these articles was written by

Arnold et al. (2010) and aimed to explore the effects of ambiguous information in companies'

initial filing on subsequent returns. This research largely revolved around the risk disclosure

section of the S-1. As discussed previously, traditional financial metrics have proven in multiple

cases to provide predictive value in determining subsequent IPO returns. The study done by

Arnold et al. took a new approach to IPO research by trying to determine if "soft" or ambiguous

information could also have predictive value. Guided by existing research on ambiguity theory,

Arnold et al. predicted that firms with higher concentrations of ambiguous language in their S-1

would yield lower returns as investors altered their portfolio weights to account for this

additional risk. They measured the ambiguity of the *Risk Factors* sections of 1398 IPOs between

'98-'05 following methods used by Goetzmann et al. (2007). In line with the group's hypothesis,

Arnold et al. found that firms with greater ambiguity (i.e., soft information) in their prospectuses

yielded lower returns than their unambiguous counterparts. This article adds significantly to the

existing literature for two reasons. First, it suggests that investors place significant value on the

ambiguous informational content of the S-1 filing and at least consider these disclosures as they make investment decisions. Second, it suggests that the market is doing an effective job of holding companies that fill their S-1s with information-light content accountable.

Later, Wyatt (2014) added to the informational content focused literature on IPO filings. Wyatt's research focused specifically on the *Use of Proceeds* section of the S-1 and aimed to determine if this section could be helpful in predicting IPO underpricing and subsequent firm survival rates. This study used content analysis tools to decompose the section into 3 categories of disclosure: growth investments, production investments, and financing activities. This data set included 241 Australian IPOs between '94 and '00 and related the three previously discussed categories with subsequent IPO performance. The study found that firms with *Use of Proceeds* sections geared towards financing activities saw slightly more underpricing and performed worse on firms' survival metrics. This research adds to the existing evidence that content analysis can be useful in predicting negative IPO outcomes. Additionally, it suggests that investors are legitimately considering the contents of specific sections of the S-1 as they make investment decisions.

Hanley and Hoberg (2010) also explored the informational content of the S-1 filing. Rather than focusing on IPO returns after their launch date, their research revolved around the pre-market pricing success or failure before IPO launch. Their research was guided by a real-world question that every IPO firm will face: should firms expend resources during the pre-market period to produce a highly informative S-1, or should they disclose the bare minimum and leave the rest for investors to figure out on their own? To analyze this question, Hanley and Hoberg used content analysis to decompose various S-1 sections into "standard" and "informative" components and compared this to pre-market pricing data. The group subsequently

found that firms with a greater degree of "informative" content suffered less underpricing during the pre-market period. This finding suggests that allocating resources towards a highly informative S-1 filings can lead to more successful IPO launches. Additionally, Hanley and Hoberg findings suggest that firms who opt for more boilerplate, less informative language can suffer negative financial consequences.

Further adding to the evidence that ambiguous language can negatively affect the IPO launch process, Park and Patel (2015) conducted very similar research on IPO pricing during the first trading day. Basing their research in signaling theory, they hypothesized that IPO underpricing would be low when the prospectus contained less ambiguous information and therefore sent a more reliable signal of the firm's quality to market participants. The group observed data from 398 IPOs between '98 and '07 and found that firms with less ambiguity (i.e., higher information clarity) did, in fact, suffer from underpricing at a significantly lower rate. While these findings suggest that firm can reduce underpricing by presenting clear information about the firm's quality, it does not appear to be a perfect metric. Park and Patel noted that the positive association between ambiguity and IPO underpricing was less pronounced when IPO firms displayed low strategic conformity with peers, operated in industries with high valuation heterogeneity, or were medium-sized. This caveat to the study's findings suggests that ambiguity alone has limited predictive value for IPO underpricing and that other measures should be explored in future research.

In one final exploration of the use of ambiguous language, Loughran and McDonald (2013) researched the effects of negative and uncertain words on the return volatility of IPO firms. Basing their research in prospect theory, these scholars anticipated that IPO firms opted to include more negative and uncertainty words would suffer greater volatility than their peers.

Their data included 1,187 completed IPOs as well as 793 withdrawn IPOs between '97 and '10. Consistent with their hypothesis, the study found that large amounts of uncertainty text led to more valuation uncertainty as measured by price revisions and return volatility. This article adds further evidence to the idea the use of uncertain language can have significant value in predicting unstable IPO launches.

HYPOTHESES DEVELOPMENT

In sum, the existing literature provides several insights that are relevant to this study. First, both IPO and linguistic based studies suggest that market participants do rely on the informational content of the S-1 filing when making portfolio allocation decisions. Additionally, the existing literature suggests that both quantitative financial metrics and qualitative measures derived from initial filing documents can hold predictive value in determining future firm performance. Second, across nearly all the literature reviewed, it is clear that investors respond negatively to ambiguous, obfuscated, and deceptive language. As a result, firms who employee these tactics appear to suffer from negative financial outcomes like poor share price performance and even complete firm failure at a significantly higher rates than their peers. In contrast, firms who take an ethics-based or ESG focus to their business operations can be rewarded with more optimistic analyst ratings, lower cost of equity, and decreased litigation risk. Firms that make a good-faith effort to adopt an ethical approach are viewed as less risky by public markets and benefit financially across several quantitative metrics.

Across the existing literature on linguistics and its relevance in predicting future firm performance, one theme stands out amongst the rest; increased use of obfuscating or ambiguous language leads to decreased investor confidence and serves as an effective predictor for poor firm performance in the long run. Based on some of the trends discussed in this paper's

14

introduction section, I believe that a new variety of obfuscation language, especially the use of technology jargon, has appeared in recent years and that studying this language may offer additional predictive value in determining firm performance post-IPO. Stated formally:

- **Hypothesis 1**: The use of obfuscating language will be associated with a greater likelihood of IPO failure.

The literature also indicates that firms can expect some upside from adopting an ethic-based approach to conducting business. As investors tend to view these firms as less risky stewards of investor capital, these firms suffer firm failure as measured across a variety of metrics at lower rates than their non-ethics-focused peers. It is worth noting some of the literature suggests that firms may be espousing an ESG focus without conducting their business operations in an ethics-based manner (i.e., "greenwashing"). While this concern is worth addressing and will be discussed further, the general tone of the literature is that firms can expect a variety of positive firm outcomes by adopting ethic-based language in their corporate disclosures. Stated formally:

- **Hypothesis 2**: The use of ethics language will be associated with a reduced likelihood of IPO failure.

While the existing literature does provide a few important takeaways as discussed above, there is a need for additional research on this topic. Almost all the existing literature seems to address well researched topics like obfuscation, linguistic complexity, and readability. Some researchers have taken their studies deeper and used content analysis to analyze specific sections of the S-1 filing document. The current study contributes to this literature as it will attempt to address the effects of one segment of ambiguous language that has not yet been assessed: buzzword usage. Furthermore, most of the existing literature relies on company data from the

late '90s through the late '00s. Therefore, many important financial periods including the '08

financial crisis and the meteoric rise of technology companies during the dot-com bubble are not

accounted for in this literature. This paper will help update the existing literature significantly

and it could help determine if market sentiment towards ambiguous language has shifted in

recent years in any meaning full way. By identifying and quantifying the newly emerged types of

obfuscation language, the research can help determine if this new variety of obfuscation

language has predictive value in determining future firm performance.

METHODS

Sample

The first step in the research process was to locate a data set that contained all the IPOs

that occurred between 2019 and 2021. This data was exported from an online data provider

(stockanalysis.com) and contained relevant data on the IPOs that occurred that year including

IPO Date, Company Name, IPO Price, and both Trading Price and Real Return as of September

29, 2021. Using the stock information feature in Microsoft Excel, industry descriptions were

added for each IPO observation. From here, all IPOs with the industry of "Holdings Company"

were removed from the data set. This step was taken to ensure special purpose acquisition

companies (SPACs) were not included in the data set, because their initial filing documents do

not contain sufficient explanations about the business the SPAC is attempting to acquire.

Additionally, following the practice of other studies of IPOs (Bhattacharya et al. 2010; Loughran

and McDonald 2013), any observation with an IPO price of less than $5 was removed. The

master IPO data set was then supplemented with data from COMPUSTAT. Data retrieved from

COMPUSTAT included a binary indicator if the shares were still publicly traded, country codes,

GICS data, along with various financial metrics.

After this initial data set consolidation, 705 observations remained. Lastly a few additional categories of observations were removed from the data set. 163 of the observations were removed because the ticker could not be found in the COMPUSTAT database. Then 136 observations were removed because the firm was a non-US filer. These foreign based firms were excluded because they file an F-1 rather than the S-1 as their initial filing document. 11 observations were also removed because COMPUSTAT identified them as the target of either a merger or reverse acquisition since their initial IPO. 19 observations were also removed because the S-1 filing for the observation could not be located. All of these changes resulted in a final sample size of 378 IPOs for analysis.

Defining IPO Failures

I measured IPO failure in two ways. The first metric for failure was if the firm was in the lowest 5% of returns out of all of the observations (i.e., returns of less than -74.4%). This type of IPO failure is referred to as "IPO Return Failure" throughout this paper. The second metric for failure was if there had been a class action litigation event recorded against the firm. Data for securities class action litigation was pulled from Stanford Law Schools Securities Class Action Clearinghouse database. This type of IPO failure is referred to as "IPO Litigation Failure".

Preliminary Identification of the Language of IPO Failure

To identify the types of language that might predict IPO failure, I used a grounded theory approach (Glaser and Strauss 1967). Specifically, I conducted a content analysis of eight S-1 business sections for IPOs that had low (negative) real returns. In theory, these firms should provide insights into the current language and buzzwords used by failed IPO firms. Firms were chosen for this sub-sample of IPOs due to their particularly poor stock performance across a variety of industries. The IPOs included in the content analysis included: Casper (CSPR),

Greenlane Holdings (GNLN), Home Point Capital (HMPT), Progenity (PROG), Root Insurance (ROOT), SmileDirectClub (SDC), ContextLogic (WISH), and ON24 (ONTF).

The content analysis involved reading and coding the business sections of the S-1s of the firms in this sub-sample line-by-line to identify common language and phrases, comparing these common terms to identify second-order categories, and finally, grouping these categories into highest-order dimensions or themes. The lists of terms associated with each second-order category then became a unique term dictionary, which I could use to assess and score the frequency of language use across each category for each company in our larger sample. Table 1 provides a summary of these terms organized by category and theme.

-- INSERT TABLE 1 HERE --

Step 1: Reading and Coding. The S-1 filing document for each of the eight observations was retrieved from their respective investor relations website. All the S-1 documents in PDF format were uploaded to Perusall, an online based document mark-up tool. The business section for each S-1 was reviewed by 2 coders working independently to identify any word or phrase that met one of more of the following criteria: (1)The word or phrase seems surprising or unusual in the context of a formal corporate filing; (2)The word or phrase occurred frequently (three or more times) across the S-1 sample set; or (3) The word or phrase was involved in metaphor or imagery in the document or contributed significantly to the tone or emotion of the document.

The two coders used Perusall's "hashtag" feature to record these key words. The Perusall tool allowed us to identify words and phrases of interest in a collaborative manor and exchange memos about potential emerging themes. We also returned to the literature to consider terms that would be suggested by previous theory. Additionally, the two coders meet on a weekly basis to discuss emerging themes and to resolve discrepancies that arose during the codification process.

It was during these conversations that we explored the possibility of expanding our key words to also include terms that were noticeably absent from the S-1 filings of our sub-set of failed IPOs but had been identified by the literature as theoretically significant. Specifically, we brainstormed terms related to ethics, ESG, and a stakeholder-orientation to add to our coding schema.

Steps 2 and 3: Categorizing and Grouping into Themes. To facilitate the process of identifying second-order categories for each of the words and phrases identified in Step 1, we recorded each word or phrase of interest on notecards and sorted them by grouping words with similar use cases and meanings to identify categories. We then evaluated these categories to group them into three themes which emerged from reviewing the sample set of S-1s.

Innovation Language. The first theme that emerged from the coding analysis of failed IPOs we labeled "innovation language". Included within this theme were the categories of tech jargon and entrepreneurial. Some examples of each type of language include: ON24 (NYSE: ONTF): *"By leveraging **artificial intelligence / machine learning**, our platform enables businesses to use this data to derive highly relevant and deep insights."* (p. 98). SmileDirectClub (NASDAQ: SDC): *"Through our **cutting-edge** teledentistry technology and vertically integrated model, we are revolutionizing the oral care industry."* (p. 99)

There are few reasons why failed firms would have relied on innovation language in their S-1's business section. First, firms could be attempting to signal their membership to the innovation focused technology sector. Since the technology sector has outperformed the broader market greatly over the past two decades, it is not surprising that newly listed firms would wish to capitalize on the industry wide tailwinds that have benefited the technology sector (S&P Global 2022).

Second, failure firms could be employing innovation language to signal their innovative and entrepreneurial nature to investors and distract attention from their perhaps weaker fundamentals. By filling their business sections with innovation focused language, these firms are signaling their ability to create new and cutting-edge products which should generate outsized shareholder returns in the long run even while these firms are incurring substantial losses along the way.

Opportunity Language. The second theme that emerged from the coding analysis of failed IPOs we labeled "opportunity language". The categories included in this theme were extreme optimism, growth language, strategy jargon, and financial terminology. Failed IPO firms may have relied on opportunity language for a number of reasons. By using terms, we classified as extreme optimism (e.g., "massive", "phenomenal") these firms may have been trying to build hype and generate emotional sentiment around their products and services. Firms may have been employing growth language to signal not only the potential size of the firm but also the potential of the market they are competing in. IPO firms may have been using strategy jargon to create the illusion of strong planning, strategic opportunities, and the like. Strategy jargon could also signal a firm's ability to leverage their core competencies into tangible returns for shareholders. Lastly, firms may have employed financial language to indicate the potential for profitability in the future while distracting from the firm's relative weaknesses in the present. Some examples of this language include: ContextLogic (NASDAQ: WISH): *"Our data science allows personalization at the individual user level at a **massive** scale"* (p. 112); ON24 (NYSE: ONTF): *"We estimate that the current **total addressable market** for our solutions is $42 billion worldwide annually,"* (p. 100); ContextLogic (NASDAQ: WISH): *"Our sales and marketing*

20

*represent a **core competency** that is essential to the success of the Wish platform,"* (p. 157);
Home Point Capital (NASDAQ: HMPT): *"We have also grown our **non-GAAP** core operating metrics for the same periods,"* (p. 114).

Stewardship Language. The third theme that emerged during the coding analysis of failed IPO firms we labeled "stewardship language". This theme included the categories of ethics language, operational language, and customer focused language. Unlike the previous two themes, we grouped these categories together because we noticed the absence of these terms among our failed companies rather than their presence. In other words, these categories and themes were inspired from the literature and our theoretical motivations rather than our coding.

Testing the Language of IPO Failure

The second phase of our data analysis was to assess how well the language categories identified in the content analysis phase described above predicted IPO failure across a broader population. To test whether increased use of innovation, opportunity, or stewardship language was associated with IPO Return Failure (H1), we ran the following logistic regression:

$$
\begin{aligned}
IPO\ RETURN\ FAILURE = \ &\beta_0 + \beta_1 INNOVATION\ LANGUAGE + \\
&\beta_2 OPPORTUNITY\ LANGUAGE - \\
&\beta_3 STEWARDSHIP\ LANGUAGE + \beta_4 WORD\ COUNT + \\
&\beta_5 READABILITY + \beta_6 TOTAL\ ASSETS + \beta_7 NET\ INCOME + \\
&\beta_8 TRADINGDAYS + \beta_9 IPO\ VALUE + Year\ Effects + \\
&Industry\ Effects + \ \varepsilon
\end{aligned}
$$

We include several control variables to account for the external factors that could contribute to a firm's membership in the IPO Failure category. We control for the length of the business section of the S-1 (*WORDCOUNT*) and the readability of an IPO firm's business sections (as measured by the Flesch Reading Ease score, *READABILITY*), because prior research suggests that firms with more obscure and complex language fair worse in public markets than

their peers (Bushee, Gow, and Taylor 2018). We also control for firm size (*TOTALASSETS*) and the size of the public offering (*IPOVALUE*) to account for the effects of firm size may have on membership in the failure category. We control for a firm's profitability (*NETINCOME*) as highly unprofitable companies should belong to the failure category at higher rates than their profitable peers. We control for a firm's exposure to the public markets (*TRADINGDAYS*) to adjust for the time each observation firm has been exposed to public markets. This exposure ranges from a low of 105 days to a high of 1008 days. We control for the year each firm went public (*Year Effects*) to adjust for the range of market and macroeconomic conditions seen between 2019 and 2021 and to adjusted for the differing market conditions seen before and after the COVID-19 pandemic. Lastly, we include indicators for industry to control for the highly varied performance seen across industries during the observed period.

To test for the effects of business section language on a firm's likelihood to be the subject of class action litigation (H2), we ran the following logistic regression:

$$
\begin{aligned}
IPO\ LITIGATION\ FAILURE = \ &\beta_0 + \beta_1 INNOVATION\ LANGUAGE + \\
&\beta_2 OPPORTUNITY\ LANGUAGE - \\
&\beta_3 STEWARDSHIP\ LANGUAGE + \beta_4 WORD\ COUNT + \\
&\beta_5 READABILITY + \beta_6 TOTAL\ ASSETS + \beta_7 NET\ INCOME + \\
&\beta_8 TRADINGDAYS + \beta_9 IPO\ VALUE + Year\ Effects + \\
&Industry\ Effects + \varepsilon
\end{aligned}
$$

Like the previous model, we include control variables for business section length (*WORDCOUNT*), ease of readability (*READABILITY*), firm size (*TOTALASSETS*), profitability (*NETINCOME*), and offering size (*IPO VALUE*) to account for the affects these variables may have on failure. Exposure to public markets (*TRADING DAYS*) was included to control for the increased likelihood of class action suits that stems from being public for a longer period. Dummy variables for the year of IPO (Year Effects) and industry membership (Industry Effect)

were included to control for changing public market performance across the observed time periods and industries.

RESULTS

Descriptive Statistics and Bivariate Correlations

Table 2 provides the descriptive statistics of the sample. We note that on average, real returns for the sample were 52%, but the range of returns was quite broad with a low of 93% losses on one end (Progenity, NASDAQ: PROG) and a high of 863% gains on the other (Silvergate Capital, NYSE: SI). Among the nine different categories of language identified through our content analysis, Tech Jargon has the greatest presence in the sample with an average word count from this list at just over 126. On the other hand, Extreme optimism was the least prevalent in the sample with an average word count from this list at only six.

-- INSERT TABLE 2 HERE --

Table 3 shows the inter-correlations of study variables. As predicted, the use of one type of Stewardship Language - ethical terms - has a significant, negative correlation with our first measure of IPO failure (IPO Return Failure). However, we do not see any significant correlations between our other language measures or our control measures and IPO Return Failure.

Regarding our second measure of IPO failure (IPO Litigation Failure), we see that two of our Stewardship Language categories - the use of operational language and of customer-focused language - both have significant positive correlations with shareholder class-action litigation events, which is in the opposite direction of our predictions. However, we do not see any significant correlations between Innovation Language or Opportunity Language and IPO Litigation Failure. All of our control measures have significant correlations with IPO Litigation Failure. We also note that several of our control measures are highly correlated with one another.

We assessed the risk of multi-collinearity in our models by running a Variance Inflation Factor (VIF) test. This test identified that our industry controls and our measure for Trading Days had high auto-correlation (scores above 10) and the average VIF was well above 1. Therefore, we ran our models again without these measures to see if our results were due to multi-collinearity and noted the significance of the results presented below were unchanged.

-- INSERT TABLE 3 HERE --

Model 1 Findings

Columns 1-3 of Table 4 present our findings on the relationship between the use of language in the S-1 filing and our first measure of IPO failure, IPO Return Failure (that is, real returns falling in the lowest 5% of the distribution of the sample). Overall, the model fit is strong, with an overall pseudo R^2 of 19.0%. Contrary to our hypothesis 1, however, increased use of Innovation Language (tech jargon and entrepreneurship language) did not have a statistically significant effect on IPO Return Failure. Likewise, hypothesis 2 that Opportunity Language would be positively correlated with IPO failure is not supported. On the other hand, our hypothesis that the usage of Stewardship Language in the S-1 business section would reduce the risk of IPO Return Failure (H2) is supported. Specifically, there is a significant negative relationship between increased usage of ethics language in the business section and IPO Return Failure (*Beta* = -0.101, *p-value* = .043). To interpret the ethics coefficient on IPO Return Failure, I estimated our model without the logarithmic transformation. The model suggests that a one-unit increase in ethics language contributes to -9.6% lower odds of IPO Return Failure, holding all other variables constant. While pre-pandemic vs. post-pandemic year controls did not have a significant effect on IPO Return Failure, some industries did have significantly higher rates of IPO Return Failure. These industries include Consumer Staples and Financials.

Model 2 Findings

Columns 4-6 of Table 4 present our findings on the relationship between the use of language in the S-1 filing and our second measure of IPO failure, subsequent class action litigation (IPO Litigation Failure). Overall, the model fit is strong, with an overall pseudo R^2 of 26.2%. Contrary to our predictions, Innovation Language did not show significance when controlling for other factors that might predict IPO Litigation Failure. However, two of our measures of Opportunity Language - growth language (*Beta* = -0.039, *p-value* = 0.007) and strategy jargon (*Beta* = 0.017, *p-value* = 0.031) have significant effects on a firm's likelihood of class action litigation. Additionally, one of our measures of Stewardship Language – customer language (*Beta* = 0.016, *p-value* = 0.021) had a significant effect on firms' likelihood of class action litigation. We note that our prediction was that increased use of Stewardship Language would have a negative rather than a positive effect on the likelihood of IPO failure, so we explore these surprising findings in more detail in the following section of the paper.

Interpreting these coefficients without the logarithmic transformation can provide more readable outputs. For the Opportunity Language theme, a one unit increase in growth language yields a 3.83% decrease in the odds of class action litigation, and a one unit increase in strategy jargon yields a 1.74% increase in the odds of class action litigation. On the Stewardship Language theme, a one unit increase in customer language results in a 1.63% increase in the odds of class action litigation. Year controls were not significant predictors of IPO litigation Failure but industry controls, in particular the Materials industry, was significant.

DISCUSSION & CONCLUSION

25

This study considers the language used in the initial filing documents for firms issuing their initial public offerings (IPO) and the effects of this language on subsequent IPO failure as measured by total returns and likelihood of class action litigation. While other scholars have explored the effects of corporate filing language on a variety of firm outcomes, this study is unique in two ways. First, I explore the effects of the language used in the *Business* section of initial filing document on subsequent IPO performance, a disclosure area that has not yet been explored according to my review of the literature. This study is also unique in that it considers IPO failure not only in terms of real returns, which has been studied previously, but also in terms of class action litigation. By including litigation as a form of failure, I can access information about how shareholders felt mislead in the IPO process and unpack the role that language played in this process. Developing a more thorough understanding about these relationships could help investors avoid IPO firms which are more likely to suffer poor market performance or high litigation risk in the future.

To examine the relationships among innovation language, opportunity language, and stewardship language and the resulting IPO outcomes they might influence, I began by exploring the existing literature on the use of language in corporate reporting. Following a thorough review of the existing literature, I identified that ambiguous or obfuscated language was associated with fraud and poor public market performance. On the other hand, stewardship language, especially ethics and a stakeholder-focus, was frequently associated with positive performance. My content analysis of a sub-set of failed IPOs also suggested that the terminology I identified as "innovation language" and "opportunity language" might capture the new ways that companies were attempting to obfuscate and mislead investors. I hypothesized that the increased use of these terms would predict poor IPO returns and/or increased risk of litigation. Likewise, I

compiled a list of "stewardship language" terms that might signal a firm's honesty and commitment to ethics. I hypothesized that the increased use of these terms would have a negative correlation with poor IPO returns and/or increased risk of litigation.

Despite the clear story from the existing literature on the effects of language on firm performance, the findings from this study are somewhat mixed. Innovation language, which included my data dictionaries for tech jargon and entrepreneurship, and the control measure for obfuscation (*READABILITY*), did not serve as an effective predictor for either of the measures of IPO failure. This non-finding could be explained by a few factors. One potential reason is that not all Innovation Language is created equal. It is possible that some Innovation Language is being use legitimately by managers to explain the operational detail of their business, while other portions of Innovation language are being using for a more sinister purposes as originally hypothesized. Additional research would be needed to tease out the informative vs. the uninformative portions of Innovation Language. Another reason for this finding could be that we have entered an era were Innovation Language offers no material information to investors regardless of industry type. It is possible that the rapid adoption of technology across all industry types has made innovation language a common and a legitimate way for managers to explain their business operations to the public rather than the obfuscation tool that was originally hypothesized. Indeed, Innovation Language had the highest frequency rate of all the data dictionaries in our sample, with firms using an average of 127 terms from the tech jargon dictionary and 51 terms from the entrepreneurship dictionary.

Our findings on the opportunity language theme were somewhat mixed across the two measures of failure. When comparing opportunity language and its relationship to return failure, we found that opportunity language was not an effective predictor of failure as hypothesized. I

expected that managers were using opportunity language to overstate the potential size and profitability of their respective firms with the hope of generating outsized share price performance. The findings, however, suggest that managers are legitimately using opportunity language to convey their firm's revenue growth prospects to potential investors. When comparing opportunity language and its relationship to litigation failure, the findings were more significant. The increased use of strategy jargon led to higher rates of litigation failure. This finding suggests that managers may be using this type of language to obfuscate their business sections and in turn distract from any fundamental issues with their firm's business model.

An example of this theory can be found in the IPO of Rocket Companies (NSYE:RKT). The *Business Section* of this mortgage loan provider features 275 occurrences of strategy jargon, among the highest in the sample. For example, Rocket's claims: *"[their] national Rocket Brand is our **competitive advantage** that is extremely difficult to replicate"*. Following their IPO however, Rocket Companies was the subject of a class action suit which alleged they failed to disclose the price war they were engaged in to boost market share. In other words, they did not have a sustainable competitive advantage but rather were cutting their prices in order to grow the business. I believe that this is an example of Rocket using strategy jargon to create the illusion of strategic advantage rather than an actual customer acquisition tool (i.e., cutting prices and compressing margins).

Surprisingly, the increased use of growth language was associated with a decreased risk of litigation failure. This finding seems to directly contradict the existing literature and suggests that managers are using this category of opportunity language appropriately as they describe the growth prospects of their firm and industry. An example of this phenomenon can be seen in the case of Sonim Technologies (NASDAQ: SONM). This ultra -rugged phone manufacturer used

growth language just 17 times in their *Business Section*, a full standard deviation below the mean. In one of the firms few references to growth language, Sonim states that *"[they] plan to expand the revenue contribution of the subscription-based products and services… potentially increasing sales to end customers"*. Following their IPO, Sonim was the subject of a class action suit where complainants claimed that Sonim failed to disclose adverse facts about its business operations and prospects. A potential explanation for this tempered use of growth language is that managements' knowledge of the firm's technology shortcomings caused them to be more cautious in using growth language to describe their prospects.

The findings on the significance of stewardship language on IPO failure in this study also raise interesting questions. As the literature would suggest, my findings show that the increased use of ethics language is associated with higher IPO returns (a reduced chance of IPO return failure). These findings suggest that the S-1, and specifically the *Business Section,* can provide material insights into management intentions to steward investor capital effectively. This finding also adds to the existing literature on ethics-focused firms. Adopting an ethics-based approach to business operations may offer downside protection against poor return performance.

An example of increased ethics language producing higher IPO returns can be seen in the IPO of Sprout Social (NASDAQ: SPT). Sprout, which produced social media management tools, used ethics language 72 times in its *Business Section*. Some examples of this frequent usage of ethics language include: *"Sprout Social **empowers** businesses around the globe to tap into the power and opportunity presented by the shift to social communication"* and *"Our success comes from our **diverse** and talented people with varied perspectives who can be their whole selves in an equitable and inclusive environment."* Following its IPO, Sprout shares yielded a 609% total return on its share price, among the highest returns in the sample. The Sprout IPO supports the

idea that firms who adopt an ethics-based approach to business operations can expect more favorable return performance.

The findings on the relationship between stewardship language and litigation failure did not align with my hypothesis, however. Here, the findings suggest that an increase in customer-focused language makes litigation failure more, not less, likely. While the explanation for this finding is somewhat unclear, it is possible that managers are using customer-focused language to oversell their commitment to customers, much like firms that "greenwash" are overselling their commitment to the environment (Cho, Guidry, Hageman & Patten 2012). Investors could be punishing firms for the perceived hypocrisy of "customer-washing" by initiating class-action lawsuits against these firms at a higher rate. An example of this "customer-washing" can be found in the *Business Section* of ContextLogic. The firm, which operates the e-commerce platform wish.com, featured a strikingly high rate of customer language in its *Business Section*. For example, the *Business Section* of ContextLogic claims *"to enhance user engagement, we incorporate fresh gamified features, user-generated content… and a wide range of products to make shopping more entertaining"*. In the class action suit filed against ContextLogic, investors claimed that ContextLogic mislead them about the strength of its business operations and overstated its monthly active user figures and customer acquisition rates. This may be an example of a firm engaging in "customer-washing" to try to convince investors that the firm has a strong customer focus even if the reality is that the firm is struggling in this area.

The are some limitations faced by this study, which are worth exploring and may offer guidance to future research in this field. The first major limitation in this study is the time horizon of the observed IPO firms. There are credible reasons to conduct a study on IPO failure over the past three years. Namely the increasing prevalence of "tech washing" made famous by

WeWork or the rise of ESG investing. However, it is worth noting that many of the observations in this study occurred during the irregular, pandemic influenced market conditions of 2020 and 2021. Although controls for year effects were included in this study, it is possible that these irregular market conditions may have affected the study findings.

Another limitation of this study is that the content analysis done to identify the key terms for study was done in an industry inclusive manner. While the lack of an industry focus resulted in interesting findings in this paper, it may be worthwhile to conduct a similar study on a specific industry segment. This focus would allow for the creation of bespoke language dictionaries for a specific industry rather than the one-size fits all approach taken in this study.

A final limitation of this study to consider is our approach to identifying the language of IPO failure. This study included a thorough analysis of eight IPO failure firms to identify the language associated with IPO failure. It is possible that conducting a content analysis on a greater number of IPO failure firms from a wider range of industries would have led to a more robust language classification scheme. One could also argue that our approach to language identification should have included a set of successful IPO firms to more accurately classify the language that is associated with both IPO failure and IPO success.

Despite these limitations, this study offers important insights into the language of IPO failure. The paper clearly demonstrates that some portions of newly identified obfuscation language can be valuable in predicting IPO firm failure. The paper also adds to the existing literature in suggesting that qualitative sections of corporate filings (like the *Business Section*) can have predictive value for future firm performance. While more research in the area will be needed going forward, this paper can serve as a starting point for future research on obfuscation language and IPO firm failure.

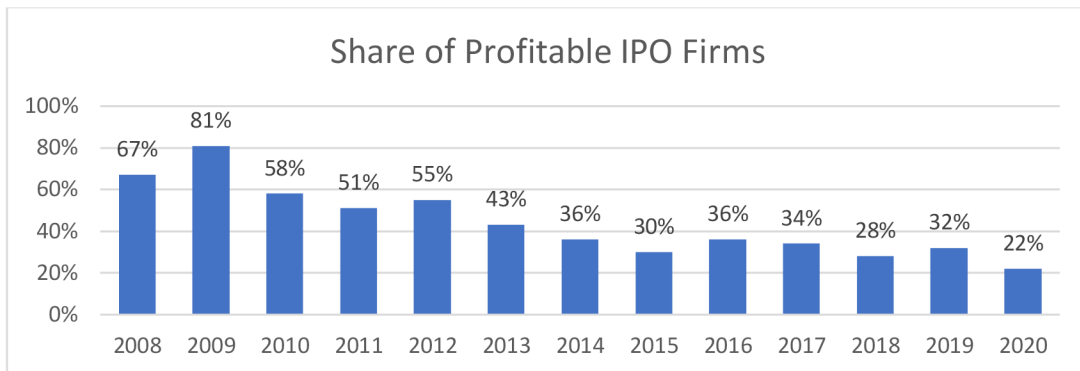**Figure 1.** Share of Profitable IPO Firms (Statista.com 2022)

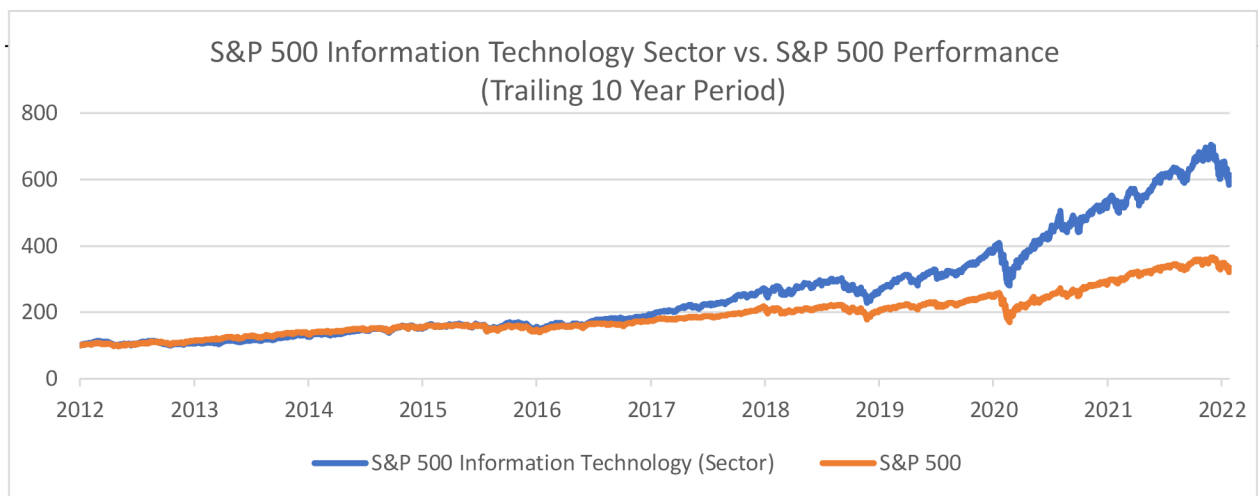**Figure 2.** S&P 500 Information Technology Sector vs. S&P 500 (Trailing 10 Year)

Table 1. Coding Themes, Categories and Words

Themes	Categories	Words
Innovation	Tech Jargon (49 words)	algorithm, algorithms, analyze, analytic, analytical, analytics, API, artificial intelligence, AI, automate, automated, automation, behavioral data, business intelligence, BI, cloud, cloud based, cloud based, data, data points, data driven, data driven, data rich, data rich, database, digital first, ecosystem, ecosystems, friction, frictionless, insight, insights, machine learning, ML, multimedia, network effects, network effects, platform, platforms, portal, social media, software, technology, technological, tech, telematics, dynamic, flywheel
	Entrepreneurial (32 words)	cutting edge, democratize, differentiator, differentiation, disruption, entrepreneur, entrepreneurial, entrepreneurship, human capital, human capital, in house development, in-house development, innovative, innovation, intellectual property, next generation, next gen, niche, pioneer, pioneering, proprietary, research, research and development, research and development, R&D, revolutionary, revolution, revolutionizing, solution, solutions, unique, uniquely
Opportunity	Extreme Optimism (10 words)	aggressive, considerable, dynamic, flywheel, maniacal, massive, massively, robust, sophisticated, tremendous
	Growth (13 words)	addressable market, tam, future, global, growth, opportunity, opportunities, potential, scalable, scalability, scale, tailwind, tailwinds
	Financial (19 words)	billion, billions, EBIT, EBITDA, metric, metrics, non-gaap, run rate, shareholder, shareholders, stockholder, stockholders, trillion, trillions, wallet share, revenue, sales, profit, income, margin
	Strategy Jargon (48 words)	affordable, affordability, best practices, best practices, best practice, best practice, business to business, B2B, business to consumer, B2C, competitive advantage, convenient, convenience, core competency, core competencies, diversification, exclusive, exclusivity, experience, experienced, first mover, focus, focused, high quality, leverage, leveraged, leveraging, premium, strategic, strategy, synergy, synergies, transform, transformative, transformational, value, value proposition, value add, world class, addressable market, TAM, global, opportunity, opportunities, potential, scalable, scalability, scale
Stewardship	Ethics (35 words)	authentic, authenticity, b corporation, b corp, culture, diversity, democratic, empathy, empathetic, empowerment, empowering, enthusiastic, enthusiasm, fair, fairness, inclusive, inclusivity, joy, joyful, passion, passionate, stakeholder, stakeholders, sustainable, sustainability, trust, trustworthy, trusting, wellness, wellbeing, ESG, environment, environmental, social, governance
	Customer (24 words)	brand equity, channel, channels, content, content rich, customer centric, customer first, customizable, customizability, engagement, go to market, targeted, hyper targeted, interactive, interactivity, millennial, millennials, one stop shop, partner, partners, personalization, point of contact, proximity, touchpoint
	Operational (17 words)	agile, agility, efficient, efficiency, flexible, flexibility, logistic, logistics, logistical, nimble, optimize, optimized, redundant, reliable, reliability, resilient, resilience

Table 2: Summary Descriptive Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Dependent Variable					
Real Return	378	0.52	1.32	-0.93	8.63
IPO Return Failure	378	0.05	0.22	0.00	1.00
IPO Litigation Failure	378	0.10	0.29	0.00	1.00
Innovation Language					
Tech Jargon	378	126.82	105.48	0.00	636.00
Entrepreneurship	378	51.06	34.01	0.00	239.00
Opportunity Language					
Extreme Optimism	378	6.00	5.48	0.00	30.00
Growth	378	68.51	36.63	7.00	412.00
Strategy Jargon	378	117.80	51.19	16.00	485.00
Financial	378	41.71	35.83	0.00	458.00
Stewardship Language					
Ethics	378	17.28	19.52	0.00	138.00
Operational	378	14.79	13.49	0.00	107.00
Customer-Focused	378	28.99	30.97	0.00	191.00
Control Measures					
Word Count	378	17922.97	10090.55	3329.00	92886.00
Readability	378	26.84	5.14	15.01	49.49
Total Assets	378	1428.43	3531.88	0.62	37752.83
Net Income	378	-50.34	364.64	-5236.00	1606.35
Trading Days	378	461.80	264.19	105.00	1008.00
IPO Value	378	2869.27	6630.17	0.00	76392.63

Table 3: Pearson Correlation Matrix

Panel 1:

Variables	Real Return	Return Failure	Litigation Failure	Tech Jargon	Entrepreneurial	Extreme Optimism
Dependent Variables						
Real Return	1					
IPO Return Failure	-0.2316*	1				
IPO Litigation Failure	-0.0721	0.2141*	1			
Innovation Language						
Tech Jargon	0.0629	-0.066	0.0269	1		
Entrepreneurship	0.0048	-0.0514	-0.0332	0.4651*	1	
Opportunity Language						
Extreme Optimism	0.0792	-0.0752	0.0842	0.5346*	0.3202*	1
Growth	-0.0563	-0.0932	-0.0678	0.3039*	0.4039*	0.3222*
Strategy Jargon	-0.0857	-0.0678	0.0282	0.3798*	0.4220*	0.3584*
Financial	0.0328	-0.0641	0.0472	0.2604*	0.3016*	0.1697*
Stewardship Language						
Ethics	0.0283	-0.1195*	0.0568	0.1616*	0.0733	0.1810*
Operational	0.062	-0.0395	0.1870*	0.3661*	0.2276*	0.3499*
Customer-Focused	-0.0267	-0.054	0.1519*	0.3319*	0.1954*	0.2559*
Control Measures						
Word Count	-0.1082*	-0.0033	-0.1203*	0.0756	0.2838*	0.1126*
Readability	0.0857	0.0241	0.1998*	-0.0523	-0.3285*	0.0283
Total Assets	-0.013	-0.0646	0.2165*	0.1002	0.0068	0.1569*
Net Income	-0.0207	-0.0006	-0.1344*	-0.1141*	0.051	-0.2202*
Trading Days	0.2729*	0.0785	0.1295*	-0.1500*	-0.0771	-0.0949
IPO Value	0.0275	-0.0472	0.2062*	0.3362*	0.0129	0.3563*

Panel 2:

Variables	Growth	Strategy Jargon	Financial	Ethics	Operational	Customer-Focused
Opportunity Language						
Growth	1					
Strategy Jargon	0.8612*	1				
Financial	0.5419*	0.4766*	1			
Stewardship Language						
Ethics	0.3405*	0.3266*	0.4619*	1		
Operational	0.1982*	0.2791*	0.2528*	0.2311*	1	
Customer-Focused	0.2133*	0.2958*	0.3112*	0.2970*	0.2649*	1

Control Measures

Word Count	0.6789*	0.6358*	0.1079*	-0.0091	-0.085	-0.1104*
Readability	-0.1036*	-0.0827	0.1820*	0.2578*	0.1039*	0.1890*
Total Assets	0.1475*	0.1816*	0.2910*	0.1408*	0.1976*	0.1113*
Net Income	-0.1061*	-0.1085*	-0.0761	-0.0928	-0.1834*	-0.0313
Trading Days	-0.1492*	-0.1718*	-0.0581	-0.1495*	-0.1170*	-0.1919*
IPO Value	0.1243*	0.1697*	0.1586*	0.1797*	0.3196*	0.1652*

Panel 3:

Variables	Word Count	Readability	Total Assets	Net Income	Trading Days	IPO Value
Control Measures						
Word Count	1					
Readability	-0.2611*	1				
Total Assets	-0.1095*	0.1592*	1			
Net Income	-0.0144	-0.2390*	-0.3194*	1		
Trading Days	0.0027	0.0795	0.0043	-0.0606	1	
IPO Value	-0.1175*	0.2563*	0.5189*	-0.5835*	-0.0856	1

Table 4: Models 1 & 2 Logistic Regressions

	Log IPO Return Failure			Log IPO Litigation Failure		
	Prediction	**Coef.**	**P-Value**	**Prediction**	**Coef.**	**P-Value**
Constant		-1.456	0.655		-5.983	0.004
Innovation Language						
Tech Jargon	(+)	0.003	0.556	(+)	0.000	0.994
Entrepreneurial	(+)	0.008	0.563	(+)	0.005	0.559
Opportunity Language						
Extreme Optimism	(+)	-0.069	0.408	(+)	0.027	0.590
Growth	(+)	-0.041	0.104	(+)	-0.039*	0.007
Strategy Jargon	(+)	0.014	0.318	(+)	0.017*	0.031
Financial	(+)	0.002	0.879	(+)	-0.012	0.286
Stewardship Language						
Ethics	(-)	-0.101*	0.043	(-)	-0.014	0.385
Operational	(-)	0.019	0.513	(-)	0.023	0.129
Customer	(-)	0.005	0.750	(-)	0.016*	0.021
Control Measures						
Word Count	?	0.000	0.903	?	0.000	0.297
Readability	(-)	-0.029	0.728	(-)	0.069	0.173
Total Assets	(-)	-0.002	0.074	(-)	0.000	0.208
Net Income	(-)	-0.004	0.080	(-)	0.000	0.816
Trading Days	?	0.000	0.990	(+)	0.004	0.106
IPO Value	(-)	0.000	0.224	(-)	0.000	0.145
Year Effects			ns			ns
Industry Effects			Yes*			Yes*
Number of Obs.			231			371
Overall Pseudo R2			0.190			0.262
Chi-squared			24.93			55.06

* Indicates statistical significance at the <= 5 percent level (two-tailed).
Model 1 Significant Industries: Consumer Staples, Financials
Model 2 Significant Industries: Materials

REFERENCES

Arnold, T., Fishe, R. P., & North, D. (2010). The effects of ambiguous information on initial and subsequent IPO returns. Financial Management, 39(4), 1497-1519.

Benson, D. F., Brau, J. C., Cicon, J., & Ferris, S. P. (2015). Strategically camouflaged corporate governance in IPOs: Entrepreneurial masking and impression management. Journal of Business Venturing, 30(6), 839-864.

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic complexity in firm disclosures: Obfuscation or information?. Journal of Accounting Research, 56(1), 85-121.

Bhattacharya, N., Demers, E., & Joos, P. (2010). The relevance of accounting information in a stock market bubble: evidence from internet IPOs. Journal of Business Finance & Accounting, 37(3-4), 291-321.

Bhabra, H. S., & Pettway, R. H. (2003). IPO prospectus information and subsequent performance. Financial Review, 38(3), 369-397.

Bauer, R., & Hann, D. (2010). Corporate environmental management and credit risk. Available at SSRN 1660470.

Bhattacharya, C. B., & Sen, S. (2004). Doing better at doing good: When, why, and how consumers respond to corporate social initiatives. California management review, 47(1), 9-24.

Cho, C. H., Guidry, R. P., Hageman, A. M., & Patten, D. M. (2012). Do actions speak louder than words? An empirical investigation of corporate environmental reputation. Accounting, Organizations and Society, 37(1), 14-25.

Courtis, J. K. (2004). Corporate report obfuscation: artefact or phenomenon?. The British Accounting Review, 36(3), 291-312.

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. The accounting review, 86(1), 59-100.

Dhaliwal, D. S., Radhakrishnan, S., Tsang, A., & Yang, Y. G. (2012). Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. The accounting review, 87(3), 723-759.

El Ghoul, S., Guedhami, O., Kwok, C. C., & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital?. Journal of Banking & Finance, 35(9), 2388-2406.

Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. Management Science, 61(11), 2549-2568.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment, 5(4), 210-233.

Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2008). The incremental information content of tone change in management discussion and analysis.

Hawn, O., & Ioannou, I. (2016). Mind the gap: The interplay between external and internal actions in the case of corporate social responsibility. Strategic Management Journal.

Heinkel, R., Kraus, A., & Zechner, J. (2001). The effect of green investment on corporate behavior. Journal of financial and quantitative analysis, 36(4), 431-449.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. Journal of financial economics, 93(1), 15-36.

Hanley, K. W., & Hoberg, G. (2010). The information content of IPO prospectuses. The Review of Financial Studies, 23(7), 2821-2864.

Larcker, D. F., & Zakolyukina, A. A. (2012). Detecting deceptive discussions in conference calls. Journal of Accounting Research, 50(2), 495-540.

Kim, Y., Park, M. S., & Wier, B. (2012). Is earnings quality associated with corporate social responsibility?. The Accounting Review, 87(3), 761-796.

Loughran, T., & McDonald, B. (2013). IPO first-day returns, offer price revisions, volatility, and form S-1 language. Journal of Financial Economics, 109(2), 307-326.

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. Organization studies, 24(3), 403-441.

Park, H. D., & Patel, P. C. (2015). How does ambiguity influence IPO underpricing? The role of the signalling environment. Journal of Management Studies, 52(6), 796-818.

Roberts, P. W., & Dowling, G. R. (2002). Corporate reputation and sustained superior financial performance. Strategic management journal, 23(12), 1077-1093.

Sharfman, M. P., & Fernando, C. S. (2008). Environmental risk management and the cost of capital. Strategic management journal, 29(6), 569-592.

Surroca, J., Tribó, J. A., & Waddock, S. (2010). Corporate responsibility and financial performance: The role of intangible resources. Strategic Management Journal, 31(5), 463-490.

Turban, D. B., & Greening, D. W. (1997). Corporate social performance and organizational attractiveness to prospective employees. Academy of management journal, 40(3), 658-672.

Wyatt, A. (2014). Is there useful information in the 'use of proceeds' disclosures in IPO

     prospectuses? Accounting & Finance, 54(2), 625-667.