Is intangibles talk informative about future returns? Evidence from 10-K filings

Alexander David [a], Amir Hosseini [a, *], Anup Srivastava [a]

[a] University of Calgary, Haskayne School of Business, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada

[*] Corresponding author. Tel.:+1 (403) 604-2149 ; E-mail addresses: alexander.david@haskayne.ucalgary.ca (A. David), seyedamirfarhang.hos@ucalgary.ca (A. Hosseini), anup.srivastava@ucalgary.ca (A. Srivastava)

December 2023

Abstract

We construct a measure of intangible intensity — intangibles talk — based on discussions of intangibles in a firm's 10-K filings. This measure likely includes managers' views about the ex-post, successful outcomes of intangibles investments and not just ex-ante investments. In general, the expected and realized outcomes of intangible investments differ because many of those investments fail while only a few would produce lottery-type payoffs. Therefore our measure is correlated with, but carries orthogonal information to, prior measures of intangibles that are largely based on expected outcomes of initial intangibles investments. We test the informativeness of our measure about future returns. Returns from long and short portfolios based on high and low values of our measure, outperform portfolios formed on book-to-market ratio as well as those based on capitalization intangibles. Our strategy generates an average annual alpha of 3.37% from 1995 to 2020. Our alphas are higher than those generated from portfolios sorted on other indicators of intangible intensity shown in the literature. We find that positive alphas are concentrated among stocks with higher arbitrage risk, proxied by idiosyncratic volatility, suggesting that stocks with higher discussion of intangibles in 10-K filings are often mispriced.

JEL classification: E22; G14; O3

Keywords: 10-K filings, Intangibles talk, Intangible intensity, Recognition versus disclosure, Portfolio returns, Idiosyncratic volatility, Limits to arbitrage

1. Introduction

The United States economy has experienced a dramatic shift toward intangible assets in recent decades. Investment in knowledge capital and organizational capacity among US firms has risen steadily (see Corrado, Hulten and Sichel, 2005; Eisfeldt and Papanikolaou, 2013; Enache and Srivastava, 2018; Ewens, Peters and Wang, 2019), allowing them to launch new products and services, or gain a competitive edge in existing marketplaces, through innovation, lower costs, and better customer relations. Because the current US Generally Accepted Accounting Principles (GAAP) requires immediate expensing of internally generated intangibles, they are largely missing from the balance sheet, and no reliable measures of firms' total intangible capital exist. As a result, knowledge firms have more mispriced securities than do firms with physical assets (see Lev and Sougiannis, 1996; Daniel and Titman, 2006; Eisfeldt and Papanikolaou, 2013; Edmans, Li and Zhang, 2014). Meanwhile, the importance of intangibles in the U.S. economy keeps increasing, as each new cohort of public firms spends more on intangibles than its predecessor cohort (Corrado, Hulten and Sichel, 2005; Srivastava, 2014, 2023).

Numerous studies attempt to address investors' problem via estimating the value of internally generated capital by capitalizing and amortizing research and development (R&D) and selling, general, and administrative (SG&A) expenses.[1] Edmans, Li and Zhang (2014) signifies the challenges of accurately defining and gauging intangible value using these methods. We extend this literature by proposing a new, text-based measure that relies on intangibles-related keywords appearing in firms' 10-K filings [following the glossary created by Filipovic and Wager (2019)]. Our method could reflect ex-post, firm-specific successful outcomes of intangibles investments, which cannot be fully captured in the capitalization of ex-ante in-

[1]Those studies using the perpetual inventory model are Hulten and Hao (2008), Peters and Taylor (2017), Lev and Srivastava (2022), Eisfeldt, Kim and Papanikolaou (2020), Iqbal et al. (2023), and Falato et al. (2022)

vestments using uniform capitalization and amortization parameters. We demonstrate the informativeness of our measure through alphas generated by long-short portfolios formed on our measure while controlling for momentum and Fama and French factors.

US GAAP requires that expenditures on internally generated intangibles be immediately expensed. The same GAAP rules permit capitalization of expenditures on property, plant, and equipment (PP&E) and acquired intangibles. As a result, information on in-house-developed intangible assets such as innovation, knowledge, and brand capital are not readily available in financial statements compared with data related to tangible capital (Belo et al., 2022). To address this accounting limitation, numerous studies in the finance and accounting literature have focused on improving measures of, and proposing new methods to estimate, intangible capital (see Peters and Taylor, 2017; Enache and Srivastava, 2018; Park, 2019; Eisfeldt, Kim and Papanikolaou, 2020; Lev and Srivastava, 2022). Most of those studies rely on a perpetual inventory model (PIM), that is, capitalizing and amortizing past R&D and SG&A expenditures, using uniform or industry-specific parameters. For example, 30% of SG&A creates future value and has a life of three years.

While the new methods, based on PIM, yield improvements over investment models that ignore in-house intangibles, they suffer from two limitations. First, no consensus exists on what percentage of intangible expenditures should be capitalized. Capitalization percentages used by those studies range from 30% to 100% (Iqbal et al., 2023). Second, any capitalized intangible stock measure, based on past expenses, does not take into account the lottery type payoff that often comes from serendipitous investments. For example, the discovery of a search formula by Google founders led to a trillion-dollar valuation company, and no amount of capitalization of past expenses would yield a number close to the value of that discovery. Stated differently, perpetual inventory models, even if correctly specified, would not capture the lottery type outcome from initial investments. Managers, however, are likely to describe the same successful discoveries and self-developed intangible assets in their communications with investors, if those assets are expected to create benefits for the company.

We contribute to the literature by identifying new measure of intangibles, which arguably tracks developed intangibles, that is, ex-post outcomes, but are not yet reflected in financial numbers presented in the balance sheet. We extract the informational content relating to intangible capital embedded in the text of a firm's 10-K filings. Financial Accounting Standards Board (FASB) Concepts Statement No. 5 prescribes strict criteria for recognition transactions in financial numbers, such as measurability, relevance, and reliability. Items that fail these criteria but are value-relevant nevertheless are often disclosed in footnotes and in the management discussion and analysis (MD&A) section of the 10-K. Managers describe their assessments of items that will impact future operations and whose discussion will enhance investors' understanding of firm operations. Any forward-looking information supplied in the MD&A section is expressly covered by the safe harbor rule, a legal provision that shields managers from liability if future projections go wrong. Merkley (2014) finds that narrative R&D disclosure in 10-K filings is positively associated with earnings forecast accuracy and predicts lower analyst forecast dispersion, suggesting that narrative disclosures reduce information asymmetry. Textual portion, thus, is particularly useful for conveying soft information (Seamons and Rouse, 1997). Hence, the textual portion of 10-K filings could provide guidance to investors on the value of internally generated intangible capital, particularly the value that cannot be recognized in the balance sheet for lack of reliability and constraints imposed by accounting rules. In addition, textual portion could describe serendipitous outcomes that would not be inferable by capitalizing past investments.

We conduct textual analysis to gauge the intensity of intangibles discussion in 10-K filings and consider it a proxy for intangible value. Our textual measure assumes that the intensity of the intangible-related discussion in 10-K filings represents the emphasis that management places on intangible development and its importance to firm operations. Our measure is the relative frequency of words on intangibles topics to total words in 10-K filings. Scaling the frequency of intangibles words by total words allows for comparison across firms with varying document sizes. We focus on three distinct categories of intangible assets: innovation assets and information technology, brand and customer relations, and human resources. Our decision to concentrate on these categories is motivated by similar classifications of intangible investments used in the literature (Lev, 2000; Corrado, Hulten and Sichel, 2005). We construct our measure of intangibles talk using a glossary of intangible terms created by Filipovic and Wager (2019). An additional benefit of using a text-based measure, relative to capitalizing past expenses, is the possibility of mapping words to and classifying them in, the three categories, while separately analyzing their informativeness with respect to future returns.

In testing the validity of our measure, we find that intangibles talk positively correlates with conventional accounting measures of intangible investments such as R&D and SG&A expenditures (both scaled by total expense). Also, intangibles talk correlates positively with indicators of intangible value called intangible capital advanced by Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020). Furthermore, our measure has a positive correlation with market-to-book ratio, which prior literature considers as proxy for intangible value not recognized on balance sheet. Considering that intangibles talk is constructed based entirely on textual material, as opposed to market or book data, these correlations point to a strong link between a firm's intangibles discussions in 10-K filings and its underlying intangible capital. We view this link as preliminary evidence that intangibles talk tracks variations in intangible capital and intensity over time.

We analyze variations in our measure across firms in Fama and French twelve industries. As expected, industries such as health care and business equipment score the highest in intangibles talk, and the lowest intangibles talk belongs to the finance and energy industries. We also decompose intangibles talk into three categories and investigate the highest-ranking industries under each of them. Results are consistent with intuition. For instance, while the healthcare industry scores the highest on intangibles talk focused on innovation assets and human resources, it scores among the lowest in the brand and customer relations category.

Our main tests of intangibles talk are informativeness with respect to future stock returns. Classic studies, as well as recent studies, investigate the mispricing of intangibles information. Lev and Sougiannis (1996) find a systematic mispricing of R&D-intensive stocks and show that incorporating that information leads to an annual return of 4.57%. Chan, Lakonishok and Sougiannis (2001) find that stocks with high R&D relative to the market value of equity deliver an average of 6.12% annual returns. They show similar results for stocks with high advertising expenses. Other studies link excess returns to patent citations (Deng, Lev and Narin, 1999), software developments (Aboody and Lev, 1998), and employee satisfaction (Edmans, Li and Zhang, 2014). Our argument on the link between intangibles talk and future returns, is similar to Edmans (2011), which points to the insufficient salience of intangibles information that could lead to it being overlooked by investors. Recent studies (Arnott et al., 2021; Choi, So and Wang, 2021; Lev and Srivastava, 2022) augment the book-to-market measure of value investing by intangible estimates and show superior returns than book-to-market based high minus low (HML) returns based on Fama and French (2015).

Prior studies on disclosure versus recognition argue that textual content on intangibles in documents such as 10-K filings are more likely to be ignored by investors than are numbers reported in income statement and balance sheet (Aboody, 1996; Davis-Friday et al., 1999; Ahmed, Kilic and Lobo, 2006). Based on this idea, we hypothesize that intangibles talk could convey at least some informative signals that likely are missed by investors. Eisfeldt, Kim and Papanikolaou (2020) show that intangible augmented value factors outperform the traditional value factor. They contribute to the literature by pushing the limits of the available accounting data to capture intangible value. We follow this idea to investigate our hypothesis. We first examine whether long-short portfolios formed on intangibles talk can generate positive, risk-adjusted returns. We then compare these results with returns generated by other benchmark value signals, namely, book-to-market ratio and its intangible augmented versions from Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020).

We begin our analysis by sorting portfolios based on intangibles talk, INT^{10K}. We follow the long-short sorting methodology presented in Fama and French (2015) and construct HML^{FF}, which captures the traditional value strategy based on book-to-market ratio. Similarly, we construct HML^{PT} and HML^{EKP} using the intangible augmented book-to-market ratio based on Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020), respectively.

We find that INT^{10K} outperforms the above three value strategies for the sample period, covering July 1995 to June 2020, while excluding the years that represent the dot-com bubble burst (2000–2001).[2] We achieve an annualized average monthly return of 5.05%. The returns are particularly high at around 7.6% in the period from 2008 to 2020. The average returns are consistently positive for each year since 2007, with the exception of 2012 and 2016 (see Fig. 1). INT^{10K} performs most strongly during the period that records the worst performance of HML^{FF} as shown in the post–financial crisis era (2008–2020) (see Eisfeldt, Kim and Papanikolaou, 2020). Our graph of cumulative returns is consistent with the growing importance of intangible value (see Fig. 2). Our portfolio's return displays steady growth over time and reaches its highest levels by the end of our study period in 2020. INT^{10K}'s outperformance relative to HML^{FF}, HML^{PT}, and HML^{EKP}, especially in recent years, suggests that incorporating textual information holds great promise as a separate source of intangible value for investors besides the data presented in accounting numbers. Furthermore, the fact that returns can be generated using textual data indicates that text-based intangibles information is not fully incorporated by the investors.

[2] We argue that the weak historical performance of INT^{10K} near the burst of the dot-com bubble results from the massive overvaluation of technology stocks in previous years. Technology stocks typically score high in intangibles talk and thus are likely picked up by our sorting methodology, which relies on intangibles talk values.

[Insert Figs.1 and 2 near here]

We conduct a more detailed examination on the relative informativeness of intangibles talk against other value indicators. We follow the strategy in Eisfeldt, Kim and Papanikolaou (2020); that is, we go long on INT^{10K} and short other intangibles-enhanced value portfolios. This strategy enables us to show that our measure captures orthogonal, and perhaps superior, information compared with augmented value strategies. We find that our long-short portfolio generates significant positive returns over HML^{FF} except for the period after the dot-com bubble (2000–2007). We find similar results when the short leg of the portfolio is changed to HML^{PT} and HML^{EKP}, providing strong evidence that our measure has additional, and arguably more unpriced, information than intangible augmented versions of value indicators documented in prior studies.

We next test whether the returns generated by intangibles talk represent premiums for risk or some other unpriced factor. We generate positive alphas while controlling for momentum (Carhart, 1997) and the three as well as the five Fama and French factors. The alpha averages 3.37% and 5.81% from 1995 to 2020 in the three- and the five-factor setting, respectively. The alpha remains positive and significant when we replace HML^{FF} with portfolios HML^{PT} and HML^{EKP} as the value factor in the Fama and French regression models. This suggests that the INT^{10K} return cannot be explained by traditional risk measures. We also create five value-weighted portfolios ranked based on intangibles talk and show that the alpha of the highest-ranking portfolio minus the lowest-ranking portfolio is 7.03% and significant.

We further examine the possibility that intangibles talk is an unpriced systematic risk factor by conducting a Fama-MacBeth two-stage regression analysis from Fama and MacBeth (1973). In the first stage, we first estimate the factor loadings of a set of test portfolios on the returns of the portfolio sorted based on intangibles talk. In the second stage, we test whether these factor loadings explain the variations in the returns of test portfolios. Our results suggest that intangibles talk is not consistently systematically priced in the cross-section of stock returns. This result is inconsistent with risk explanation.

We test mispricing as an explanation for return informativeness of intangibles talk. We rely on the idea that higher idiosyncratic volatility (IVOL) should amplify returns that result from mispricing. Prior studies show that higher IVOL leads to greater arbitrage risk, which, in turn, limits the ability of rational investors to correct mispricing (Pontiff, 1996; Stambaugh, Yu and Yuan, 2015)[3] In the case of intangible-intensive stocks, under higher IVOL market conditions, rational investors would bid less aggressively against overlooking of intangibles information. Accordingly, we argue that to the extent that intangible–intensive stocks are mispriced because of overlooked information, the amplification of such an effect should be more evident under market conditions with higher arbitrage risk, proxied by high IVOL. To test this, we first measure firm-level IVOL by estimating the volatility of the residuals from the three factors of Fama and French for daily returns (Ang et al., 2006). We then sort stocks by IVOL and estimate long-short returns based on our measure in each IVOL-sorted portfolio. Our double-sorted portfolio returns, based on IVOL and intangibles talk, are consistent with the idea of mispricing. We find that alphas of portfolios sorted on intangibles talk increase in IVOL, are most significant for high IVOL stocks, and are insignificant at low IVOL stocks. We therefore conclude that intangible–intensive stocks experience greater mispricing under high arbitrage risk, as evidenced by the strongest INT 10K abnormal returns among high IVOL stocks.

Our study mainly contributes to the emerging stream of literature, aiming to create alternative measures of intangible intensity, and using them to earn investment returns. Our study shows that managers could convey their value of intangible assets, through discus-

[3]They show that among overpriced (underpriced) stocks, the ones with the highest IVOL are the most overpriced (underpriced). Another recent study, Birru and Young (2020), utilizes IVOL to show stronger return predictions from investor irrationality.

sions and disclosures, particularly the value that is not represented in balance sheet and is not interpretable through capitalizing past investments. Our results also demonstrate that such information is mispriced by investors, especially in stocks characterized by high arbitrage risk.

In addition to contributing to intangibles-based value literature, our study is closely related to the literature on the use of textual analysis to predict returns (see Tetlock, 2007; Garcia, 2013; Jiang et al., 2019). We primarily rely on the bag of words approach in our textual analysis, which is the standard method used in the literature since Loughran and McDonald (2011). Most studies in the textual analysis literature measure tone sentiment or uncertainty. Our study in this regard deviates from such studies and is closer to those that use specialized glossaries to gauge the intensity of discussion surrounding particular topics.

This paper proceeds as follows. In Section 2, we describe the methodology used to construct intangibles talk and discuss and report its values across industries and firm characteristics. Section 3 presents the results from our portfolio analysis. Section 4 examines whether the common risk factors explain the portfolio returns. Section 5 tests the possibility that intangibles talk and other measures of intangible intensity are an unpriced systematic risk factor. Section 6 compares the performance of the portfolio sorted based on intangibles talk with technology stocks. In Section 7, we present and test the mispricing hypothesis that links the abnormal returns of our portfolio to market conditions with limits to arbitrage. Section 8 concludes.

2. Data and methodology

2.1 SEC filings

The corpus or textual data for our analysis come from 10-K filings submitted to the Securities and Exchange Commission (SEC) by 12,184 public firms from January 1994 to December 2021. This adds up to a total of approximately 107,000 10-K filings in our analysis. The filings are collected using the Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system. [4] In the parsing process, numbers, tables, and figures are removed for the raw text to be ready for textual analysis. To reduce the noise in the text, we remove all the stop words.[5]

To construct intangibles talk, we use the bag of words approach. A filing's vector of words is analyzed against a glossary of intangibles terms (see Table 1) based on the recently developed intangibles words list in Filipovic and Wager (2019). See the Appendix for examples. The list is derived from several studies on intangible assets such as Hall (2009) and Lev (2005, 2012). A more detailed description of how our intangibles word list was developed can be found in Filipovic and Wager (2019).

[Insert Table 1 near here]

We focus on three broad categories of intangible assets: innovation assets and information technology, brand and customer relations, and human resources. Lev (2005) claims that information technology has grown in importance in recent years with the rise of Internet platforms and software solutions. To account for this aspect of the organizational capital in intangibles talk, we create a combined category called innovation assets and information technology that is closer to what Wyatt (2008) calls technology resources. Table 2 shows that words under

[4] We use the parsed documents publicly available on the Loughran-McDonald website at: `https://sraf.nd.edu/loughranmcdonald-master-dictionary/`

[5] Examples of stop words are "a," "the," "are," "and," "could," and "would."

the innovation assets and information technology category include terms like databases, websites, and platforms.

[Insert Table 2 near here]

Table 2 contains the intangible word list belonging to the other two categories. In the brand and customer relations category, we include reputation, brands, and relations. In the last category, human resources, we emphasize skills, abilities, and competencies. Table 3 shows that most variance in intangibles words frequency across filings is related to the innovation assets and information technology category. This category contributes 47% to the total variance, and the words belonging to it account for 54% of the words in the glossary.

[Insert Table 3 near here]

We define intangibles talk for 10-K filing i with n intangibles words as the sum of the frequency across all intangibles words divided by the total number of words in the filing:

$$\textit{Intangibles talk}_{\mathbf{i}} = \frac{\sum_{\mathbf{j=1}}^{\mathbf{n}} \textit{Frequency of intangible word}_{\mathbf{j}}}{\textit{Total words}_{\mathbf{i}}} \tag{1}$$

The ratio, measured on a firm-year basis, reflects the relative intensity of discussion surrounding intangibles topics in filings. As Filipovic and Wager (2019) assert, the intangibles word list is neither optimized nor reverse-engineered to fulfill return maximization objectives. In a similar fashion, we do not select words from the intangibles word list that maximize our portfolio returns using machine learning techniques.

Fig. 3 plots the top 30 most frequent intangibles words across firms in our sample. Some of the most common words, such as ”employee,” ”customers,” and ”data,” can be found under various discussion topics that are not always related to intangible assets. Our glossary contains 128 words, of which the top three in Fig. 3 account for almost 20% of the total frequency while their share of the total frequency would be 2.3% under the equally distributed case. This illustrates the power law distribution of words frequency in natural languages, also known as Zipf's law.[6]

[Insert Fig. 3 near here]

2.2 Intangibles talk across industries

Fig. 4 depicts a summary of intangibles word frequency along with intangibles talk across the twelve Fama and French industries. Healthcare and business equipment rank the highest based on the median values of intangibles talk. The lowest-ranking industries are finance and energy. Intangible intensity in the healthcare and business equipment industries likely reflects their principal sources of competitive advantage, such as patents, data, technical knowhow, and information technology. In contrast, finance and energy rely heavily on tangible assets and financial capital. Later, we delve deeper into each industry's most frequent intangibles words to identify the main components of intangibles talk in that industry. Fig. 4, shows that the absolute frequency of intangibles words is high for some industries and the relative frequency (intangibles talk) can be low in some cases because of varying document sizes across industries.

[Insert Fig. 4 near here]

The portion of intangibles talk's three categories across the twelve industries reveal the concentration of each class of intangible assets in the economy. Fig. 5 shows that the consumer non-durables industry ranks high in the brand and customer relations category, indicating the importance of brands, sales and distribution network, and customer satisfaction in

[6] According to Zipf's law the frequency of words is proportional to the inverse of their ranking: f(r) $\propto \frac{1}{r^{\alpha}}, with \alpha \approx 1$ (see Mandelbrot, 1961).

that industry. The healthcare industry ranks relatively low in the brand and customer relations category but ranks high in the human resources and innovation assets categories. This points to the central role that qualified professionals play in the healthcare industry. They are responsible for the quality and efficiency of services. In addition, patents for new drugs and medical devices are important sources of revenue in the pharmaceutical industry.

[Insert Fig. 5 near here]

Another important feature of our measure is that it indicates the distribution of intangible value across firms and industries. Fig. 6 plots the top ten most frequent intangibles words by twelve industries. Words such as ”software”, ”data”, and ”technology,” that are associated with information technology appear to be frequent and concentrated in the business equipment industry. Another word, ”employee,” appears as the most frequent word across most industries. Admittedly, the context in which common words such as ”employee” and ”data” are discussed in 10-K documents is relevant to classify them as indicators of intangibles value. Our measure suffers from the limitation that it does not identify the context in which the words are used. Nonetheless, when a group of words that refer to a particular category of intangible assets is concentrated in one industry, it likely indicates the importance of that category to that industry's firms. For instance, the word ”customer” is relatively common across the majority of industries. Closely related words such as ”brand,” ”advertising,” ”franchise,” and ”marketing” are less common and are concentrated in industries such as consumer non-durables and wholesale and retail. This indicates the importance of brand and customer relations in consumer goods industries.

[Insert Fig. 6 near here]

Overall, the textual analysis shows that emphasis on different categories of intangible assets varies across industries, as captured by our measure. The variance also aligns with the expected industry characteristics, consistent with intuition.

2.3 Validity of intangibles talk measure

In this section, we examine the relationship between intangibles talk and firm characteristics. Table 4 reports the time series average of median firm characteristics for firms in the high, middle, and low range of intangibles talk. Prior studies consider R&D to total expenses, SG&A to total expenses, and enhancements in book value because of capitalized intangibles as proxies for intangible intensity.[7] We benchmark our measure against these proxies, by examining whether they vary in predicted directions across quantiles of firms formed based on our measure. All three proxies exhibit an increase by quantiles based on intangibles talk with notably high values amongst the firms in the third quantile. We also examine variation of the proxies across quantiles formed based on the three categories of intagibles talk.

[Insert Table 4 near here]

R&D to total expenses shows the highest variation with our measure in the innovation assets and information technology category, with a jump from 0.01 to 0.13 between the second and third quantiles. The increase from the second to third quantile is sizable for all three proxies of intangible intensity: R&D, SG&A, and intangible capital. This shows consistency between intangibles talk and proxies considered in the literature.

We also examine variations in firm characteristics, such as size, leverage, and profitability. No clear relationship is discernible between intangibles talk and sales-to-asset ratio, which

[7] R&D and SG&A are reported as expenses and occur over the normal course of business operations (i.e., flow variables). We divide them by total expenses instead of total assets to capture the variations in the flow of intangible investment over each year. This allows us to compare SG&A and R&D across firms with large and small asset bases.
Intangible capital which is constructed based on the methods in Peters and Taylor (2017) (intangible capital^{PT}) and Eisfeldt, Kim and Papanikolaou (2020) (intangible capital^{EKP}), is a stock of knowledge and organizational capital that accumulates over time through investments reported in SG&A and R&D. We use both versions of intangible capital in our analysis. Intangible capital^{PT} and Intangible capital^{EKP} are available at the GitHub page affiliated with Eisfeldt, Kim and Papanikolaou (2020): https://github.com/edwardtkim/intangiblevalue

represents asset turnover or efficiency in utilization of assets. However, debt to total assets falls with intangibles talk, showing that firms in the high quantile are the least leveraged. The opposite holds for the profitability-to-total assets ratio, with its highest value in the third quantile of intangibles talk. This aligns with the research on the relation between R&D and profitability (Lev and Sougiannis, 1996). However, it is consistent with Curtis, McVay and Toynbee (2020) who document a decline in the relation between R&D and profitability since the 1980s.

We also examine the traditional book-to-market ratio, along with its intangible augmented versions put forth by Peters and Taylor (2017) (book-to-market^{PT}) and Eisfeldt, Kim and Papanikolaou (2020) (book-to-market^{EKP}).[8] We observe that book-to-market ratio drops along the quantiles, with the lowest book-to-market ratios concentrated in the top quantiles of intangibles talk. This shows consistency between our measure of intangible value and the one implied by the book-to-market ratio (which is the opposite of market-to-book ratio). With respect to intangible augmented book-to-market ratios, their highest values are in the lowest quantile of intangibles talk, with less variation across quantiles relative to the as-reported book-to-market ratio itself. This arguably indicates that the intangibles-enhanced version of the book-to-market ratio is a less accurate indicator of intangible value than the as-reported one, because book value has already been at least partly corrected for intangibles.

We provide further evidence that intangibles talk is correlated with other indicators of intangible value in our panel regression analysis in Table 5. One standard deviation drop in the book-to-market ratio on average is associated with an approximately 0.18 standard deviation increase in intangibles talk. To capture the relation between SG&A and intangibles talk, we subtract R&D from SG&A because R&D is a constituent item of SG&A (Enache and Srivastava, 2018). As expected, our results show that SG&A (without R&D) and R&D,

[8]The formula is book-to-market^{PT} = (book value of equity + intangible capital^{PT} + goodwill)/market value of equity. Similarly, book-to-market^{EKP} uses intangible capital^{EKP} in the formula.

both scaled by total expenses, are positively related to intangibles talk. A standard deviation increases in SG&A and R&D to total expense ratios corresponds to about a 0.3 standard deviation rise in intangibles talk. A similar result holds for intangible capital based on both definitions of the variable from Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020).

[Insert Table 5 near here]

We view the results from Table 5 as evidence that intangibles talk and conventional indicators of intangible value are strongly correlated. This strengthens our main assumption that intangibles talk is a proxy for intangible value and intensity across firms. We consider whether our measure conveys additional information on future returns than do current proxies.

3. Returns analysis

In examining the returns performance of portfolios sorted based on intangibles talk, we seek to determine whether intangibles talk is informative about future returns. We use factor-mimicking portfolios to consider the informativeness of intangibles talk by analyzing portfolio returns over time while controlling for the common risk factors. We follow the long-short sorting method, described in Fama and French (2015), to construct a value-weighted portfolio. We use intangibles talk as an investment signal to identify the long and short portfolios every year in a sample of NYSE, Amex, and Nasdaq stocks, with data available from the Center for Research in Security Prices (CRSP). We analyze the period starting from 1995 to June 2020. [9] In this section, we describe our sorting methodology and long-short portfolio performance. We then conduct numerous analyses using several subperiods, such as before

[9] Given data requirements from the previous year, our portfolio sorting method generates returns from July 1995 based on intangibles talk values available from 1994.

and after the dot-com bubble burst and the global financial crisis. We then benchmark the returns of our portfolio against the value strategies based on book-to-market ratio, and its augmented versions with capitalized intangibles.

3.1 Intangibles talk and value strategies

We follow Fama and French in constructing long-short portfolios except that we use intangibles talk as the sorting variable instead of book-to-market ratio. We sort firms as of June 30 of each year, based on intangibles talk calculated from the 10-K filings for the fiscal year that ended in the previous calendar year. This method assumes that at least the December fiscal year-end firms have published their annual report by June 30 of the next year, consistent with Fama and French (2015). The portfolio (INT^{10K}) is constructed based on last reported intangibles talk. We identify stocks above the 70th percentile of intangibles talk, and put them in the long portfolios while shorting those below the 30th percentile.[10] Portfolios are held constant from July 1, following the June 30 portfolio formation date, to June 30 of the next year, except for delisted stocks. Monthly returns are calculated for each long and short portfolio, by value weighting returns of their constituent securities using their share in total market cap at the end of December of the previous year. Annual returns are calculated using monthly returns from January to June, based on portfolios formed in June of the previous year, and from July to December, based on portfolios formed in June of this year.[11]

We first compare the performance of INT^{10K} against portfolios formed based on book-

[10] A more detailed description of the sorting method can be found in Fama and French (2015). Using the NYSE median market cap as the breaking point, for the long leg of our portfolio, we average the returns of two portfolios, namely big and small stocks with high intangibles talk, and repeat the same procedure for the short leg of our portfolio.

[11] We test the validity of our sorting methodology, by constructing a portfolio based on book-to-market ratio as defined in Fama and French (2015). We achieve a 95% correlation with the HML^{FF} returns reported in the Kenneth R. French data library: (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). We use the same sorting methodology to construct other portfolios described in this study based on various firm-level intangible value indicators including our measure.

to-market ratio and its two intangible augmented versions (Peters and Taylor, 2017; Eisfeldt, Kim and Papanikolaou, 2020). In this analysis, HML^{FF} represents a value strategy that takes into account only reported value of assets. HML^{PT} and HML^{EKP} incorporate book value enhanced by non-reported intangible assets estimated with capitalization of past intangible investments. Notably, the definition of intangible capital used to calculate the augmented versions differs between Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020) and so do the returns generated by HML^{PT} and HML^{EKP}.

Performance statistics in Table 6 report the average monthly returns of the four portfolios for the period between July 1995 and June 2020. None of the returns is statistically significant from zero on average. Nevertheless, the returns from INT^{10K} are the highest, and the Sharpe ratios suggest that HML^{EKP} and INT^{10K} provide the best-performing strategies. Fig. 2 shows that the returns from INT^{10K} drop around the dot-com bubble, arguably because many intangible intensive companies suffered large negative returns, and numerous others were delisted. We also report the portfolio performances excluding 2000 and 2001, the peak bubble years. INT^{10K} now outperforms all the value strategies with statistically significant average annualized returns of 5.05%.

[Insert Table 6 near here]

A more detailed examination of the subperiods is reported in the last three columns of Table 6. The performance of INT^{10K} is stronger than other value strategies from 1995 to 1999, but no portfolio shows any statistically significant returns. During the pre–financial crisis period from 2000 to 2007, INT^{10K} performs poorly. The returns of HML portfolios are positive with high Sharpe ratios but still are not statistically significant from zero. Examining the post–financial crisis period (after 2008) is important because of HML^{FF} underperforming year after year during this period (Lev and Srivastava, 2022). INT^{10K} delivers the highest returns between 2008 to 2020 with a statistically significant return of 7.63% and a Sharpe ratio of 0.92, the highest amongst all four strategies. The results support two ideas: the growing importance of intangible value in the economy and our ex–post measure of intangibles talk capturing intangible capital better than the ex-ante capitalization methods of Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020).

Results based on intangibles talk follow the pattern presented in Eisfeldt, Kim and Papanikolaou (2020), showing that value investing based on intangible augmented book-to-market ratio outperform HML^{FF}, particularly in the post–financial crisis era. More importantly, the results demonstrate that our measure captures orthogonal, and arguably superior, information on intangibles compared with the ones based on capitalization of past expenses. Arguably, text-based measures either better capture ex-post value creation or are more likely to be ignored than value interpretable by capitalizing past expenses.

3.2 Other intangible intensity indicators

We turn to other indicators of intangible value from the literature and compare their portfolio performance against the performance of INT^{10K}. We create four portfolios based on four intangible proxies: SG&A expenses and R&D expenses, and both versions of intangible capital from Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020). We use intangible capital based on capitalized investment (scaled by total assets) to create the last two portfolios, not the book-to-market ratios augmented with intangibles. In constructing intangible capital, Peters and Taylor (2017) already account for R&D and SG&A expenses to a degree, and we include the two variables separately as well. We do this to account for any direct intangible value signal from these items that could be informative about future returns. We sort portfolios as of the end of June every year based on the ratio of R&D and SG&A to total expenses reported at the end of the last fiscal year. Similar to Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020), we sort portfolios based on intangible capital as of the end of the fiscal year in the previous calendar year. We divide these measures of intangible capital by total assets at the beginning of that fiscal year. We name them INT^{PT} and INT^{EKP} to distinguish them from HML^{PT} and HML^{EKP}. This exercise allows us to compare the returns of portfolios sorted on other measures of intangible intensity, which is a more comparable strategy to INT^{10K}.

The returns of these four portfolios, along with INT^{10K} are reported in Table 7. In the full sample between 1995 and 2020, none of the portfolios generates statistically significant returns. Similar to intangible augmented HML returns, when we exclude the years around the dot-com bubble, most strategies perform better, with INT^{10K} slightly outperforming the rest. Columns 3 to 5 do not report any significant outperformance by any of the portfolios, and the period from 2008 to 2020 presents significant positive returns for the INT^{10K}, INT^{PT}, INT^{EKP}, and $R\&D$ portfolios. The returns of INT^{10K}, and INT^{PT} are essentially the same, with INT^{10K} having a slightly larger Sharpe ratio. Therefore, the results from Table 7 show that intangibles talk is at least as informative about future returns as the other measures of intangible intensity in the literature, if not superior.

[Insert Table 7 near here]

3.3 Returns across categories of intangibles talk

We break down intangibles talk into its three categories based on the words associated with each category (see Table 2). We repeat the sorting strategy based on each category of intangibles talk and report the returns. The returns of the portfolios based on these three categories are presented in Table 8. The results are similar to INT^{10K} over the full period and the sub-periods. The highest returns are associated with the first category of intangibles talk related to innovation assets and information technology. However, the returns under all the categories are positive and significant when dot-com bubble years are excluded and become economically important especially in recent years. This indicates that the informativeness of intangibles talk is not limited to any single category and that the effect is present across all the categories.[12]

[Insert Table 8 near here]

4. Intangibles talk and common risk factors

We have established that INT^{10K} is informative about future returns and outperforms value strategies based on traditional and intangible augmented book-to-market ratio, especially in recent years. We next examine whether the returns associated with intangibles talk are simply premiums for known risk factors. We first test the hypothesis that the return for INT^{10K} is compensation for systematic risk. We regress our portfolio's returns against the systematic factors discussed in the literature, namely the three and five factors in the Fama and French (2015) models plus momentum. We control for all the factors in the regression:

$$\begin{aligned} R_t = \alpha + \beta_{MKT} \times MKT_t + \beta_{SMB} \times SMB_t + \beta_{HML} \times HML_t + \\ \beta_{RMW} \times RMW_t + \beta_{CMA} \times CMA_t + \beta_{UMD} \times UMD_t + \epsilon_t, \end{aligned} \quad (2)$$

where R_t is the return of INT^{10K} in month t, and α is the intercept that captures the abnormal returns after controlling for risk factors (alpha hereafter). MKT_t, SMB_t, HML_t, RMW_t, CMA_t, and UMD_t are, respectively, the returns of the market, size, value, profitability, investment, and momentum portfolios taken from Ken French's website. The stan-

[12] We also sort stocks based on intangibles talk within industries, using industry-specific benchmarks. We use the twelve Fama and French industry classifications and form our portfolio using stocks from each industry. The returns are insignificant and mostly positive except in the finance industry, for which the returns are positive and statistically significant. This indicates that the returns generated based on intangibles talk is not primarily a within-industry phenomenon, unlike value (see Asness, Porter and Stevens, 2000).

dard errors are estimated using Newey and West (1987), which allows for serially correlated and heteroskedastic error terms. The alphas are reported in Table 9 for the period between July 1995 and June 2020. The first two columns report the excess returns over the Fama and French three and five factors plus momentum. In the first column, as the baseline regression, we use HML^{FF} as the value factor. In the third and fourth columns, we use HML^{PT} and HML^{EKP} as the value factor, respectively, to account for intangibles-enhanced book value used in prior studies. The alphas remain positive and statistically significant in the first four columns, with the highest excess return, on an annualized basis, reported at 8.35%, and the lowest at 3.37%. The positive and significant alphas show that the informativeness of intangibles talk with respect to future returns is not fully explained by the common risk factors: market, size, value, profitability, investment, and momentum. This result rules out the possibility that the returns associated with intangibles talk capture the effects of profitability (RMW). Profitability is positively associated with intangibles talk (see Table 4). In columns 5 to 7 in Table 9, we also include INT^{PT}, INT^{EKP}, $R\&D$ and $SG\&A$ portfolios in our regressions. The excess returns remain positive and statistically significant, indicating that the informativeness of intangibles talk is not captured by other measures of intangible intensity constructed based on capitalizing past reported R&D and SG&A expenses.

[Insert Table 9 near here]

The abnormal returns are similar for each category of intangibles talk as well. Table 10 reports the alphas over the Fama and French factors for the three portfolios corresponding to each category of intangibles talk. The highest alpha belongs to innovation assets and information technology with 3% and 5.3%, respectively, for the Fama and French three and five factors plus momentum. Therefore, while the magnitude of alpha varies across categories, the abnormal returns associated with intangibles talk are not limited to a particular class of intangible assets.

[Insert Table 10 near here]

The limited availability of textual data from 10-K filings restricts the exploration of excess returns from before 1995. However, the results here are comparable with the 3.48% excess returns (over the three Fama and French factors plus momentum) generated by the value-weighted portfolio that picks stocks based on employee satisfaction, reported between 1984 and 2009 by Edmans (2011). Similarly, the excess returns associated with R&D and advertising expenses reported in Lev and Sougiannis (1996) is around 4.57% in their sample, which goes back to 1975. Hence, our results point to the link between intangible value and subsequent returns that seem to be persistent over the most recent years.

Table 11 better illustrates the relatively cheap (expensive) valuation of stocks with high (low) levels of intangibles talk in the market. We construct five portfolios of stocks ranked based on intangibles talk and estimate the alphas over the five-factor model. In addition, we investigate whether the observed excess returns in Table 9 are associated with intangible intensity in general. We repeat the exercise based on other intangible intensity indicators discussed so far and report their alphas over the five-factor model as well. These excess returns are presented in Table 11.

[Insert Table 11 near here]

The highest-ranking portfolio based on intangibles talk generates significant positive alpha, while a significant negative alpha is reported for the lowest-ranking portfolio, suggesting that intangible intensive stocks, as measured by intangibles talk, are underpriced. The results for the portfolio of stocks ranked based on R&D to total expenses also confirm the previous findings by reporting a positive and significant alpha for the highest-ranking portfolio. When stocks are ranked based on intangible capital to total assets using the Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020) measures, the alphas are not significant for the highest-ranking portfolios, but the lowest-ranking portfolio in the case of intangible capital^{PT} generates a significant negative alpha.

Regardless of the intangibles measured used to generate alphas, high versus low intangibles talk better illustrates the directional abnormal returns than other indicators such as R&D and capitalized intangibles. The results from this section and Section 3 are consistent with our hypothesis that the information on intangibles embedded in textual disclosures of 10-K filings is not fully considered by investors. However, another plausible explanation is that intangibles talk, along with other measures of intangible intensity from the literature that can lead to abnormal returns, are proxies for a new systematic risk factor that is not explained by the commonly known risk factors in the literature.

5. Is intangibles talk a systematic risk factor?

To investigate the extent to which intangible intensity, as measured by intangibles talk and other intangible-augmented value factors, can be considered systematic risk factors that influence the cross-sectional variation in stock returns, we start by constructing aggregate value-weighted intangible factors using the method of Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020), as well as intangibles talk developed in this paper. We simply include these factors in the Fama and MacBeth two-stage regressions as in Fama and MacBeth (1973) to estimate the price that investors pay to bear the risk of these intangible factors.

To ensure that our estimated prices of intangibles' risk are robust, we use alternative sets of test portfolios. By examining various sets of test portfolios, we reduce the likelihood of finding 'false' factors in any one set of test portfolios as pointed out by Lewellen, Nagel and Shanken (2010). We conduct two-stage regressions for each set of portfolios in the Kenneth R. French data library, which has at least ten portfolios.[13] Overall, we have $K = 36$ sets

[13] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.

of test portfolios, listed in Table12. The test portfolios are created using a diverse set of firm characteristics, such as size, market-to-book, investment, repurchases, and several others. The sets of portfolios, therefore, are unlikely to be based on similar sets of firms.

[Insert Table 12 near here]

We use the Fama-Macbeth (FM) two-stage procedure as in Fama and MacBeth (1973) to estimate the prices of intangibles risk. Consider the portfolio set $k \in 1, 2, \cdots, K$. We obtain monthly returns on the factors (MKT, SMB, HML, and UMD) from the Kenneth R. French data library. Each month, risk loadings are estimated using the current and previous thirty-five monthly returns for the first-stage equation:

$$\begin{aligned} R^p_{t-s} = \alpha_{p,t} + \beta^{MKT}_{p,t} \times MKT_{t-s} + \delta_{p,t} \times INT_{t-s-1} \\ + \beta^{SMB}_{p,t} \times \textbf{\textit{SMB}}_{t-s} + \beta^{HML}_{p,t} \times \textbf{\textit{HML}}_{t-s} + \beta^{UMD}_{p,t} \times \textbf{\textit{UMD}}_{t-s} + \epsilon_{p,t-s}, \end{aligned} \quad (3)$$

for $p = 1, ..., N_k$, $s = 0, \cdots, 35$. In this equation, $\boldsymbol{R}^p_t$ is the monthly return of the p^{th} portfolio at t, N_k is the number of portfolios of stocks in portfolio set k, and INT_t is either the INT^{PT}, INT^{EKP}, or INT^{10K} factor. Next, for the second-stage, we estimate for each month the prices of risk at date t for portfolio set k using the equation

$$\begin{aligned} R^p_t = \kappa^k_t + \pi^k_t \times \beta^{MKT}_{p,t} + \omega^k_t \times \delta_{p,t} \\ + \phi^{SMB,k}_t \times \beta^{SMB}_{p,t} + \phi^{HML,k}_t \times \beta^{HML}_{p,t} + \phi^{UMD,k}_t \times \beta^{UMD}_{p,t} + \epsilon_{t,p}, \end{aligned} \quad (4)$$

for $p = 1, ..., N_k$. Finally, following Cochrane (2001), we calculate the time-series mean of the estimated prices of risk and the estimated variance (our equations illustrate the case of the

html

price of intangibles risk), for portfolio set k as

$$\hat{\omega}^k = \frac{1}{T}\sum_{t=1}^{T}\hat{\omega_t^k} \tag{5}$$

and,

$$\sigma^2(\hat{\omega^k}) = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\omega_t^k} - \hat{\omega^k})^2, \tag{6}$$

respectively. The prices of risk of the other factors are estimated analogously.

In Fig. 7, we report the average prices of risk from the second stage regression for each of the factors in the top panels. As can be seen, the prices of risk for each of the three factors are very similar; they are almost all positive but small, mostly less than 0.06. The middle panels show the t-statistics of the estimated coefficients. As seen, most of the estimated coefficients are not statistically significant at the five percent level. Overall, neither of the factors seem to be consistently systematically priced in the cross-section of stock returns. The bottom panels show the R^2 of the 2nd stage regression for each of the three factors (along with the four-factor model). The R^2s using each factor are very similar and range between 20% and 55% for the different sets of portfolios. We, therefore, conclude that there is no compelling evidence that either of the intangibles-based factors significantly improves the factor specification of the four-factor model.

[Insert Fig. 7 near here]

6. Intangibles talk versus. technology stocks

A plausible explanation for abnormal returns of the portfolio sorted based on intangibles talk, INT^{10K}, can be attributed to the strong performance of technology stocks over the past decade. Because technology firms rely more on intangible assets, discussions of intangibles in documents such as 10-K filings are more likely ot be from this sector, which elevates their measured intangibles talk. This is what we find when we plot the distribution of intangibles talk values across the twelve Fama and French industry classifications (see Figs. 4 and 5). Notably, the business equipment industry ranks the highest among the twelve industries in intangibles talk, especially in the category of innovation assets and information technology. This suggests that the portfolio sorted based on intangibles talk may partially capture the strong performance of technology stocks by taking a long position in stocks in the highest percentiles of intangibles talk in the market. To address this concern, we conduct two exercises to separate the effects of intangibles talk from the outperformance of technology stocks relative to other industries.

For our first exercise, we construct an INT^{10K} portfolio independently within each of the twelve Fama and French industries and investigate the magnitude of its abnormal returns against the common risk factors. Given that INT^{10K}'s strong performance reported in the previous sections may come from technology stocks outperformance relative to other industries, we should not observe a statistically significant alpha for INT^{10K} within the business and equipment industry. This, however, is not the case based on the results reported in Table 13. The alphas of INT^{10K} are significant in only four industries, with its highest values reported in the business and equipment industry at around 6% annually. Therefore even among technology stocks, returns performance can be significantly enhanced by sorting stocks based on high versus low intangibles talk. Other industries with significant alphas are health care, finance, and other industries based on the twelve Fama and French industry classifications.

[Insert Table 13 near here]

Our second exercise uses a blind portfolio strategy in which we take long positions on technology stocks and short stocks from other industries. We call this portfolio $Tech^{portf}$. Like other portfolios in this paper, $Tech^{portf}$ is value-weighted and re-balanced at the end of June every year. We first investigate whether such a strategy would yield a significant and positive abnormal return against the common risk factors. We then test whether $Tech^{portf}$ can explain INT^{10K}'s abnormal returns and vice versa.

Table 14 reports the results for this exercise. $Tech^{portf}$ has a significant and positive alpha over the Fama and French five factors plus momentum. This alpha, however, is explained when we include INT^{10K} in the regression specifications in columns (3) and (4). Also, in the last two columns, the positive alpha of INT^{10K} remains significant when we include $Tech^{portf}$ in our regression specifications. The results from Table 14 show that including INT^{10K} in factor regressions renders the alpha of $Tech^{portf}$ insignificant, while INT^{10K}'s alpha remains significant even after including $Tech^{portf}$ in factor regressions. Overall a blind strategy of buying technology stocks and shorting other stocks, even though achieving a positive alpha, does not explain the alpha linked to intangibles talk.

[Insert Table 14 near here]

The results from this section show that a portfolio strategy based on intangibles talk leads to significant positive returns even among technology stocks that cannot be explained by the common risk factors. Furthermore, the positive alpha of a portfolio strategy with an emphasis on technology stocks is explained by the portfolio return generated using intangibles talk as an investment signal.

7. Do intangibles returns represent mispricing?

Is return informativeness of intangibles talk related to mispricing associated with limits to arbitrage? To answer this question, we rely on IVOL as a measure for limits to arbitrage that could amplify mispricing. To the extent that the INT^{10K} alphas come from limits to arbitrage for mispriced stocks, they should increase in IVOL.

Diversification of idiosyncratic risks is central to the capital asset pricing model, yet its limiting effect on arbitrage is well proven in the literature. Stambaugh, Yu and Yuan (2015) claim that adverse price moves are more likely under higher IVOL and, therefore, high IVOL is a source of arbitrage risk. This is because capital-constrained investors are forced to close their positions prematurely, under high IVOL conditions, before subsequent price corrections occur. We follow Stambaugh, Yu and Yuan (2015) as well as subsequent studies that consider IVOL as a proxy for mispricing (see Cao and Han, 2016; Birru and Young, 2020) to examine whether our results differ under different IVOL conditions.

Stambaugh, Yu and Yuan (2012) investigate the effects of IVOL on mispricing by constructing double-sorted portfolios based on IVOL and a mispricing measure. They examine eleven return anomalies identified in the literature that survive the three factors of Fama and French. Their results indicate that the degree of mispricing is higher among high IVOL stocks in that the returns generated by going long on putative underpriced stocks and going short on putative overpriced stocks increases in IVOL. They also show that the negative IVOL-return relation (the so-called idiosyncratic volatility puzzle) exists among overpriced stocks and that the reverse is true for underpriced stocks (positive IVOL-return relation).

To the extent that returns from intangibles talk are because of mispricing of intangible-related information in 10-K filings, we expect an amplification of such an effect under high IVOL conditions. In addition, in line with the results in Stambaugh, Yu and Yuan (2012) on the IVOL-return relation, we expect the IVOL-return relation to be negative among stocks that score the lowest in intangibles talk (that is, the putative overpriced stocks) and to be positive for the highest in intangibles talk stocks (that is, putative underpriced stocks).

We test our hypothesis by creating twenty-five double-sorted portfolios, five times five each, based on intangibles talk and IVOL. We then estimate the alphas of each portfolio over the five factors of Fama and French to examine the abnormal returns. Table 15 reports the alphas for the twenty-five portfolios. The difference between abnormal returns of the highest and lowest-ranking portfolios (high minus low alpha) based on intangibles talk (rows) is not significantly different from zero among stocks with low IVOL. Meanwhile, the alpha difference for the bottom two rows with the highest IVOL is significant and positive. To account for size effects, we also report the alphas for the twenty-five portfolios separately for small and big firms. Table 16 shows that the results are similar for small and big firms separately. The results for the big firms imply the same, with the alpha difference being 19.47% among stocks with the highest IVOL.

[Insert Table 15 and 16 near here]

In Fig. 8, we plot the strong positive relation between IVOL and INT^{10K} abnormal returns over Fama and French's five factors. The highest alpha belongs to INT^{10K}, constructed using the stocks in the highest IVOL decile in our sample. More importantly, a positive, and increasing, trend appears in the magnitude of alpha across IVOL deciles. Alphas go up to above 50% average annualized monthly returns in the highest decile of IVOL.

[Insert Fig. 8 near here]

Another observation from Table 15 and Table 16 comes from the sorting of stocks based on IVOL (columns), demonstrating the negative relation between IVOL and subsequent returns. This aligns with Ang et al. (2006) and the literature on the idiosyncratic volatility puzzle (negative IVOL-return relation). However, this relation disappears among stocks with the highest intangibles talk, for both small and big firms. This result aligns with Stambaugh, Yu and Yuan (2015), which shows that the negative relation of IVOL and returns is stronger for overpriced (low intangibles talk) stocks and is positive among underpriced (high intangibles talk) stocks.

Overall, the results suggest that the abnormal returns associated with intangibles talk is significant and growing in size with higher levels of IVOL and, thus, with limits to arbitrage. The positive relation between alpha and IVOL is also in line with Birru and Young (2020), which finds that investor sentiment is a stronger predictor of subsequent returns (mispricing) in the presence of uncertainty (measured through IVOL).

The abnormal returns of INT^{10K} reported in Section 4 indicate that the informativeness of intangibles talk is not captured by the common risk factors. The sorting exercise provides strong evidence for the association between IVOL and INT^{10K} abnormal returns. This supports the mispricing hypothesis and therefore ties our findings to the literature that examines mispricing in relation to the limits to arbitrage.

8. Conclusion

We construct a textual measure of intangibles talk using 10-K filings and test its informativeness for future stock returns. We devise a long-short portfolio based on our measure and report significant and positive returns between 1995 and 2020, after excluding the years 2000 and 2001. We compare our portfolio's returns with value strategies based on traditional book-to-market ratio and its intangible–augmented versions reported in the literature and document a superior performance, especially in recent years.

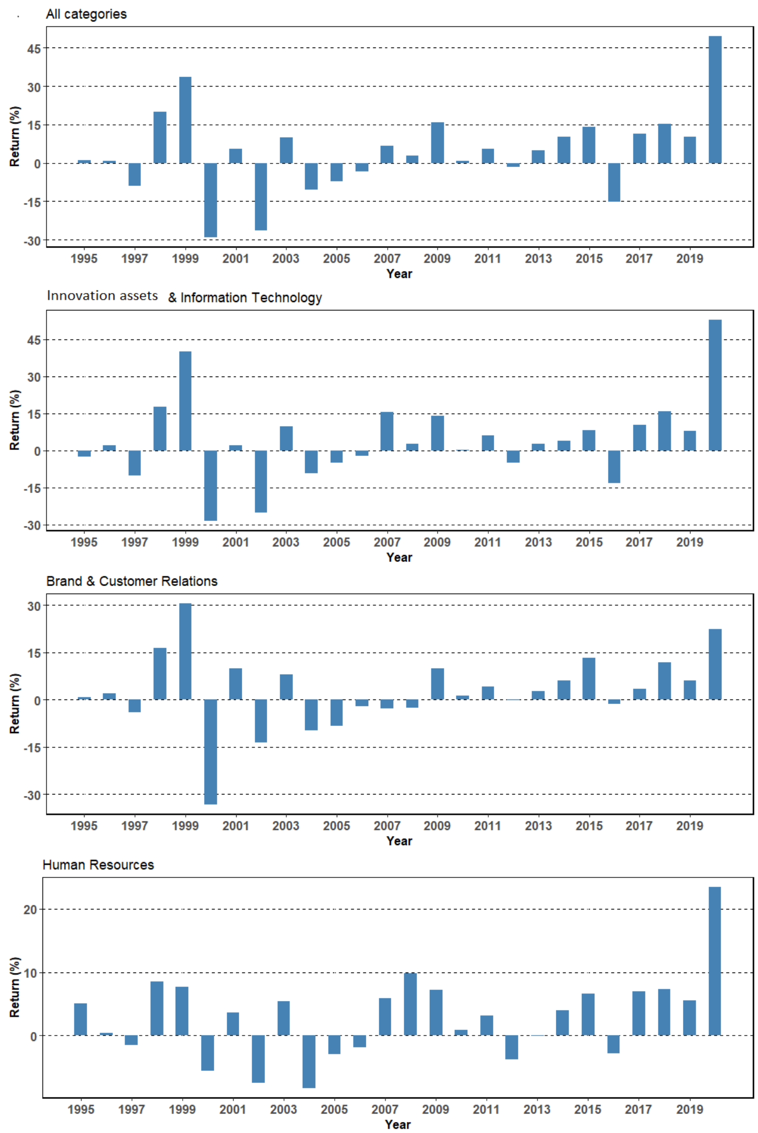
We test whether the common risk factors from Fama and French explain intangibles talk's informativeness. Our results indicate a positive and significant alpha over the three- and five–factor models from Fama and French in addition to momentum. We achieve the same results when we replace the value factor in the regressions with its intangible augmented versions from the literature. An implication of the abnormal returns of our portfolio is that investors do not fully price intangibles information disclosed in 10-K filings, perhaps due to the challenging nature of detecting, defining, and valuing intangible assets from that information. Our results align with similar findings on mispricing of R&D and employee satisfaction in stock valuations.

We test the mispricing hypothesis for intangibles talk's abnormal returns in the presence of limits to arbitrage. We find that the abnormal returns associated with intangibles talk is present only among stocks with high levels of IVOL, which measure arbitrage risks. This supports our mispricing hypothesis and shows that the limits to arbitrage exacerbate the mispricing of intangibles talk.

Our contribution to the literature is twofold. First we show that value-relevant information on successful intangibles investments can be measured from firms' textual disclosures. Second, this information is not fully incorporated in stock prices by investors.

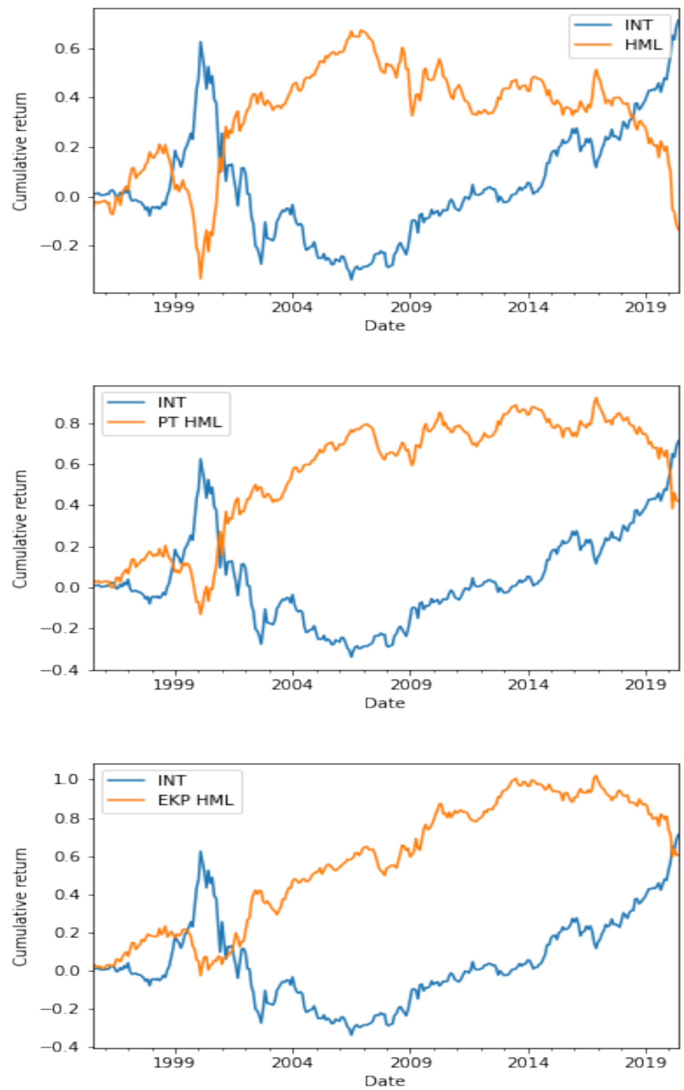
Tables and figures

Figure 1: Average Monthly Return Performance



The figure above plots the average monthly returns of INT^{10K} for each year between July 1995 to June 2020. For each category stocks are sorted based on intangibles talk measure in that category. The returns are in percent per year (monthly return multiplied by twelve). Detailed calculations are described in Section 3.

Figure 2: Cumulative Return Performance



The figure above plots the cumulative returns of one dollar invested in INT^{10K} in comparison to value strategies HML^{FF}, HML^{PT}, and HML^{EKP} for the period between July 1995 to June 2020. Detailed calculations are described in Section 3.

Table 1: Intangibles words glossary

ABILITIES	CUSTOMER RELATION	INNOVATION	NETWORKS	TALENT
ABILITY	CUSTOMERS	INNOVATIONS	PATENT	TALENTS
ADVERTISING	DATA	INNOVATOR	PATENTED	TEAM
ALGORITHM	DATABASE	INNOVATORS	PATENTS	TEAMS
AUTHORSHIP	DATABASES	INTELLECTUAL	PLATFORM	TEAMWORK
AUTHORSHIPS	DESIGN		PLATFORMS	TECHNOLOGIES
BRAND	DESIGNS	INTELLECTUAL PROPERTY	PRESENCE	TECHNOLOGY
BRANDING	DISCOVERIES	INTERNET	PRODUCTIVITY	TRADE MARK
BRANDS	DISCOVERY	INTERNET ACTIVITIES	PROTECTED DESIGN	TRADE MARKS
CLIENT	EMPLOYEE	INTERNET ACTIVITY	PROTECTED DESIGNS	TRADE NAME
CLIENT RELATIONS	EMPLOYEES	INVENT	REGISTERED DESIGN	TRADE SECRET
CLIENTS	EXPERIENCE	INVENTED	REGISTERED DESIGNS	TRADE SECRETS
COMPETENCE	EXPERT	INVENTING	RELATION	TRADEMARK
COMPETENCES	EXPERTISE	INVENTION	RELATIONS	TRADEMARKS
COMPETENCIES	EXPERTS	INVENTIONS	RELATIONSHIP	TRAINING
COMPETENCY	FORMULA	INVENTS	RELATIONSHIPS	USER
CONNECTIONS	FORMULAE	KNOWHOW	REPUTATION	USERS
CONNECTIVITY	FRANCHISE	KNOWLEDGE	RESEARCH	WEBSITE
CONSUMER	FRANCHISES	LABEL	RESEARCHES	WEBSITES
CONSUMERS	HUMAN	LABELS	SITE VISITS	WORKFORCE
COPYRIGHT	HUMAN CAPITAL	LICENCE	SKILL	
COPYRIGHTS	HUMAN RESOURCES	LICENCES	SKILLS	
CUSTOMER	INNOVATE	LOGO	SOFTWARE	
CUSTOMER BASE	INNOVATE PARTNERS	LOYALTY	SOLUTION	
CUSTOMER BASES	INNOVATED	MARKETING	SOLUTIONS	
CUSTOMER LIST	INNOVATES	NAMES	SYSTEM	
CUSTOMER LISTS	INNOVATING	NETWORK	SYSTEMS	

This table shows 128 words and terms used to calculate intangibles talk measure for each 10-K filings. The words in the table are based on Filipovic and Wager (2019). Calculations of intangibles talk measure is described in Section 2.

Table 2: Intangibles words by categories of intangibles

Innovation assets & information technology		Brand & customer relations	Human resources
INNOVATE PARTNERS	INNOVATOR	CLIENT RELATIONS	HUMAN CAPITAL
INTELLECTUAL PROPERTIES	INNOVATORS	CUSTOMER BASE	HUMAN RESOURCES
INTELLECTUAL PROPERTY	INTELLECTUAL	CUSTOMER BASES	ABILITIES
INTERNET ACTIVITIES	INTERNET	CUSTOMER LIST	ABILITY
INTERNET ACTIVITY	INVENT	CUSTOMER LISTS	COMPETENCE
PROTECTED DESIGN	INVENTED	CUSTOMER RELATION	COMPETENCES
PROTECTED DESIGNS	INVENTING	ADVERTISING	COMPETENCIES
REGISTERED DESIGN	INVENTION	BRAND	COMPETENCY
REGISTERED DESIGNS	INVENTIONS	BRANDING	EMPLOYEE
SITE VISITS	INVENTS	BRANDS	EMPLOYEES
TRADE MARK	KNOWHOW	CLIENT	EXPERIENCE
TRADE MARKS	KNOWLEDGE	CLIENTS	EXPERT
TRADE NAME	LICENCE	CONNECTIONS	EXPERTISE
TRADE SECRET	LICENCES	CONNECTIVITY	EXPERTS
TRADE SECRETS	NETWORK	CONSUMER	HUMAN
ALGORITHM	NETWORKS	CONSUMERS	PRODUCTIVITY
AUTHORSHIP	PATENT	CUSTOMER	SKILL
AUTHORSHIPS	PATENTED	CUSTOMERS	SKILLS
COPYRIGHT	PATENTS	FRANCHISE	TALENT
COPYRIGHTS	PLATFORM	FRANCHISES	TALENTS
DATA	PLATFORMS	LABEL	TEAM
DATABASE	RESEARCH	LABELS	TEAMS
DATABASES	RESEARCHES	LOGO	TEAMWORK
DESIGN	SOFTWARE	LOYALTY	TRAINING
DESIGNS	SOLUTION	MARKETING	WORKFORCE
DISCOVERIES	SOLUTIONS	NAMES	
DISCOVERY	SYSTEM	PRESENCE	
FORMULA	SYSTEMS	RELATION	
FORMULAE	TECHNOLOGIES	RELATIONS	
INNOVATE	TECHNOLOGY	RELATIONSHIP	
INNOVATED	TRADEMARK	RELATIONSHIPS	
INNOVATES	TRADEMARKS	REPUTATION	
INNOVATING	WEBSITE	USER	
INNOVATION	WEBSITES	USERS	
INNOVATIONS			

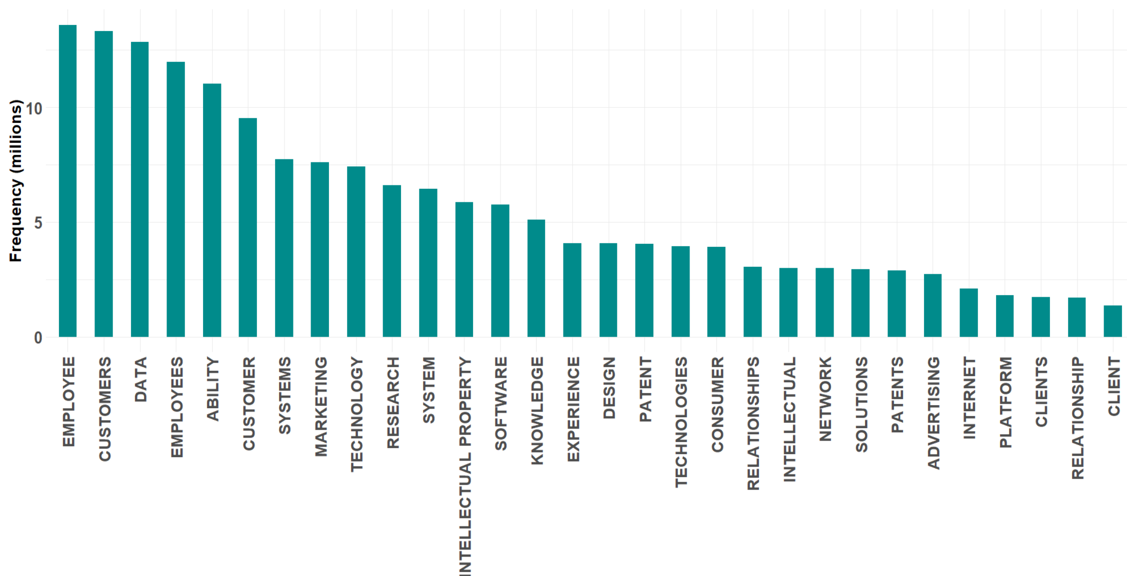
This table shows the list of words to calculate intangibles talk measure for each 10-K filing, by three categories of intangibles. The words in the table are based on Filipovic and Wager (2019). Calculations of intangibles talk measure is described in Section 2.

Table 3: Variance shares of three intangibles categories in intangibles talk measure

Category	Portion of glossary	Variance share	Top words
Innovation assets & information technology	54%	46.74%	Data, System, Technology, Research, Intellectual Property
Brand & customer relations	27%	28.6%	Customers, Marketing, Consumer, Advertising, Relationship
Human resources	19%	24.66%	Employee, Ability, Experience Expertise, Talent

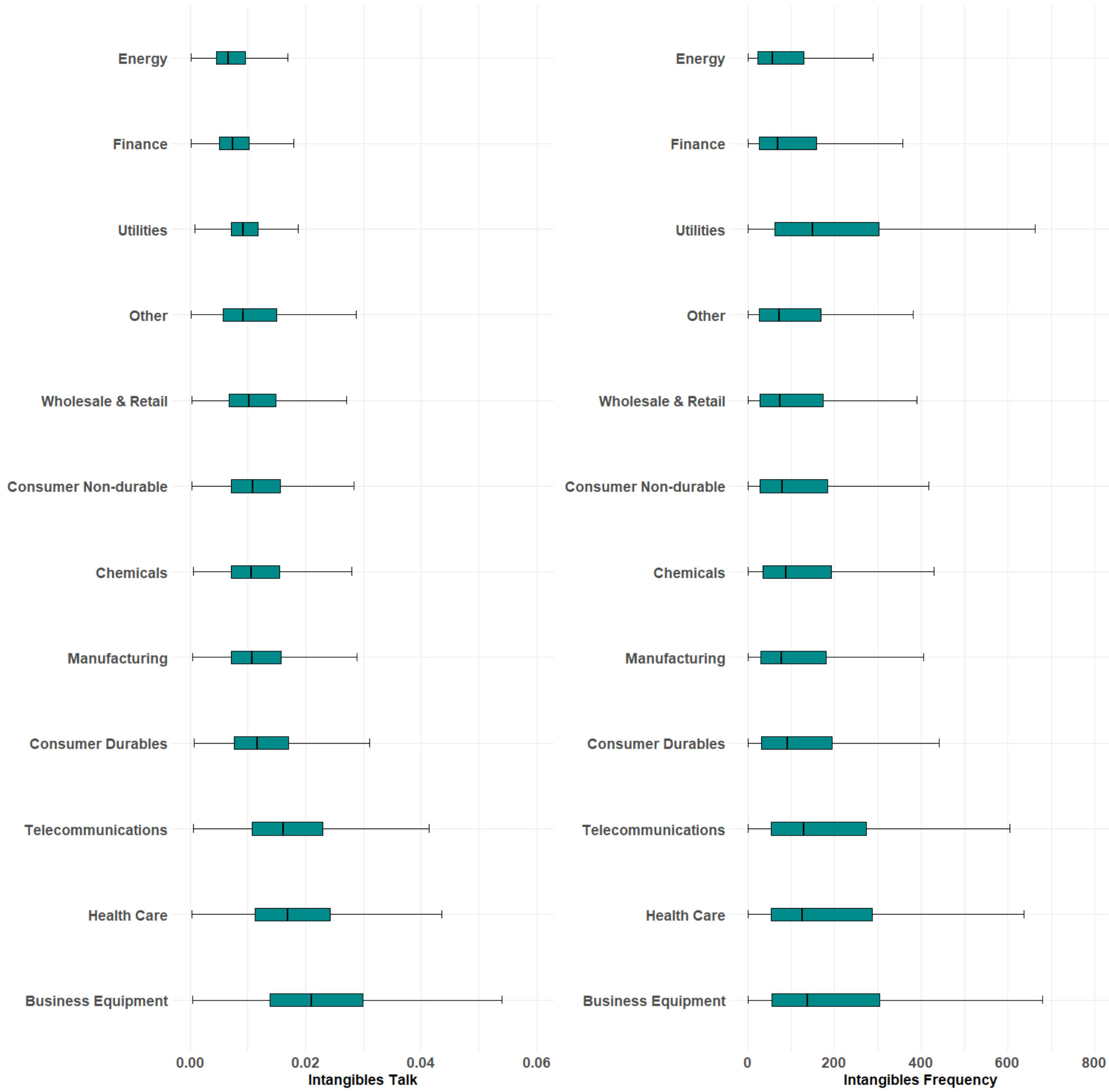
The table reports the share of the total variance coming from each category of intangibles words. The total variance is [$\Sigma_j \, v(j)$] summed across all the words in our glossary and $v(j)$ is the variance of frequency of word j across all the filings in our sample.

Figure 3: Most frequent terms appearing in intangibles talk measure across all firms



This figure shows the frequency of words appearing in intangibles talk measure. The words in the table are based on Filipovic and Wager (2019). Calculations of intangibles talk measure is described in Section 2.

Figure 4: Intangibles talk measure for each Fama and French twelve industries

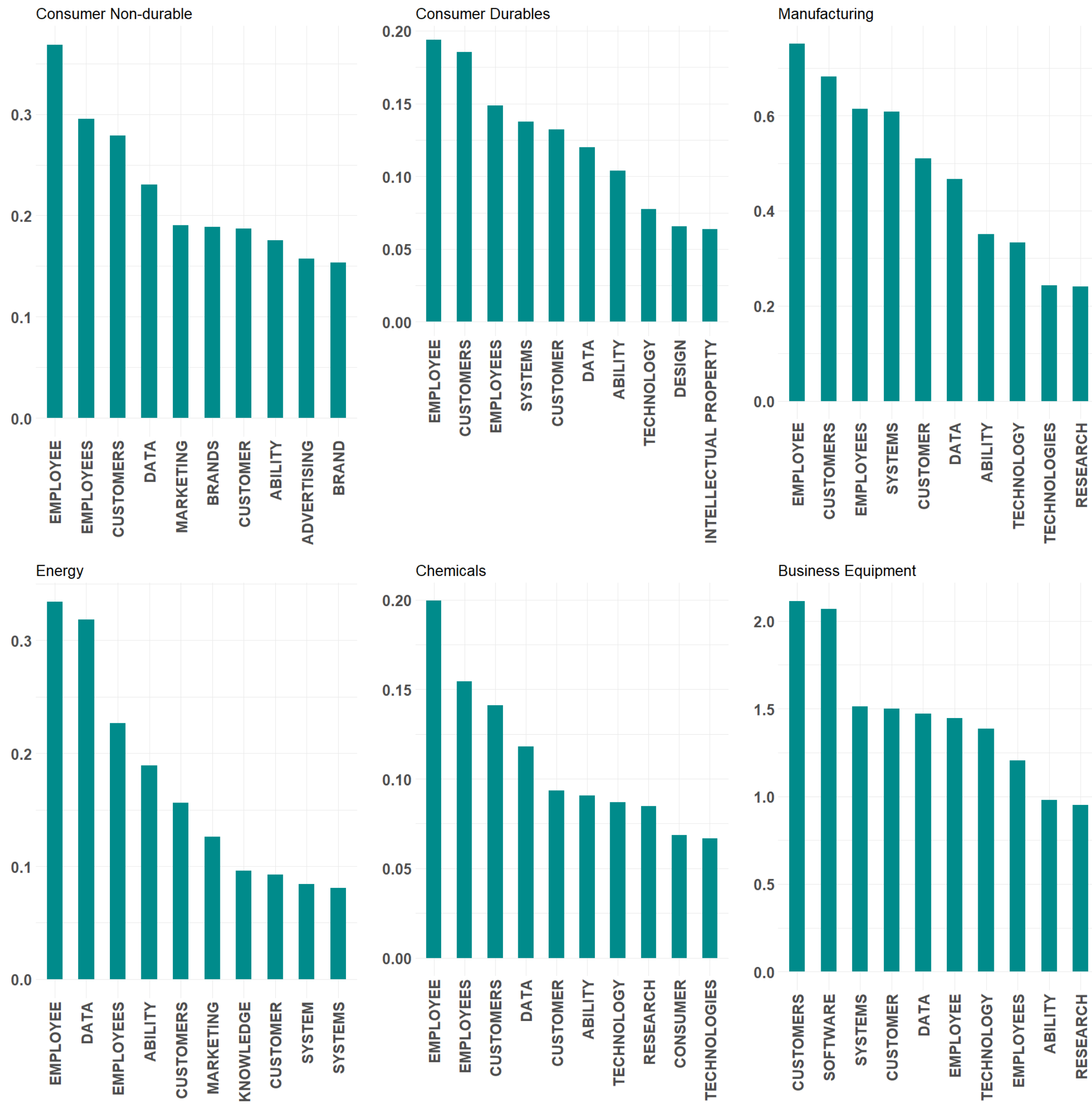


This figure shows the Box and Whisker plot for intangibles talk measure for each Fama and French industry category, with dark central line in the middle representing the median value, and the box representing variation from 25th to 75th percentile. Calculation of intangibles talk measure is described in Section 2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019).

Figure 5: Distribution of intangibles talk measure by three intangible categories in Fama and French twelve industries

This figure shows the Box and Whisker plot for intangibles talk measure for each of the three intangible categories, in each Fama and French industry category. Dark central line in the middle representing the median value, and the box representing variation from 25th to 75th percentile. Calculation of intangibles talk measure is described in Section 2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019).

Figure 6: Most frequent terms entering intangibles talk measure by Fama and French twelve industries



This figure shows the frequency of words appearing in intangibles talk measure, by Fama and French industry categories. The words in the figure are based on Filipovic and Wager (2019). Calculation of intangibles talk measure is described in Section 2.

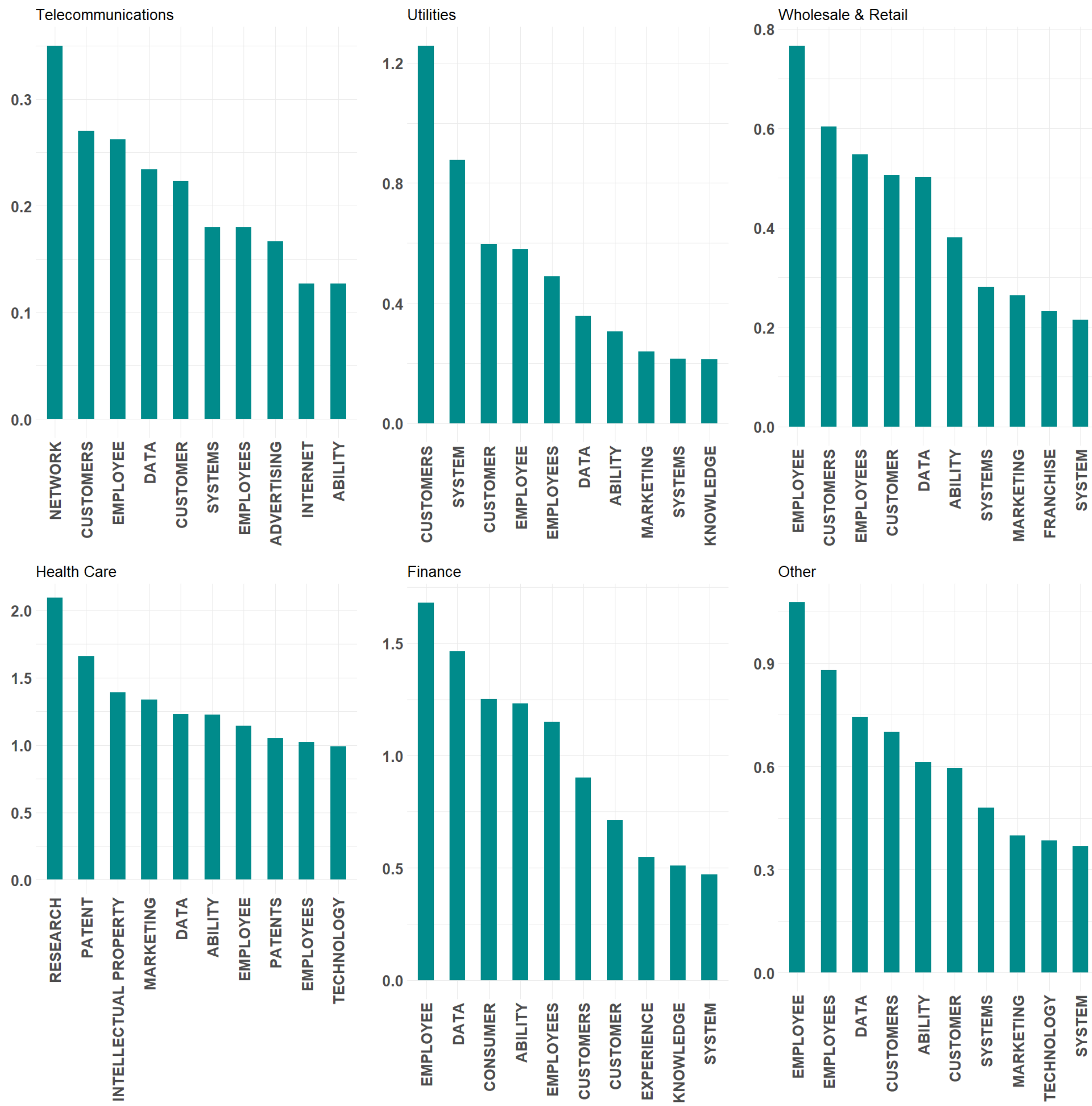

Telecommunications
0.0
0.1
0.2
0.3
NETWORK
CUSTOMERS
EMPLOYEE
DATA
CUSTOMER
SYSTEMS
EMPLOYEES
ADVERTISING
INTERNET
ABILITY
Utilities
0.0
0.4
0.8
1.2
CUSTOMERS
SYSTEM
CUSTOMER
EMPLOYEE
EMPLOYEES
DATA
ABILITY
MARKETING
SYSTEMS
KNOWLEDGE
Wholesale & Retail
0.0
0.2
0.4
0.6
0.8
EMPLOYEE
CUSTOMERS
EMPLOYEES
CUSTOMER
DATA
ABILITY
SYSTEMS
MARKETING
FRANCHISE
SYSTEM
Health Care
0.0
0.5
1.0
1.5
2.0
RESEARCH
PATENT
INTELLECTUAL PROPERTY
MARKETING
DATA
ABILITY
EMPLOYEE
PATENTS
EMPLOYEES
TECHNOLOGY
Finance
0.0
0.5
1.0
1.5
EMPLOYEE
DATA
CONSUMER
ABILITY
EMPLOYEES
CUSTOMERS
CUSTOMER
EXPERIENCE
KNOWLEDGE
SYSTEM
Other
0.0
0.3
0.6
0.9
EMPLOYEE
EMPLOYEES
DATA
CUSTOMERS
ABILITY
CUSTOMER
SYSTEMS
MARKETING
TECHNOLOGY
SYSTEM

Table 4: Summary statistics for firms sorted on intangibles talk measure

	Intangibles Talk (All Categories)			Innovation assets & Information Technology		
	Low 30	Mid 40	High 30	Low 30	Mid 40	High 30
R&D to total expenses	0	0.02	0.12	0	0.01	0.13
SG&A to total expenses	0.24	0.23	0.38	0.27	0.21	0.4
Intangible capital$^{\mathrm{PT}}$ to total assets	0.06	0.43	0.73	0.05	0.42	0.76
Intangible capital$^{\mathrm{PT}}$ to sales	0.34	0.44	0.88	0.34	0.39	0.99
Sales to total assets	0.34	0.96	0.8	0.32	1.04	0.75
Sales to stockholder's equity	1.01	1.98	1.37	1.02	2.2	1.25
Price to sales	0.05	0.03	0.07	0.05	0.03	0.08
Debt to EBITDA	2.22	1.3	0.15	2.35	1.37	0.19
Debt to total assets	0.26	0.2	0.07	0.26	0.2	0.07
Profitability to total assets	0.09	0.28	0.35	0.08	0.29	0.33
Investment to physical capital	0.09	0.09	0.12	0.08	0.09	0.12
Market Cap (millions)	401.16	566.35	382.35	377.97	655.49	364.4
Book to market^{FF}	0.73	0.51	0.38	0.73	0.51	0.37
Book to market^{EKP}	1.2	1.21	1.12	1.23	1.22	1.07
Book to market^{PT}	0.85	0.73	0.63	0.87	0.73	0.6

	Brand & Customer Relations			Human Resources		
	Low 30	Mid 40	High 30	Low 30	Mid 40	High 30
R&D to total expenses	0.08	0.04	0.06	0.02	0.05	0.09
SG&A to total expenses	0.22	0.3	0.32	0.26	0.28	0.32
Intangible capital$^{\mathrm{PT}}$ to total assets	0.14	0.23	0.6	0.12	0.37	0.58
Intangible capital$^{\mathrm{PT}}$ to sales	0.41	0.53	0.62	0.43	0.53	0.64
Sales to total assets	0.45	0.68	0.95	0.53	0.73	0.85
Sales to stockholder's equity	1.1	1.3	1.75	1.21	1.43	1.57
Price to sales	0.05	0.06	0.05	0.05	0.05	0.05
Debt to EBITDA	1.46	1.05	0.58	1.64	1.11	0.22
Debt to total assets	0.21	0.18	0.11	0.23	0.18	0.09
Profitability to total assets	0.12	0.21	0.36	0.16	0.24	0.32
Investment to physical capital	0.09	0.09	0.11	0.09	0.1	0.11
Market Cap (millions)	447.57	387.42	480.18	425.67	401.57	473.28
Book to market^{FF}	0.58	0.61	0.46	0.66	0.56	0.44
Book to market^{EKP}	1.04	1.24	1.3	1.21	1.2	1.14
Book to market^{PT}	0.75	0.8	0.71	0.83	0.76	0.67

This table summarizes the characteristics of firms sorted by intangibles talk measure above 70th percentile, between 30th to 70th percentile, and the bottom 30th percentile. The values are the time-series average of the median firm characteristics within each bucket. The sample period is from January 1994 to December 2019. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019).

Table 5: Association between intangibles talk measure and other proxies for intangible intensity

	Dependent variable:					
	Intangibles Talk$_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
Book-to-Market$_t$	−0.176					0.006
	(−56.96)					(1.03)
$\frac{(SG\&A_t - R\&D_t)}{Total\ Expenses_t}$		0.299				0.255
		(88.05)				(40.63)
$\frac{R\&D_t}{Total\ Expenses_t}$			0.256			0.520
			(55.33)			(59.45)
$\frac{Intangible\ Capital_t^{PT}}{Total\ Assets_{t-1}}$				0.372		−0.512
				(112.43)		(−28.51)
$\frac{Intangible\ Capital_t^{EKP}}{Total\ Assets_{t-1}}$					0.370	0.525
					(108.74)	(29.98)
Observations	99,112	82,713	43,986	84,283	81,927	31,923
Adjusted R^2	0.032	0.086	0.065	0.130	0.126	0.218

In this table, we report the pooling regression with firm-level intangibles talk as the dependent variable:

$$\text{Intangibles Talk}_{i,t} = \boldsymbol{\alpha_i} + \boldsymbol{\beta} X_{i,t} + \epsilon_{i,t} \tag{7}$$

where $X_{i,t}$ are firm-level book-to-market, (SG&A - R&D)/Total Expense, R&D/Total Expense, intangible capital$^{\mathrm{PT}}$. The panel covers the period between January 1994 to December 2021. To separately capture the effect of only SG&A we use (SG&A - R&D) since R&D is already included in SG&A. All the variables are annual, with their negative values dropped, winsorized at 1% from above, and normalized by dividing by standard deviation. The error terms are clustered at the firm level and time fixed effects are accounted for. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019).

Table 6: Risk and returns from intangible talk measure versus intangibles-enhanced value strategies

		Full Sample (1995-2020)	Full Sample (Exc. Dot-com Bubble)	Dot-com Bubble (2000-2001)	1995-1999	2000-2007	2008-2020
		(1)	(2)	(3)	(4)	(5)	(6)
INT^{10K}	Ret	3.79	**5.05**	-10.69	10.78	-6.13	**7.63**
		(1.36)	(2.3)	(-0.85)	(1.52)	(-1.22)	(3.15)
	σ	12.13	9.31	29.24	9.86	17	8.26
	Sharpe	0.31	0.54	-0.37	1.09	-0.36	0.92
HML^{FF}	Ret	0.22	-2.18	**27.82**	-3.41	9.64	-4.51
		(0.07)	(-0.91)	(2.49)	(-0.44)	(1.66)	(-1.38)
	σ	11.13	9.44	21.75	10.12	12.33	10.4
	Sharpe	0.02	-0.23	1.28	-0.34	0.78	-0.43
HML^{PT}	Ret	2.01	0.06	**24.51**	-1.83	9.21	-1.21
		(0.81)	(0.03)	(2.65)	(-0.37)	(1.8)	(-0.42)
	σ	10.24	8.53	21.12	8.31	11.94	9.54
	Sharpe	0.2	0.01	1.16	-0.22	0.77	-0.13
HML^{EKP}	Ret	2.55	2.32	5.26	0.91	5.72	1.12
		(1.27)	(1.26)	(1.46)	(0.24)	(1.55)	(0.4)
	σ	8.56	8.22	12.01	7.39	9	8.67
	Sharpe	0.3	0.28	0.44	0.12	0.64	0.13

In this table, we summarize the risk and return associated with intangibles talk measure and other measures of intangible value documented in the literature. INT^{10K} is the portfolio sorted based on the intangible talk. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019). HML^{FF} is sorted based on traditional book-to-market value. HML^{PT} and HML^{EKP} are sorted based on intangible augmented book-to-market calculated using Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020) methods. Calculation of portfolio returns is described in Section3. The values in parentheses are Newey-West T-statistics that test the difference in the means from zero. The full sample is from July 1995 to June 2020. We exclude the years 2000 and 2001 in the sample, that is, the bursting of dot-com bubble. The returns are in percent per year (monthly return multiplied by twelve).

Table 7: Outperformance of intangible talk measure over other intangibles-based investment strategies

		Full Sample (1995-2020)	Full Sample (Exc. Dot-com Bubble)	Dot-com Bubble (2000-2001)	1995-1999	2000-2007	2008-2020
		(1)	(2)	(3)	(4)	(5)	(6)
INT^{10K}	Ret	3.79	**5.05**	-10.69	10.78	-6.13	**7.63**
		(1.36)	(2.3)	(-0.85)	(1.52)	(-1.22)	(3.15)
	σ	12.13	9.31	29.24	9.86	17	8.26
	Sharpe	0.31	0.54	-0.37	1.09	-0.36	0.92
$SG\&A^{portf.}$	Ret	0.53	1.44	-9.91	9.39	**-7.65**	2.58
		(0.25)	(0.75)	(-1.16)	(1.89)	(-2.52)	(1.17)
	σ	10.29	9.1	19.41	9.4	12.01	9.11
	Sharpe	0.05	0.16	-0.51	1	-0.64	0.28
$R\&D^{portf.}$	Ret	3.29	**4.92**	-15.39	13.58	-6.18	**5.65**
		(0.94)	(2.16)	(-0.92)	(1.61)	(-0.87)	(2.44)
	σ	14.43	11.32	33.79	12.72	20.21	9.56
	Sharpe	0.23	0.43	-0.46	1.07	-0.31	0.59
INT^{PT}	Ret	3.61	**4.52**	-6.44	6.36	-3.87	**7.62**
		(1.45)	(2.07)	(-0.59)	(0.77)	(-1.1)	(2.74)
	σ	11.09	9.35	22.94	11.83	12.81	9.45
	Sharpe	0.33	0.48	-0.28	0.54	-0.3	0.81
INT^{EKP}	Ret	3.44	**4.26**	-5.6	5.52	-3.07	**7.02**
		(1.62)	(2.12)	(-0.59)	(0.75)	(-1.04)	(2.67)
	σ	9.83	8.58	18.87	10.69	10.72	8.83
	Sharpe	0.35	0.5	-0.3	0.52	-0.29	0.79

In this table, we summarize the risk and return associated with textual intangible value and other measures of value and intangible value in the literature. SG&A and R&D portfolios are sorted based on (SG&A - R&D)/Total Expense and R&D/Total Expense. INT^{PT} and INT^{EKP} are sorted based only on the capitalized intangibles based, respectively in Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020) methods, both scaled by total assets. Calculation of portfolio returns is described in Section3. The values in parentheses are Newey-West T-statistics that test the difference in the means to be zero. The returns are monthly in percent per year (monthly return multiplied by twelve).

Table 8: Risk and returns from intangible talk measure for different subperiods and different intangible categories

Category		Full Sample (1995-2020) (1)	Full Sample (Exc. Dot-com Bubble) (2)	Dot-com Bubble (2000-2001) (3)	1995-1999 (4)	2000-2007 (5)	2008-2020 (6)
Innovation assets & Information Technology	Ret	3.52	**4.9**	-12.43	10.35	-4.77	**6.36**
		(1.16)	(2.19)	(-0.88)	(1.19)	(-0.81)	(2.5)
	σ	12.64	9.62	30.78	10.93	17.73	8.49
	Sharpe	0.28	0.51	-0.4	0.95	-0.27	0.75
Brand & Customer Relations	Ret	2.47	**3.69**	-11.57	10.84	-6.11	**4.95**
		(1.15)	(2.14)	(-1.15)	(1.58)	(-1.63)	(3.15)
	σ	9.89	7.55	23.91	7.88	13.77	6.85
	Sharpe	0.25	0.49	-0.48	1.38	-0.44	0.72
Human Resources	Ret	**2.47**	**2.71**	-0.28	**4.46**	-1.38	**4.21**
		(2.06)	(2.35)	(-0.06)	(2.4)	(-0.69)	(2.74)
	σ	6.29	5.71	11.15	4.92	8.08	5.3
	Sharpe	0.39	0.47	-0.03	0.91	-0.17	0.8

In this table, we summarize the risk and return associated with intangible talk measure, by three intangible categories. Calculation of intangibles talk measure is described in . Words used to calculate intangibles talk measure come from Filipovic and Wager (2019). Calculation of portfolio returns is described in Section3. The values in parentheses are Newey-West T-statistics that test the difference in the means to be zero. The full sample is from July 1995 to June 2020. We exclude the years 2000 and 2001 in the sample, excluding the dot-com bubble. The returns are monthly in percent per year (monthly return multiplied by twelve).

Table 9: Risk-adjusted future returns (alphas) from intangible talk measure.

	Dependent variable: INT^{10K}						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\alpha(\%)$	**3.369****	**5.810*****	**6.372*****	**8.346*****	**2.408*****	**2.831*****	**1.718***
	(2.351)	(4.028)	(4.310)	(4.098)	(3.013)	(3.352)	(1.871)
Mkt-RF	0.108**	0.033	0.055	−0.030	0.062***	0.066***	0.090***
	(2.380)	(0.904)	(1.381)	(−0.526)	(3.693)	(3.771)	(3.506)
SMB	0.153*	0.009	0.091	−0.027	−0.070	-0.073	0.142**
	(1.937)	(0.120)	(0.981)	(−0.354)	(−1.339)	(−1.351)	(2.477)
HML^{FF}	−0.825***	−0.651***			−0.034	−0.115***	−0.373***
	(−10.021)	(−9.046)			(−0.773)	(−2.600)	(−6.115)
HML^{PT}			−0.718***				
			(−8.661)				
HML^{EKP}				−0.220**			
				(−2.177)			
RMW		−0.426***	−0.385***	−0.656***	−0.324***	−0.415***	0.033
		(−4.646)	(−3.646)	(−6.538)	(−5.275)	(−6.773)	(0.402)
CMA		−0.050	−0.005	−0.485***	−0.117	−0.079	0.052
		(−0.467)	(−0.051)	(−2.881)	(−1.649)	(−1.120)	(0.674)
UMD	−0.098	−0.076	−0.041	−0.012	−0.053*	−0.055**	−0.071**
	(−1.503)	(−1.460)	(−1.045)	(−0.187)	(−1.913)	(−2.127)	(−2.476)
INT^{PT}					0.800***		
					(23.075)		
INT^{EKP}						0.797***	
						(20.431)	
$R\&D^{portf.}$							0.351***
							(6.920)
$SG\&A^{portf.}$							0.335***
							(6.124)
Observations	300	300	300	300	288	288	300
Adjusted R^2	0.622	0.681	0.669	0.533	0.861	0.853	0.790

In this table, we report portfolio alphas and betas by regressing the returns of INT^{10K} against factor models. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019). Calculation of portfolio returns (INT^{10K}) is described in Section3. Columns (1) and (2) use the Fama and French (2015) three and five factors and momentum factor from Carhart (1997). HML portfolios is based on traditional book-to-market. Columns (3) and (4) replace HML with its intangible augmented versions from Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020). Columns (5) and (6) include INT^{PT} and INT^{EKP} sorted based only on the capitalized intangibles based, respectively, in Peters and Taylor (2017) and Eisfeldt, Kim and Papanikolaou (2020) methods, both scaled by total assets. Column (7) includes portfolios sorted based on R&D/Total Expense and (SG&A - R&D)/Total Expense. We include Newey-West T-statistics. The sample is monthly from July 1995 to June 2020. All coefficients are reported in percentage per year (monthly returns multiplied by twelve).

Table 10: Risk-adjusted future returns (alphas) from intangible talk measure, by intangible categories

	Dependent variable: Categorical INT^{10K}					
Category:	Innovation assets &information technology	Band & customer relation	Human resources	Innovation assets & information technology	Band & customer relation	Human resources
$\alpha(\%)$	3.000**	2.279*	2.286**	5.298***	4.674***	2.619**
	(2.091)	(1.757)	(2.458)	(4.045)	(3.019)	(2.581)
Mkt-RF	0.109**	0.087**	0.039	0.039	0.007	0.034
	(2.498)	(1.973)	(1.576)	(1.177)	(0.168)	(1.387)
SMB	0.182***	0.042	0.064*	0.039	−0.068	0.012
	(2.921)	(0.814)	(1.661)	(0.625)	(−1.269)	(0.232)
HML^{FF}	−0.913***	−0.531***	−0.334***	−0.755***	−0.337***	−0.335***
	(−13.108)	(−4.835)	(−10.228)	(−13.933)	(−4.075)	(−7.020)
RMW				−0.419***	−0.346***	−0.135*
				(−4.992)	(−4.541)	(−1.766)
CMA				−0.016	−0.173	0.126
				(−0.166)	(−1.369)	(1.207)
UMD	−0.084	−0.090	−0.030	−0.064	−0.070	−0.027
	(−1.284)	(−1.234)	(−0.751)	(−1.223)	(−1.135)	(−0.741)
Adjusted R^2	0.704	0.389	0.379	0.759	0.453	0.417
obs	300	300	300	300	300	300

In this table, we report portfolio alphas and betas by regressing the returns from each category of INT^{10K} against factor models. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019). Calculation of portfolio returns (INT^{10K}) is described in Section3. Columns (1) through (3) use the Fama and French (2015) three factors and momentum factor from Carhart (1997). Columns (4) through (6) use the Fama and French (2015) five factors and momentum factor from Carhart (1997). We include Newey-West T-statistics. The sample is monthly from July 1995 to June 2020. All coefficients are reported in percentage per year (monthly returns multiplied by twelve).

Table 11: Risk-adjusted returns of portfolios sorted on intangible intensity indicators

Intangibles Talk →	Low	2	3	4	High	H-L
$\alpha(\%)$	**-3.09**	-0.2	0.09	0.79	**3.94**	**7.03**
	(-3.12)	(-0.19)	(0.11)	(1.08)	(3.65)	(4.18)
R&D to Total Expenses →	Low	2	3	4	High	H-L
$\alpha(\%)$	0.76	-2.51	0.37	-0.15	**4.81**	**4.04**
	(0.5)	(-1.48)	(0.3)	(-0.13)	(3.58)	(2.4)
SG&A to Total Expenses →	Low	2	3	4	High	H-L
$\alpha(\%)$	-1.71	-1.05	0.37	2.13	1.031	**2.54**
	(-1.02)	(-0.84)	(0.3)	(1.74)	(1.02)	(2.38)
Intangible Capital^{PT} to Total Assets →	Low	2	3	4	High	H-L
$\alpha(\%)$	6.01	6.62	**9.12**	**7.97**	**8.85**	2.84
	(1.35)	(1.85)	(2.51)	(2.48)	(2.53)	(0.94)
Intangible Capital^{EKP} to Total Assets →	Low	2	3	4	High	H-L
$\alpha(\%)$	6	**6.89**	**7.71**	**7.82**	**9.63**	3.63
	(1.36)	(1.97)	(2.06)	(2.41)	(2.91)	(1.35)

This table reports risk-adjusted returns (alphas) from five-factor Fama and French model and momentum factor from Carhart (1997). In each regression, the dependent variable is the excess returns (over risk-free rate) of value-weighted portfolios sorted on one intangible intensity at a time: intangibles talk, R&D and SG&A to total expenses, Intangible Capital^{PT} to total assets, and Intangible Capital^{EKP} to total assets. The regression model is:

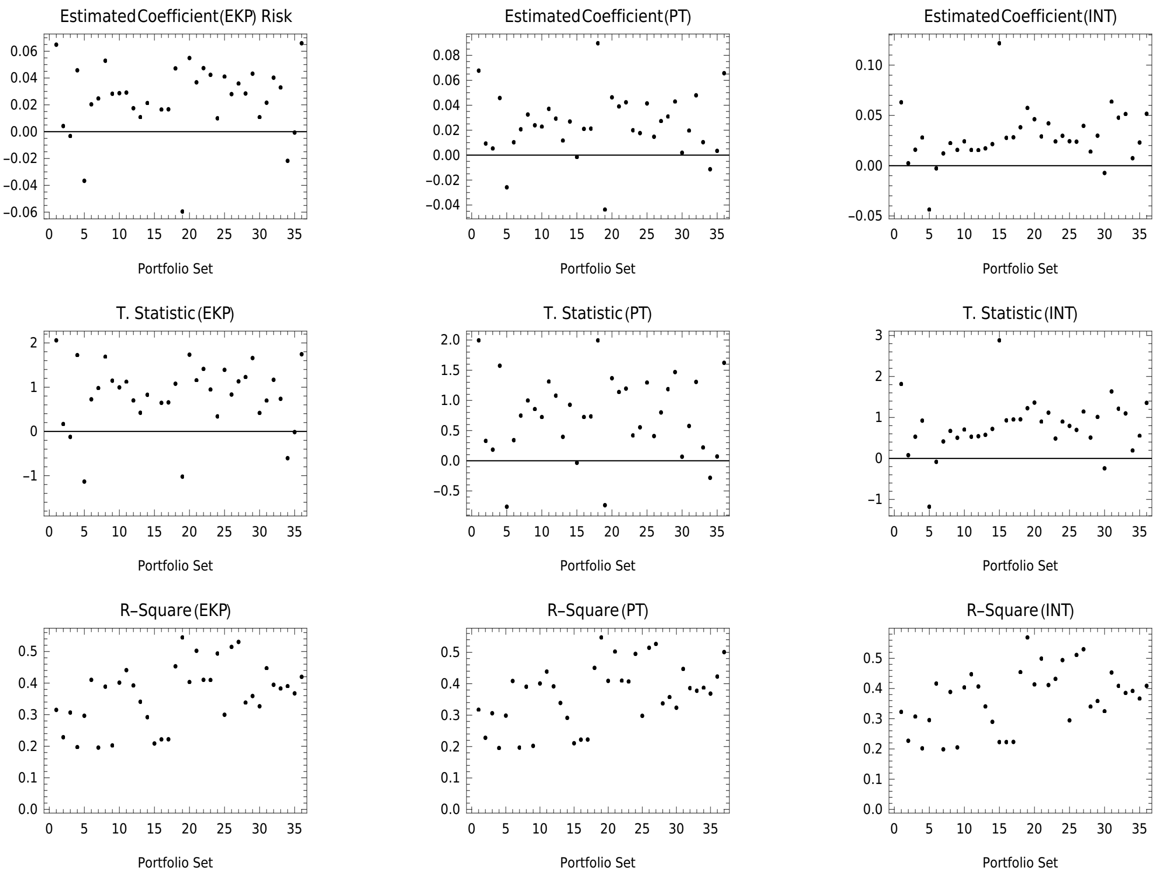
$$\mathrm{R}_{i,t} = \boldsymbol{\alpha_i} + \boldsymbol{\beta_{MKT}} MKT_t + \boldsymbol{\beta_{SMB}} SMB_t + \boldsymbol{\beta_{HML}} HML_t + \boldsymbol{\beta_{RMW}} RMW_t + \boldsymbol{\beta_{CMA}} \mathbf{CMA_t} + \boldsymbol{\beta_{UMD}} \mathbf{UMD_t} + \epsilon_{\mathbf{i,t}}$$

Where $R_{i,t}$ is the excess return (over risk-free rate) of a value-weighted portfolio in month t that is long in stocks belonging to the i^{th} (i from 1 to 5) quantile based on intangible intensity measures. Portfolio returns calculation is described in Section 3. Only the alphas from the regressions are reported for expositional purposes. We include Newey-West T-statistics in parentheses. The last column displays alphas from going long on highest quantile (5^{th}) and going short on the lowest quantile (1^{st}) of intangible-intensity measures. The sample is monthly from July 1995 to June 2020.

Table 12: Kenneth French's Sets of Portfolios (July 1995 - June 2020)

Number	Name
1	25 Portfolios Based on Size and Accruals
2	25 Portfolios Based on Book-to-Market and Investment
3	25 Portfolios based on Book-to-Market and Operating Profitability
4	25 Portfolios Based on Operating Profitability and Investment
5	100 Portfolios Based on Size and Book-to-Market
6	25 Portfolios Based on Size and Book-to-Market
7	100 Portfolios Based on Size and Investment
8	25 Portfolios Based on Size and Investment
9	100 Portfolios Based on Size and Operating Profitability
10	25 Portfolios Based on Size and Operating Profitability
11	10 Industry Portfolios
12	12 Industry Portfolios
	17 Industry Portfolios
14	30 Industry Portfolios
15	38 Industry Portfolios
16	48 Industry Portfolios
17	49 Industry Portfolios
18	10 Portfolios Based on Long-Term Reversal
19	10 Portfolios Based on Momentum
20	25 Portfolios based on Size and Long-Term Reversal
21	25 Portfolios Based on Size and Momentum
22	25 Portfolios Based on Size and Short-Term Reversal
23	10 Portfolios Based on Short Term Reversal
24	25 Portfolios Based on Size and Market Beta
25	25 Portfolios Based on Size and Net Share Issuance
26	25 Portfolios Based on Size and Residual Variance
27	25 Portfolios Based on Size and Variance
28	32 Portfolios Based on Book-to-Market, Investment, and Size
29	32 Portfolios Based on Book-to-Market, Operating Profitability and Size
30	32 Portfolios Based on Operating Profitability, Investment, and Size
31	10 Portfolios Based on Book-to-Market
32	10 Portfolios Based on Cash Flow Divided by Price
33	10 Portfolios Based on Dividend Yield
34	10 Portfolios Based on Earnings Divided by Price
35	10 Portfolios Based on Investment
36	10 Portfolios Based on Operating Profitability

Figure 7: Price of Intangibles Risk in Different Sets of Portfolios Created by Kenneth French



The price of intangibles risk is estimated from Fama-Macbeth two-stage regressions for each set of portfolios listed in Table12 and shown on the x-axis of each panel. Each month, risk loadings are estimated from the first-stage regression for portfolio p

$$\mathbf{R^p_{t-s}} = \alpha_{\mathbf{p,t}} + \beta^{\mathbf{MKT}}_{\mathbf{p,t}} \times \mathbf{MKT_{t-s}} + \delta_{\mathbf{p,t}} \times \mathbf{INT_{t-s-1}} + \beta^{\mathbf{SMB}}_{\mathbf{p,t}} \times \mathit{SMB}_{\mathbf{t-s}} \\ + \beta^{HML}_{p,t} \times \mathit{HML}_{t-s} + \beta^{UMD}_{p,t} \times \mathit{UMD}_{t-s} + \epsilon_{p,t-s}$$

for $p = 1, ..., N_k$, $s = 0, \cdots, 35$. In this equation, R^p_t is the monthly return of the p^{th} portfolio at t, N_k is the number of portfolios of stocks in portfolio set k, and INT_t is either the INT^{PT}, INT^{EKP}, or INT^{10K} factor. For the second-stage, each month we estimate the prices of risk at date t for portfolio set k using the equation

$$R^p_t = \kappa^k_t + \pi^k_t \times \beta^{MKT}_{p,t} + \omega^k_t \times \delta_{p,t} + \phi^{SMB,k}_t \times \beta^{SMB}_{p,t} + \phi^{HML,k}_t \times \beta^{HML}_{p,t} + \phi^{UMD,k}_t \times \beta^{UMD}_{p,t} + \epsilon_{t,p},$$

for $p = 1, ..., N_k$. We calculate the time-series mean of the estimated prices of risk and the estimated variance (our equations illustrate the case of the price of intangibles risk), for portfolio set k as

$$\hat{\omega}^k = \frac{1}{T} \sum_{t=1}^{T} \hat{\omega^k_t}; \qquad \sigma^2(\hat{\omega^k}) \frac{1}{T^2} \sum_{t=1}^{T} (\hat{\omega^k_t} - \hat{\omega^k})^2,$$

respectively. The top panels report the estimated price of risk for each intangible, the middle panels report the t-statistics of the coefficients, and the bottom panel reports the R^2 of the 2nd stage regression, for the INT^{PT}, INT^{EKP}, and INT^{10K} factors, respectively.

Table 13: Risk-adjusted future returns (alphas) from intangible talk measure, by twelve Fama and French industry classification

	Dependent variable: INT^{10K} *in 12 Fama–French industries*					
	NoDur	Dur	Manuf	Enrgy	Chems	BusEq
$\alpha(\%)$	−2.736	0.688	2.721	−2.183	−2.127	**6.079****
	(−1.221)	(0.226)	(1.473)	(−0.750)	(−0.716)	(2.530)
Mkt-RF	0.238***	−0.165**	−0.092*	0.046	−0.132	−0.100*
	(4.983)	(−2.464)	(−1.887)	(0.648)	(−1.384)	(−1.745)
SMB	0.101	−0.103	0.016	0.279**	−0.180**	−0.191**
	(1.143)	(−1.127)	(0.330)	(2.414)	(−2.041)	(−1.997)
HML	0.059	−0.176**	−0.101	−0.034	−0.174	−0.033
	(0.723)	(−2.113)	(−1.182)	(−0.248)	(−1.334)	(−0.396)
RMW	0.135	−0.268**	−0.264***	−0.118	−0.184	−0.327***
	(1.479)	(−2.460)	(−3.746)	(−1.274)	(−1.572)	(−3.065)
CMA	−0.097	−0.227	−0.115	−0.056	0.263	−0.324*
	(−0.731)	(−1.420)	(−0.843)	(−0.309)	(1.311)	(−1.841)
UMD	−0.001**	−0.00001	0.0001	−0.001	0.0004	0.0005
	(−2.145)	(−0.018)	(0.322)	(−1.449)	(0.727)	(0.975)
Observations	300	300	300	300	300	300
Adjusted R^2	0.134	0.059	0.134	0.082	0.049	0.151
	Telcm	Utils	Shops	Hlth	Finance	Other
$\alpha(\%)$	−2.679	0.070	0.809	**5.451***	**3.952*****	**5.346****
	(−0.680)	(0.054)	(0.455)	(1.872)	(3.548)	(2.489)
Mkt-RF	−0.133	−0.120	0.156***	0.003	0.010	−0.097*
	(−1.148)	(−1.439)	(2.766)	(0.057)	(0.275)	(−1.861)
SMB	−0.403**	−0.122	0.150**	0.055	0.064	−0.284***
	(−2.519)	(−1.305)	(2.311)	(0.518)	(1.193)	(−3.596)
HML	−0.427***	−0.047	−0.204***	−0.270***	−0.447***	−0.195*
	(−3.874)	(−0.520)	(−3.872)	(−2.687)	(−4.647)	(−1.839)
RMW	0.135	−0.268**	−0.264***	−0.118	−0.184	−0.327***
	(1.479)	(−2.460)	(−3.746)	(−1.274)	(−1.572)	(−3.065)
CMA	−0.097	−0.227	−0.115	−0.056	0.263	−0.324*
	(−0.731)	(−1.420)	(−0.843)	(−0.309)	(1.311)	(−1.841)
UMD	−0.001	0.0003	−0.001*	0.001*	−0.0002	0.002***
	(−0.393)	(1.236)	(−1.960)	(1.814)	(−0.858)	(3.059)
Observations	300	300	300	300	300	300
Adjusted R^2	0.144	0.060	0.192	0.317	0.319	0.262

In this table, we report portfolio alphas and betas by regressing the returns of INT^{10K} sorted using the stocks within each of the twelve industries based on Fama and French industry classifications. Calculation of intangibles talk measure is described in Section2. Words used to calculate intangibles talk measure come from Filipovic and Wager (2019). Calculation of portfolio returns (INT^{10K}) is described in Section3. In all the regressions, the specification includes Fama and French (2015) five factors and momentum factor from Carhart (1997). We include Newey-West T-statistics. The sample is monthly from July 1995 to June 2020. All coefficients are reported in percentage per year (monthly returns multiplied by twelve).

Table 14: Risk-adjusted future returns (alphas) for intangibles talk measure vs. technology stocks

	Dependent variable:					
	$\text{Tech}^{portf.}$				INT^{10K}	
	(1)	(2)	(3)	(4)	(5)	(6)
$\alpha(\%)$	2.113	**6.689*****	-1.234	1.474	**2.290****	**2.552****
	(0.994)	(3.380)	(-0.819)	(0.965)	(2.134)	(2.285)
Mkt_RF	0.272***	0.119***	0.165***	0.090***	-0.031	-0.025
	(6.420)	(2.787)	(5.417)	(2.790)	(-1.359)	(-1.057)
SMB	0.246***	0.043	0.094**	0.035	0.027	-0.012
	(4.342)	(0.761)	(2.313)	(0.817)	(0.918)	(-0.373)
HML	-1.055^{***}	-0.679^{***}	-0.236^{***}	-0.095	-0.285^{***}	-0.320^{***}
	(-18.706)	(-9.351)	(-3.838)	(-1.421)	(-6.789)	(-6.980)
RMW		-0.644^{***}		-0.261^{***}		-0.113^{**}
		(-8.360)		(-4.128)		(-2.381)
CMA		-0.362^{***}		-0.318^{***}		0.127**
		(-3.511)		(-4.095)		(2.176)
UMD	-0.171^{***}	-0.133^{***}	-0.074^{***}	-0.064^{**}	-0.010	-0.012
	(-4.551)	(-3.905)	(-2.748)	(-2.474)	(-0.516)	(-0.607)
INT^{10K}			0.993***	0.897***		
			(17.403)	(15.063)		
$\text{Tech}^{portf.}$					0.511***	0.487***
					(17.403)	(15.063)
Observations	300	300	300	300	300	300
Adjusted R^2	0.622	0.698	0.813	0.830	0.813	0.820

In this table, the portfolio $Tech^{portf.}$ is the value-weighted portfolio that goes long in stocks belonging to the Business Equipment sector based on the Fama and French twelve industry classifications and goes short stocks from other industries at the end of June each year. Columns(1) and (2) report the abnormal returns of $Tech^{portf.}$ against the Fama and French (2015) three and five factors and momentum factor from Carhart (1997). Columns (3) and (4) includes INT^{10K} in the specifications. Columns (5) and (6) report the abnormal returns of INT^{10K} against the Fama and French (2015) three and five factors and momentum factor from Carhart (1997) including $Tech^{portf.}$ in the specifications as well. We include Newey-West T-statistics. The sample is monthly from July 1995 to June 2020. All coefficients are reported in percentage per year (monthly returns multiplied by twelve).

Table 15: Risk-adjusted returns (Alphas) from portfolios double sorted on idiosyncratic Volatility (IVOL) and intangibles talk

Intangibles → Talk	Low	2	3	4	High	$[H-L]$	
Low IVOL	-1.83	1.38	0.61	0.67	1.32	3.16	(1.64)
2	-2.42	0.64	-0.54	1.98	1.86	4.29	(1.64)
3	-2.51	-3.16	1.19	-2.69	5.64	**8.15**	(3.23)
4	-8.78	-2.8	-2.26	0.06	6.07	**14.84**	(3.76)
High IVOL	-9.88	-9.43	1.06	2.2	4.34	**14.22**	(3.33)
$[H-L]$	**-8.05**	**-10.81**	0.45	1.52	3.01		
	(-2.22)	(-2.73)	(0.13)	(0.59)	(1.35)		

This table reports risk-adjusted returns (alphas) from five-factor Fama and French model. In each regression, the dependent variable is the excess returns (over risk-free rate) of value-weighted portfolios double sorted (5 by 5) on intangible talk measure and idiosyncratic volatility (IVOL).The regression model is:

$$\mathrm{R}_{i,t} = \boldsymbol{\alpha_i} + \boldsymbol{\beta_{MKT}} MKT_t + \boldsymbol{\beta_{SMB}} SMB_t + \boldsymbol{\beta_{HML}} HML_t + \boldsymbol{\beta_{RMW}} RMW_t + \boldsymbol{\beta_{CMA}} \mathbf{CMA_t} + \boldsymbol{\beta_{UMD}} \mathbf{UMD_t} + \epsilon_{\mathbf{i,t}}$$

Where $R_{i,t}$ is the excess return (over risk-free rate) of a value-weighted portfolio in month t that is long in stocks belonging to one of the 25 groups. Columns show results from lowest to highest quantile of intangible talk measure. Rows show results from lowest to highest quantile of IVOL. Portfolio returns calculation is described in Sections 3. We measure firm-level IVOL by estimating the volatility of the residuals from the three factors of Fama and French for daily returns. Only the alphas from the regressions are reported for expositional purposes. We include Newey-West T-statistics in parentheses. The last column displays alphas from going long on highest quantile (5^{th}) and going short on the lowest quantile (1^{st}) of intangibles talk measure for IVOL quantiles. The sample is monthly from July 1995 to June 2020.

Table 16: Risk-adjusted returns (Alphas) from portfolios double sorted on idiosyncratic Volatility (IVOL) and intangibles talk, by firm size

Small Firms

Intangibles → Talk	Low	2	3	4	High	$[H-L]$	
Low IVOL	-0.07	2.09	0.6	1.1	-0.61	-0.54	(-0.22)
2	1.91	1.51	0.15	-0.5	3.38	1.47	(0.76)
3	-0.61	0.17	-0.73	-0.65	2.97	3.58	(1.89)
4	-4.09	-0.61	-0.02	-1.2	3.36	**7.45**	(2.84)
High IVOL	-7.87	-5.41	-3.82	-1.41	4.02	**11.89**	(3)
$[H-L]$	**-7.8**	**-7.5**	-4.42	-2.51	4.63		
	(-2.13)	(-2.35)	(-1.5)	(-0.87)	(1.33)		

Big Firms

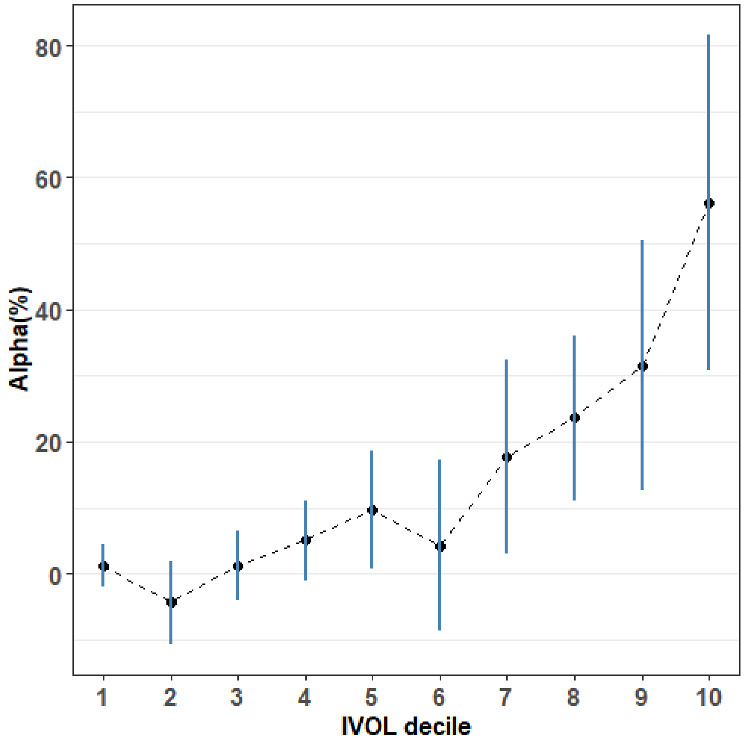
Intangibles → Talk	Low	2	3	4	High	$[H-L]$	
Low IVOL	-1.86	1.32	0.61	0.63	1.33	3.19	(1.64)
2	-2.62	0.81	-0.61	2.11	1.82	4.43	(1.63)
3	-2.54	-3.57	1.51	-3.1	5.79	**8.33**	(3.04)
4	-9.93	-3.94	-3.2	0.34	6.56	**16.50**	(3.75)
High IVOL	-13.07	-9.9	6.69	3.14	6.4	**19.47**	(3.28)
$[H-L]$	**-11.22**	**-11.22**	6.09	2.51	5.07		
	(-2.06)	(-2.23)	(1.21)	(0.55)	(1.76)		

This table reports Table 15 results by dividing sample at the median of firm size. This table reports risk-adjusted returns (alphas) from five-factor Fama and French model. In each regression, the dependent variable is the excess returns (over risk-free rate) of value-weighted portfolios double sorted (5 by 5) on intangible talk measure and idiosyncratic volatility (IVOL).The regression model is:

$$\mathrm{R}_{i,t} = \boldsymbol{\alpha_i} + \boldsymbol{\beta_{MKT}} MKT_t + \boldsymbol{\beta_{SMB}} SMB_t + \boldsymbol{\beta_{HML}} HML_t + \boldsymbol{\beta_{RMW}} RMW_t + \boldsymbol{\beta_{CMA}} \mathbf{CMA_t} + \boldsymbol{\beta_{UMD}} \mathbf{UMD_t} + \epsilon_{\mathbf{i,t}}$$

Where $R_{i,t}$ is the excess return (over risk-free rate) of a value-weighted portfolio in month t that is long in stocks belonging to one of the 25 groups. Columns show results from lowest to highest quantile of intangible talk measure. Rows show results from lowest to highest quantile of IVOL. Portfolio returns calculation is described in Sections 3. We measure firm-level IVOL by estimating the volatility of the residuals from the three factors of Fama and French for daily returns. Only the alphas from the regressions are reported for expositional purposes. We include Newey-West T-statistics in parentheses. The last column displays alphas from going long on highest quantile and going short on the lowest quantile of intangibles talk measure for IVOL quantiles. The last column displays alphas from going long on highest quantile (5^{th}) and going short on the lowest quantile (1^{st}) of intangibles talk measure for IVOL quantiles. The sample is monthly from July 1995 to June 2020.

Figure 8: Risk-adjusted return (Alphas) from portfolios sorted on intangibles talk measure, by decile of idiosyncratic volatility (IVOL)



The table above plots the Fama-French five factor alpha of ten INT^{10K} portfolios by IVOL decile. Every year stocks are sorted into deciles based on their IVOL and INT^{10K} is constructed separately using the stocks in each decile. Calculation of portfolio returns (INT^{10K}) is described in Section3. We measure firm-level IVOL by estimating the volatility of the residuals from the three factors of Fama and French for daily returns. Alphas are estimated using annual returns (monthly returns multiplied by twelve), based on Fama and French regressions discussed in Section4. The blue line represents the (μ ± 1.96 × NeweyWest standard errors). The sample is monthly, covering the period between July 1995 to June 2020.

Appendix

Excerpts from example Form 10-K filings containing intangibles terms:

"Our license and development arrangements with customers typically require significant customization of our ***intellectual property*** *components. As a result, we recognize the revenue from the* ***license*** *and the revenue from the development services as a single performance obligation over the period in which the development services are performed. We measure progress to completion based on actual cost incurred to date as a percentage of the estimated total cost required to complete each project. If a loss on an arrangement becomes probable during a period, we record a provision for such loss in that period." (NVIDIA CORPORATION, Form 10-K, January 26, 2020)*

"We purchase and roast high-quality whole bean coffees that we sell, along with handcrafted coffee and tea beverages and a variety of fresh food items, through company-operated stores. We also sell a variety of coffee and tea products and ***license*** *our* ***trademarks*** *through other channels such as licensed stores, grocery stores, and national food service accounts. In addition to our flagship Starbucks* ***brand****, our portfolio also includes Tazo® Tea, Seattle's Best Coffee®, and Starbucks VIA® Ready Brew." (Starbucks Corporation, Form 10-K, October 2, 2011)*

"At FedEx, it is our people—our greatest asset—that give us our strong ***reputation*** *and stand at the heart of our success. In addition to our superior physical and information* ***networks****, FedEx has an exemplary* ***human network****. Across the globe, our team members are united by our passion to deliver the FedEx Purple Promise—to make every FedEx experience outstanding—and our People–Service–Profit principles. " (FedEx Corporation , Form 10-K, May 31, 2022)*

"We have obtained ***patents*** *in the U.S. and other countries. Because of the fast pace of* ***innovation*** *and product development, and the comparative pace of governments' patenting processes, our products are often obsolete before the* ***patents*** *related to them expire; in some cases, our products may be obsolete before the* ***patents*** *related to them are granted. As we expand our products into new industries, we also seek to extend our patent development efforts to patent such products. " (INTEL Corporation, Form 10-K, December 26, 2015)*

"We also connect consumers with public transportation networks. We use this same ***network, technology****, operational excellence and product expertise to connect shippers with carriers in the freight industry. " (UBER TECHNOLOGIES, INC., Form 10-K, December 31, 2020)*

References

Aboody, David. 1996. “Recognition versus disclosure in the oil and gas industry.” *Journal of Accounting Research* 34:21–32.

Aboody, David and Baruch Lev. 1998. “The value relevance of intangibles: The case of software capitalization.” *Journal of Accounting Research* 36:161–191.

Ahmed, Anwer S, Emre Kilic and Gerald J Lobo. 2006. “Does recognition versus disclosure matter? Evidence from value-relevance of banks’ recognized and disclosed derivative financial instruments.” *Accounting Review* 81(3):567–588.

Ang, Andrew, Robert J Hodrick, Yuhang Xing and Xiaoyan Zhang. 2006. “The cross-section of volatility and expected returns.” *Journal of Finance* 61(1):259–299.

Arnott, Robert D, Campbell R Harvey, Vitali Kalesnik and Juhani T Linnainmaa. 2021. “Reports of value’s death may be greatly exaggerated.” *Financial Analysts Journal* 77(1):44–67.

Asness, Clifford S, R Burt Porter and Ross L Stevens. 2000. “Predicting stock returns using industry-relative firm characteristics.” *Social Science Research Network.* `"https://papers.ssrn.com/sol3/papers.cfm?abstract_id=213872"`.

Belo, F, V Gala, J Salomao and MA Vitorino. 2022. “Decomposing firm value (0898-2937).” *Journal of Fnancial Economics 143 (2), 619–639* .

Birru, Justin and Trevor Young. 2020. “Sentiment and uncertainty.” *Fisher College of Business Working Paper* (2020-03):010.

Cao, Jie and Bing Han. 2016. “Idiosyncratic risk, costly arbitrage, and the cross-section of stock returns.” *Journal of Banking and Finance* 73:1–15.

Carhart, Mark M. 1997. “On persistence in mutual fund performance.” *Journal of Finance* 52(1):57–82.

Chan, Louis KC, Josef Lakonishok and Theodore Sougiannis. 2001. “The stock market valuation of research and development expenditures.” *Journal of finance* 56(6):2431–2456.

Choi, Ki-Soon, Eric C So and Charles CY Wang. 2021. “Going by the Book: Valuation Ratios and Stock Returns.” *Social Science Network.* `"https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3854022."`.

Cochrane, JH. 2001. “Asset Pricing. Princeton University Press, Princeton, NJ.”.

Corrado, Carol, Charles Hulten and Daniel Sichel. 2005. Measuring capital and technology: An expanded framework. In *Measuring capital in the new economy*. University of Chicago Press pp. 11–46.

Curtis, Asher, Sarah McVay and Sara Toynbee. 2020. “The changing implications of research and development expenditures for future profitability.” *Review of Accounting Studies* 25:405–437.

Daniel, Kent and Sheridan Titman. 2006. "Market reactions to tangible and intangible information." *Journal of Finance* 61(4):1605–1643.

Davis-Friday, Paquita Y, L Buky Folami, Chao-Shin Liu and H Fred Mittelstaedt. 1999. "The value relevance of financial statement recognition vs. disclosure: Evidence from SFAS No. 106." *Accounting Review* 74(4):403–423.

Deng, Zhen, Baruch Lev and Francis Narin. 1999. "Science and technology as predictors of stock performance." *Financial Analysts Journal* 55(3):20–32.

Edmans, Alex. 2011. "Does the stock market fully value intangibles? Employee satisfaction and equity prices." *Journal of Financial Economics* 101(3):621–640.

Edmans, Alex, Lucius Li and Chendi Zhang. 2014. Employee satisfaction, labor market flexibility, and stock returns around the world. Technical report National Bureau of Economic Research, Cambridge, MA.

Eisfeldt, Andrea L and Dimitris Papanikolaou. 2013. "Organization capital and the cross-section of expected returns." *Journal of Finance* 68(4):1365–1406.

Eisfeldt, Andrea L, Edward Kim and Dimitris Papanikolaou. 2020. Intangible value. Technical report National Bureau of Economic Research, Cambridge, MA.

Enache, Luminita and Anup Srivastava. 2018. "Should intangible investments be reported separately or commingled with operating expenses? New evidence." *Management Science* 64(7):3446–3468.

Ewens, Michael, Ryan H Peters and Sean Wang. 2019. Measuring intangible capital with market prices. Technical report National Bureau of Economic Research, Cambridge, MA.

Falato, Antonio, Dalida Kadyrzhanova, Jae Sim and Roberto Steri. 2022. "Rising intangible capital, shrinking debt capacity, and the US corporate savings glut." *Journal of Finance* 77(5):2799–2852.

Fama, Eugene F and James D MacBeth. 1973. "Risk, return, and equilibrium: Empirical tests." *Journal of Political Economy* 81(3):607–636.

Fama, Eugene F and Kenneth R French. 2015. "A five-factor asset pricing model." *Journal of Financial Economics* 116(1):1–22.

Filipovic, Zoran and Alexander F Wager. 2019. "The intangibles song in takeover announcements: Good tempo, Hollow Tune." *Social Science Research Network*. "https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3330623.".

Garcia, Diego. 2013. "Sentiment during recessions." *Journal of Finance* 68(3):1267–1300.

Hall, Richard. 2009. The strategic analysis of intangible resources. In *Knowledge and strategy*. Routledge, London, England, pp. 181–195.

Hulten, Charles R and Xiaohui Hao. 2008. What is a Company Really Worth? Intangible Capital and the" Market to Book Value" Puzzle. Technical report National Bureau of Economic Research, Cambridge MA.

Iqbal, Aneel, Shivaram Rajgopal, Anup Srivastava and Rong Zhao. 2023. "Value of internally generated intangible capital." *Management Science (forthcoming)* .

Jiang, Fuwei, Joshua Lee, Xiumin Martin and Guofu Zhou. 2019. "Manager sentiment and stock returns." *Journal of Financial Economics* 132(1):126–149.

Lev, Baruch. 2000. *Intangibles: Management, Measurement, and Reporting*. Brookings Institution Press, Washington, DC.

Lev, Baruch. 2005. "Intangible assets: Concepts and measurements. In: Kempf-Leonard, K. (Ed.), Encyclopedia of Social Management." 2.

Lev, Baruch. 2012. *Winning Investors Over: Surprising Truths about Honesty, Earnings Guidance, and Other Ways to Boost Your Stock Price*. Harvard Business Review Press, Boston, MA.

Lev, Baruch and Anup Srivastava. 2022. "Explaining the recent failure of value investing." *Critical Finance Review* 11(2):333–360.

Lev, Baruch and Theodore Sougiannis. 1996. "The capitalization, amortization, and value-relevance of R&D." *Journal of Accounting and Economics* 21(1):107–138.

Lewellen, Jonathan, Stefan Nagel and Jay Shanken. 2010. "A skeptical appraisal of asset pricing tests." *Journal of Financial Economics* 96(2):175–194.

Loughran, Tim and Bill McDonald. 2011. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." *Journal of Finance* 66(1):35–65.

Mandelbrot, Benoit. 1961. "On the theory of word frequencies and on related Markovian models of discourse. In:Structure of language and its mathematical aspects." *American Mathematical Society, Providence* 12:190–219.

Merkley, Kenneth J. 2014. "Narrative disclosure and earnings performance: Evidence from R&D disclosures." *The Accounting Review* 89(2):725–757.

Newey, Whitney K and Kenneth D West. 1987. "Hypothesis testing with efficient method of moments estimation." *International Economic Review* pp. 777–787.

Park, Hyuna. 2019. "An intangible-adjusted book-to-market ratio still predicts stock returns." *Critical Finance Review* 25(1):207–236.

Peters, Ryan H and Lucian A Taylor. 2017. "Intangible capital and the investment-q relation." *Journal of Financial Economics* 123(2):251–272.

Pontiff, Jeffrey. 1996. "Costly arbitrage: Evidence from closed-end funds." *Quarterly Journal of Economics* 111(4):1135–1151.

Seamons, Quinton F and Robert W Rouse. 1997. "MD&A: How to use safe harbor to report "soft" information." *Journal of Corporate Accounting and Finance* 9(1):81–90.

Srivastava, Anup. 2014. "Why have measures of earnings quality changed over time?" *Journal of Accounting and Economics* 57(2-3):196–217.

Srivastava, Anup. 2023. "Trivialization of the bottom line and losing relevance of losses." *Review of Accounting Studies (forthcoming)* .

Stambaugh, Robert F, Jianfeng Yu and Yu Yuan. 2012. "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics* 104(2):288–302.

Stambaugh, Robert F, Jianfeng Yu and Yu Yuan. 2015. "Arbitrage asymmetry and the idiosyncratic volatility puzzle." *Journal of Finance* 70(5):1903–1948.

Tetlock, Paul C. 2007. "Giving content to investor sentiment: The role of media in the stock market." *Journal of Finance* 62(3):1139–1168.

Wyatt, Anne. 2008. "What financial and non-financial information on intangibles is value-relevant? A review of the evidence." *Accounting and Business Research* 38(3):217–256.