Expected Returns Based on Corporate Risk Disclosure

Allen Huang
Hong Kong University of Science and Technology

Rongyang Ma
Tsinghua University

Haifeng You*
Tsinghua University

Abstract

This study examines the efficacy of risk factor disclosures in corporate annual reports for estimating expected returns. Using BERTopic, a neural network-based topic modeling algorithm, we extract quantifiable risk exposure vectors from the textual risk factor disclosures. We find that peer firms identified based on their proximity in the exposure vectors exhibit a greater explanatory power for variations in stock returns, valuations, profitability, and other financial ratios compared to traditional industry classifications, suggesting that these vectors capture firms' exposures to underlying economic shocks. We then construct an expected return proxy using principal component regressions based on these disclosure-based risk exposures. Our results show that this novel risk-based proxy outperforms state-of-the-art, characteristics-based expected return proxies in terms of correlation with future stock returns and cross-sectional measurement error variance. This research demonstrates the potential of applying sophisticated machine learning techniques to corporate disclosures for enhancing our understanding of firm risk and expected returns.

Keywords: risk factor disclosures, peer firms, risk exposure vectors, expected return proxy, BERTopic, principal component regression

JEL Codes: G12, G32, M41

* Corresponding author, email: hyou@sem.tsinghua.edu.cn. We thank participants at China Journal of Accounting Research 2024 Annual Conference, and the 2nd Joint Conference on Statistics and Data Science in China (JCSDS) for helpful comments. All errors are our own.

1. Introduction

Classic financial theory posits that expected returns should compensate for risk. The Capital Asset Pricing Model (CAPM), for example, contends that CAPM beta represents the sole source of systematic risk, thereby dictating expected returns. However, prevailing research (e.g. Fama and French 1996) has largely discredited the CAPM by demonstrating that it cannot explain cross-sectional variation in stock returns. Subsequent studies have identified a range of firm characteristics that predict stock returns (Cochrane, 2011; Harvey, Liu, and Zhu, 2016; Hou, Xue, and Zhang, 2020), but these characteristics are often categorized as "anomalies," leaving it unclear whether the associated premia truly reflect risk compensation. This paper introduces a novel approach to estimating expected returns based on the textual disclosure of significant risk factors in corporate annual reports (10-K filings). We develop a risk-based expected return proxy (ERP) and test its performance against state-of-the-art proxies in the existing literature.

Section 1A of the Form 10-K, mandated by the SEC under the heading "Risk Factors," requires companies to list all material risks that could impact their operations significantly. Previous research indicates that these disclosures are substantive reflections of the risks companies face and provide valuable information to investors (Kravet and Muslu, 2013; Campbell et al., 2014; Hope, Hu, and Lu, 2016). Unlike firm characteristics inferred from financial statements and market variables, these disclosures offer clear economic insights, encompassing a range of risk types such as operational, financial, regulatory, environmental, and social.

We employ a deep learning-based topic modeling algorithm, BERTopic, to analyze these risk disclosures and transform them into quantifiable risk exposure metrics.

Specifically, we categorize each paragraph into a distinct topic, corresponding to a specific risk factor. We then create a risk factor exposure vector representing the proportion of text in Section 1A dedicated to each identified topic. If our metric effectively captures firms' exposure to potential economic shocks, we would expect that firms with similar risk profiles exhibit higher stock price comovement and fundamental correlation. To confirm this intuition, we utilize cosine similarity of the risk factor exposure vectors to assess similarity, finding that firms with more similar risk exposures are more likely to belong to the same industry. Moreover, we find that risk-based peers (RBPs)—defined as the ten firms with the highest cosine similarity in risk exposure to a focal firm within the same year— provide a superior explanation of the cross-sectional variations in stock returns, valuation, profitability, and other fundamental metrics compared to traditional industry peers. These findings validate that the risk factor vector derived from Section 1A of the 10-K filings effectively captures firms' underlying risk exposures.

Next, we proceed to construct a proxy for expected stock returns using the disclosure-based risk exposures. To do so, we first employ a Principal Component Analysis (PCA) to reduce the dimensionality of the risk factor exposures extracted from the prior year's Item 1A disclosure. We retain the first five principal components (PCs), standardize them cross-sectionally to have a mean of zero and a standard deviation of one. These PCs represent firms' exposure or loading to the corresponding latent risk factors. We then estimate a cross-sectional regression of monthly stock returns on the standardized PCs. The resulting slope coefficients capture the return to the corresponding risk or factor return/premium (Fama and

French, 2020).[1] The fitted value of the regression, which equals the sum of the products of pure factor returns and the factor loading, represents firms' expected returns that are attributable to their disclosed risk factors. To mitigate the effect of noise, we define the risk-based expected return proxy (ERP) as the moving average of the monthly fitted values over the past 12 months.

We benchmark our risk-based ERP against several characteristics-based ERPs developed by recent literature.[2] Specifically, we consider three characteristics-based ERPs: (1) JLR, as in Lee et al. (2021), based on Lewellen (2015), which is estimated using three firm characteristics: the log of market capitalization (size), book-to-market ratio (bm), and the cumulative stock returns from 12 months to 2 months prior to the current month; (2) LPV, as in Lyle and Wang (2015) and Chattopadhyay et al. (2022), which is based on bm, gross return on equity (roe), and the standard deviation of prior-month daily returns (vol); and (3) FFC, as in Fama and French (2020), which is based on size, bm, operating profitability (op), and the rate of growth of assets (inv). When estimating ERPs using these characteristics models, we follow these studies by taking the average of the coefficients (weights) of the historical Fama and MacBeth regressions over the past 36 months and then applying the average coefficients to contemporaneous firm characteristics.

[1] MSCI uses this approach to estimate factor return in its Barra risk model. It can be shown that the slope coefficient on a particular risk factor in the cross-sectional regression represents a return to a zero-investment long-short portfolio with a unit, i.e., one, exposure to the corresponding factor, and zero exposure to all other factors (characteristics) that are included in the regression. Due to this appealing feature, these coefficients are often called pure factor returns (e.g. Menchero 2010; Menchero, Orr and Wang 2011).

[2] Recent studies have demonstrated that characteristics-based expected return proxies (ERPs) outperform other commonly used alternatives, such as Implied Cost of Capital (ICCs) and factor-based ERPs (e.g., Lyle and Wang, 2015; Fama and French, 2020; Lee, So, and Wang, 2021; Chattopadhyay, Lyle, and Wang, 2022). However, most of the characteristics used in these studies are empirically motivated, such as book-to-market, profitability, and stock return momentum. Therefore, it remains unclear whether the expected return associated with these characteristics is attributable to firm risk.

To evaluate the efficacy of our risk-based ERP, we compare its performance to characteristics-based ERPs across several dimensions: (1) the slope coefficient from cross-sectional regressions of one-month-ahead realized returns on each ERP (Lyle and Wang 2015; Chattopadhyay et al. 2022); (2) the magnitude of return spread between the extreme deciles of these ERPs, providing insight into the economic significance of the cross-sectional relationships; and (3) cross-sectional measurement error variances (MEV) following the methodology of Lee et al. (2021).[3]

Our empirical analyses reveal that despite being derived from an unsupervised machine learning algorithm blind to stock returns, our disclosure-based risk factor exposure yields an ERP with significant explanatory power for future realized stock returns.[4] This positive association contrasts with the often-negative relationship observed between future realized stock returns and ICC and factor-based ERPs (e.g., Chattopadhyay et al. 2022). Notably, while the benchmark characteristics-based ERPs demonstrate a robust positive association with future stock returns before 2008 (Lyle and Wang 2015; Chattopadhyay et al. 2022), this relationship weakens substantially and often become insignificant after 2008.

Examining the realized return spread between extreme deciles, our risk-based ERP generates significant monthly hedge returns of 1.15% and 0.9% on an equal- and value-weighted basis, respectively. In contrast, none of the characteristics-based ERP strategies yield significant hedge returns after 2008. Furthermore, the risk-based ERP consistently

[3] Lee et al. (2021) also study the time-series measurement error variance of various ERPs. Since our disclosure risk exposures are standardized cross-sectionally, it lacks the ability to predict time-series variation in expected returns. We therefore refrain from comparing the time-series MEV of the alternative ERPs.

[4] In contrast, characteristics-based ERPs are based on financial ratios that prior literature shows to be predictive of future stock returns (e.g. size, bm, gross profitability, return momentum). These results were initially considered as stock market anomalies before they were included in various empirical asset pricing models.

performs at least as well as, and often surpasses, characteristics-based proxies in terms of cross-sectional MEV.

Our results demonstrate robustness across various subgroup sizes and subsample periods. Additionally, we conduct a range of supplementary analyses using alternative specifications, which yield qualitatively similar findings. Specifically, our results hold when evaluating ERP performance using stock returns at months t+2 and t+3 (instead of t+1), or when using 9- and 6-month moving averages (instead of 12 months) to compute the risk-based ERP. We also find that reducing the dimension of the risk factor exposure vector to 3, 10, 20, or 30 using PCA (instead of five), and re-estimating ERP does not alter our conclusions.

Building on the findings of Chattopadhyay et al. (2022) that a set of over 70 characteristics do not systematically improve the predictive ability of LPV ERPs, we investigate the incremental usefulness of disclosure-based risk factor exposures. Our analysis reveals that incorporating these exposures significantly enhances the explanatory power of the resulting ERP measures with respect to future stock returns. Moreover, decomposition analysis indicates that the component attributable to disclosure-based risk factors exhibits stronger and more robust predictive power than the component attributable to financial characteristics.

Our paper makes several contributions to the literature. First, we add to the extensive research on ERPs by developing a novel ERP based on risk factors explicitly disclosed by firms.[5] These factors are directly linked to firms' fundamental risk, making them conceptually well-suited to be determinants of expected returns. We demonstrate that our risk-based ERP performs at least as well as, and often surpasses, state-of-the-art characteristics-based ERPs in explaining realized returns. As a result, our measure has the potential to benefit not only researchers but also practitioners including financial analysts and investors.

Second, we contribute to the literature on the decision usefulness of risk factor disclosure. Prior studies have shown that investors react to risk factor disclosure, indicating that it provides value-relevant information about firms' underlying risks (Kravet and Muslu 2013; Campbell et al. 2014; Hope, Hu, and Lu 2016; Chiu, Guan, and Kim 2018). We extend this literature by demonstrating that mandatory risk factor disclosure facilitates the important task of estimating expected returns.

Third, we contribute to the growing literature applying machine learning technology in accounting and finance. While most studies focus on predicting earnings and stock returns (e.g., Gu et al. 2020; Chen, Cho, Dou, and Lev 2022; Cao and You 2024), estimating pricing kernels (e.g., Kelly, Pruitt, and Su 2019; Kozak, Nagel, and Santosh 2020; Chen, Pelger, and Zhu 2024), or analyzing sentiments (Huang et al. 2022), we apply an unsupervised neural topic modeling machine learning algorithm to construct expected returns. Our results show that this approach is a promising tool for converting unstructured textual disclosures into

[5] In addition to the studies focusing on the development and evaluation of various ERPs measures (e.g., Claus and Thomas 2001; Gebhardt et al. 2001; Gode and Mohanram 2003; Easton 2004; Easton and Monahan 2005; Botosan, Plumlee and Wen 2011; Hou et al. 2012; Li and Mohanram 2014; Lyle and Wang 2015; Lee et al. 2021; Chattopadhyay et al. 2022), there is a literature examining factors that affect ICCs and expected returns. Interested readers may refer to Chattopadhyay et al. 2022 for a comprehensive review of the literature on the use of ICCs in accounting and finance research.

numerical risk factor exposures, which could be of significant utility to investors and researchers.

The remainder of the paper are structured as follows: Section II reviews the related literature. Section III describes our research methodology. Section IV presents the empirical results. Section V concludes.

2. Related literature

2.1. Literature on expected return proxies

ERPs play a crucial role in various finance applications, including portfolio management, risk assessment, and corporate investment and finance decisions, as they provide valuable insights into the expected future performance of assets and inform investors' decision-making. The theoretical underpinning of ERPs is rooted in the CAPM and its extensions, which posit that expected returns should be a function of systematic risk, as measured by beta. However, the empirical performance of CAPM has been mixed at best (e.g. Fama and French, 1996). In response, the literature has shifted its focus to alternative approaches for estimating ERPs, which can be roughly categorized into the following three groups: (1) implied cost of capital proxies, (2) factor-based proxies, and (3) characteristics-based proxies.

Implied Cost of Capital Proxies

ICC proxies, developed in the accounting literature (e.g., Botosan, 1997; Gebhardt, Lee, and Swaminathan, 2001), estimate expected returns by equating a firm's market value to the present value of its expected future cash flows, thereby providing a market-based estimate of the firm's cost of equity. The ICC estimation methodology has evolved over time, with various valuation models being employed to calculate ICCs (Claus and Thomas, 2001;

Gebhardt et al., 2001; Gode and Mohanram, 2003; Easton, 2005). However, the reliance on analysts' earnings forecasts has limited the applicability of analyst-forecast-based ICCs to firms with analyst coverage. To address this limitation, subsequent studies have proposed the use of model-based earnings forecasts, which not only expands the scope of ICC estimation but also results in measures that outperform analyst-based ICCs in certain contexts (Hou, Van Dijk, and Zhang, 2012; Li and Mohanram, 2014).

Factor-Based Proxies

Factor-based proxies estimate ERPs by employing asset pricing models that link expected returns to a range of factor premiums. A prime example of this approach is the Fama-French three-factor model (Fama and French, 1993) and its subsequent extensions, such as the Carhart four-factor model (Carhart, 1997) and the five-factor model (Fama and French, 2015). Another notable example is the *q*-factor model developed by Hou, Xue, and Zhang (2015), which incorporates an investment factor. The estimation of factor-based ERPs involves calculating the sum of the products of their respective factor premiums and corresponding factor sensitivities. Factor premiums are frequently determined as the hedge returns to characteristic-sorted portfolios, such as the SMB (small minus big) and HML (high minus low) portfolios introduced by Fama and French (1993). Factor sensitivities, on the other hand, can be estimated through time-series regressions of securities' excess returns on the aforementioned factor premiums.

Characteristics-Based Proxies

Characteristics-based proxies offer a practical and intuitive approach to estimating expected returns by leveraging firm-specific characteristics, such as size, book-to-market ratio, and momentum. This approach has gained considerable traction in the finance literature due to its simplicity and ease of implementation. A notable example is the study by Lewellen (2015), which employs a linear combination of firm characteristics, including size and book-to-market ratio, to estimate expected returns. The weights assigned to each characteristic are derived from cross-sectional regressions of historical stock returns on these characteristics. Further extending this methodology, Lyle and Wang (2015) adopt a log-linear present-value framework to enhance the estimation of expected returns based on firm characteristics.

Empirical performance of the ERPs

Despite their widespread adoption, empirical evidence indicates that many ERPs perform poorly. Notably, ICC proxies are often compromised by inaccurate and biased forecasts of future fundamentals (e.g., Chan, Karceski, and Lakonishok, 2003). The noise and bias in these forecasts significantly impair ICC proxies, rendering them uninformative or even biased. Consistent with this assertion, Easton and Monahan (2005) examine seven accounting-based ICC proxies and find that none of them exhibit a positive association with realized returns in the cross-section. Chattopadhyay et al. (2022) report similar findings in an international context, underscoring the limitations of ICC-based ERP estimates across diverse markets. Additionally, Lee et al. (2021) demonstrate that ICCs are unreliable in the cross-section and fail to outperform a trivial ERP that assumes a uniform expected return across all firms.

Factor-based ERPs also exhibit underwhelming performance, likely due to the difficulties in estimating firm-level factor loadings and using past returns to estimate expected factor premiums (Fama and French, 2020). Chattopadhyay et al. (2022) find that factor-based ERPs using Fama and French’s three- and five-factor models have negative associations with realized stock returns in most of the 29 countries in their sample, including the U.S. Furthermore, Lee et al. (2021) show that factor-based ERPs perform poorly in estimating treatment effects in the cross-section, with significantly positive cross-sectional mean error variance.

Both Lee et al. (2021) and Chattopadhyay et al. (2022) document that characteristics-based ERPs outperform ICC- and factor-based ERPs, demonstrating better performance in predicting cross-sectional variations of realized returns and estimating treatment effects in the cross-section. Chattopadhyay et al. (2022) estimate several characteristics-based ERPs (LPV ERPs) using the log-linear present-value relation between a firm's current stock price, an accounting valuation anchor, and its expected future growth (Lyle and Wang, 2015; Chattopadhyay et al., 2022). Their findings show that LPV ERPs are positively associated with future returns in 26 out of 29 equity markets studied. Lee et al. (2021) take this approach further by combining a version of LPV ERP with JLR ERP, estimated using size, book-to-market, and momentum factors as in Lewellen (2015). The resulting characteristics-based expected return (CER) demonstrates superior performance in estimating cross-sectional treatment effects.[6]

In addition to their poor empirical performance, it remains unclear whether these ERPs truly reflect investors' risk-return tradeoff. Most factor- and characteristics-based proxies are derived from firm characteristics that the asset pricing literature identifies as having significant cross-sectional return predictive power. However, these characteristics-

[6] For a separate research objective, Fama and French (2020) also shows that cross-section factors provide better descriptions of average returns than time-series models that using time-series factors. The cross-sectional factor model in Fama and French (2020) is similar to the characteristics-based ERPs while the time-series factor models are similar to the factor-based ERPs in spirit.

return associations are often introduced as “anomalies” and lack a compelling theoretical basis as proxies for non-diversifiable risks.[7] For instance, Daniel and Titman (1997) demonstrate that it is the characteristics, rather than the covariance structure of returns, that explain the cross-sectional variations in stock returns associated with size and book-to-market. Moreover, the choice of characteristics and their weights can significantly impact the performance of these proxies, raising concerns about *p*-hacking (Harvey, Liu, and Zhu 2016).

2.2. Literature on risk factor disclosure

In 2005, the SEC mandated that publicly-traded firms discuss “the most significant factors that make the company speculative or risky” in a new section of their annual reports, namely Item 1A.[8] This mandate aimed to provide investors with useful information about firm risks. Following this mandate, several academic studies have documented that these risk factor disclosures indeed reflect firm risks and are informative to capital market participants. For instance, Campbell et al. (2014) find that firms facing greater risks disclose more risk factors and that the type of risk a firm faces influences the extent to which it describes that specific risk type in its disclosures.

Theoretical models predict that investors should value disclosure about firms’ risk factors because they allow investors to adjust expectations of the variance of cash flows and better estimate the distribution of future cash flows (Heinle and Smith 2017; Heinle, Smith and Verrecchia 2018; Smith and So 2022; Smith 2024). Indeed, prior studies document that

[7] With respect to ICC proxies and the LPV ERPs, the implied required rate of returns inferred from stock prices may also reflect sentiment driven over- or under-pricing rather than rational pricing of systematic risks.

[8] Prior to the mandate, firms only need to disclose risk factors in their registration statements for equity and debt offerings, i.e., form S-1.

equity investors find risk factor disclosure informative. Kravet and Muslu (2013) find that risk factor disclosure is related to return volatility, trading volume, and forecast volatility, suggesting that it increases investors' risk perceptions. Campbell et al. (2014) find that risk factor disclosure increases beta and return volatility but reduce bid-ask spreads. Filzen (2015) finds that quarterly updates to risk factor disclosures are associated with negative market reactions and that these updates are able to predict future negative earnings shocks. Similarly, Lyle, Riedl and Siano (2023) document that additions and removals of risk factors from the disclosure decrease the uncertainty surrounding firm risks. Hope, Hu and Lu (2016) find that the market reacts more strongly to risk factors that are more specific. Risk factor disclosure is also informative to investors in the debt market. Chiu, Guan and Kim (2018) find the credit default swap spreads decrease significantly after the risk factor disclosure was mandated by the SEC, and that the effect is concentrated in disclosure pertinent to financial and idiosyncratic risks.

Furthermore, recent studies document that other stakeholders also incorporate risk factor disclosure in their decision making. For instance, Chiu, Kim and Wang (2019) find that more informative disclosure of customers' risk factors is associated with less under- or overinvestment by suppliers. Kim, Wang and Wu (2023) show that managers increase their firms' pro-environmental activities after the SEC requires firms to disclose climate change risk. Similarly, Dai, Landsman and Peng (2024) find that managers learn new information about their own risks and disclose this information in risk factors to satisfy lenders' demand for transparency.

Nonetheless, regulators and practitioners also question that some of the risk factor disclosure may be boilerplate (IRRC 2016; SEC 2016). Empirical researches have documented some evidence supporting this argument. For instance, Kravet and Muslu (2013) find that some firm-level risk disclosures are likely to be boilerplate. In line with this claim, Beatty et al. (2015) and Dyer et al. (2017) find that risk factor disclosure accounts for a substantial increase in the length of 10-Ks over time. Not surprisingly, Beatty, Cheng and Zhang (2019) document a decrease in the relevance of risk discussion following the 2008 financial crisis, and a lower association between risk factor disclosure and future realized operating risks.

In addition, firms may use risk factor disclosure not as a way to inform investors but as meaningful cautionary statements to protect themselves against securities litigations (Nelson and Pritchard 2016). Consistent with this argument, Cazier, McMullin and Treu (2021) find that lengthier and more boilerplate risk factor disclosures are less likely to be considered inadequate under judicial and regulatory review, and that industry peers borrow risk factor languages that are assessed as adequate in judicial review. Another study by Huang, Shen and Zang (2022) document that firms that began to disclose risk factors after the SEC mandate are more likely to discuss forward-looking information and release more positive news, presumably due to the additional legal protection from the risk factor disclosure. Thus, whether disclosure-based risk factor exposure provides useful information about the expected future returns and facilitate the estimation of ERPs remains an empirical research question.

3. Methodology

3.1. Neural topic modelling

Recent accounting and finance studies have used topic modeling to group economically similar discussions (e.g., Dyers et al. 2017; Huang et al. 2018; Brown et al.

2020). Most of these papers use traditional topic modeling algorithms, such as Latent Dirichlet Allocation (LDA), which assumes that each topic has its own distribution of words and that based on the pattern of word distributions, we can assign texts to different topics (Blei et al. 2003). These algorithms typically rely on a bag-of-words representation, treating words as independent of each other without considering word sequences. This approach, also known as one-hot encoding, results in high-dimensional vector representations of text but does not account for similarities in word meanings. For instance, words like "customer" and "supplier," despite being closely related, are treated as having the same distance in the vector space as unrelated words like "competition" and "cybersecurity." Additionally, these algorithms struggle to accommodate changes in word meaning across different contexts, such as the word "bank" in "river bank" versus "bank deposit."

In recent years, researchers in computational linguistics have introduced neural word embeddings, such as word2vec, which use small-dimensional vectors—usually fewer than 1,000—to represent words based on their meanings inferred from surrounding words (Mikolov et al. 2013). More recent advancements include transformer-based word embeddings, like Bidirectional Encoder Representations from Transformers (BERT), which can consider longer texts, such as sentences, during training. These models have achieved substantial improvements in various natural language processing tasks (Devlin et al. 2019). Recent studies in computational linguistics have leveraged neural word embeddings to group long texts, such as one or more sentences, by their meanings, leading to the development of neural topic models.

In this paper, we use one of the most advanced neural topic modeling approaches, BERTopic, to discover the topics of risk factor disclosures. Introduced by Grootendorst

(2022), BERTopic generates topic representations from text through three modules. The first module, document embedding, converts each given document (or paragraph in our case) into an embedding based on the appearance and sequence of words using a contextualized embedding model. Following Grootendorst (2022), we use a Sentence-Transformer embedding model that produces 384-dimensional embeddings for each paragraph. Sentence-Transformer models are specifically trained for semantic similarity tasks, which closely align with the goal of topic modeling—grouping texts with similar meanings.[9]

In the second step, BERTopic uses a dimensionality reduction technique called Uniform Manifold Approximation and Projection to reduce the 384-dimensional embedding vector obtained from the first stage to a 5-dimensional vector (Becht et al. 2019; Allaoui et al. 2020; McInnes, Healy, and Melville 2020). Reducing the dimensionality of the embedding vector is necessary before the final clustering step, which relies on distances between vectors, because distances become meaningless when dimensions are too high (Aggarwal, Hinneburg, and Keim 2001; Beyer et al. 1999).

In the third and final step, BERTopic employs a hierarchical clustering method called Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) to group similar paragraphs (or documents) based on their embeddings and assign each cluster to a topic (Campello, Moulavi, and Sander 2013). Compared to other clustering algorithms such as K-Means, density-based clustering does not require pre-specifying the number of

[9] Sentence-transformer models are based on sentence-BERT architecture. Unlike the original BERT developed by Google (Devlin et al. 2019) which uses two sentences (texts) as inputs and does not compute independent sentence embeddings, sentence-BERT modifies BERT to output embeddings for individual sentences (texts) (Reimers and Gurevych 2019; 2020). The sentence-transformer model we use is available at Hugging Face: https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

clusters and can better handle noise. Additionally, HDBSCAN's hierarchical structure can accommodate clusters of different densities and simplify the clusters into an easily interpretable hierarchy.[10]

After these three steps, BERTopic assigns paragraphs that belong to each cluster into one topic. To facilitate us interpreting the economic meaning of the topics, we represent each topic using its most unique words. Specifically, we aggregate documents in each cluster (topic) and measure the importance of a word to a topic consisting with all paragraphs it includes using class-based Term-Frequency Inverse Document Frequency (c-TF-IDF).[11] Recent studies such as Egger and Yu (2022) show that BERTopic outperforms traditional topic modeling algorithms including LDA. Other practical advantages of BERTopic compared to LDA is that it does not require researchers to pre-process the texts or determine the number of topics, thus reducing subjectivity in the results.

3.2. Risk based peers

After we use BERTopic to assign a topic to each paragraph of Item 1As, we can convert firms' textual disclosure of risk factors into numerical risk factor exposure. Specifically, we represent each firms' risk factor exposure as a vector of topics based on the proportion of texts assigned to each topic. For example, assuming that there are five topics in

---

[10] Researchers can choose which specific models to use in BERTopic's modules. For example, in first step, researchers can choose from a number of embedding models (e.g., universal sentence encoder) to produce the embedding of the texts. Similarly, one can select the dimensionality reduction algorithms (e.g., principal component analysis) and cluster algorithms (e.g., k-Means) in the second and third steps respectively. We use the default choices following the original BERTopic model in all three steps (Grootendorst 2022).

[11] For a word x within a topic c, its c-TF-IDF ($w_{x,c}$) is calculated using the formula: $w_{x,c} = tf_{x,c} \times log\left(1 + \frac{A}{f_x}\right)$, where $tf_{x,c}$ is the frequency of word x in topic c; f_x is the frequency of word x in across all topics; and A is the average number of words per topic.

a year, a firm with 250, 0, 450, 0, and 300 characters of discussion in these five topics, respectively, would have a risk factor exposure of [0.25, 0, 0.45, 0, 0.3]. In essence, we assume that firms discuss risks that poses a larger threat to itself in more details, consistent with the finding in Campbell et al. (2014). Therefore, our risk factor exposure definition measures not only whether a firm is subject to a specific type of risk, but also the magnitude of the threats from that risk.

To test the validity of the risk factor exposure, we identify peers with similar risk factor exposures (or risk-based peers, RBPs) for each firm year observation in our sample. If the risk factor exposures capture firms' exposures to underlying economic shocks, we would expect firms with similar risk factors to be economically related and therefore have higher comovements in stock prices and fundamental performance. Each year, we calculate the pair-wise cosine similarity of risk factor exposure for all firms. For each firm, we then identify the top ten firms with the highest cosine similarity as its RBPs.[12]

3.3. Risk-based expected return proxy

Using the risk factor exposures generated from the BERTopic algorithm, we construct a risk-based expected return proxy (RB_ERP) through the following procedures:

i. Each month from July of year t+1 to June of t+2, we obtain the beginning-of-month risk factor exposure estimated from Item 1A in 10-Ks filed in the prior year.

[12] Peer firms are important for setting benchmarks in valuation and compensation. Prior studies in accounting and finance identify peer firms using concepts such as common analysts, production complementary, employee attention, investor perception, technology links (Lee, Ma and Wang 2015; Lee et al. 2019; Kaustia and Rantala 2021; Lee et al. 2024; Li 2024). Our study offers a new approach to identify peer firms, which can be potentially useful as alternative benchmarks for equity valuation and compensation contracting.

ii. We perform Principal Component Analysis (PCA each month using the above risk factor exposure to reduce the dimension of the exposures, retaining only the first five principal components.[13] Our results suggest that the first five PCs explain approximately 50% of the variation in the original data. We standardize the five PCs with z-score transformation so that they have a mean of zero and standard deviation of one.

iii. We conduct cross-sectional regressions of the monthly stock returns on the five (beginning-of-the-month) standardized PCs each month. As Fama (1976) and Fama and French (2020) suggest, the slope coefficients from the cross-sectional regression can be interpreted as factor premium, which is the return on a zero-investment portfolio that has a net exposure of one to the corresponding PC and zero exposure to other PCs.

iv. Using the standardized PCs as factor sensitivity (loading) and the slope coefficients from the above regression as factor premiums, we compute the cross-sectional expected returns attributable to the latent factors as the fitted value of the regression.

v. To alleviate the effect of random noise in the monthly estimate, we use the 12-month moving average of the above monthly estimates as our primary risk-based expected return proxy (RB_ERP) and use it for the next 12 months.

We use characteristics-based ERPs as the benchmarks because recent literature indicates that they outperform ICC- and factor-based ERPs. Our first benchmark, LPV, is

[13] BERTopic identifies between 126 to 356 topics in risk factor disclosures. It's extremely unlikely that all the topics represent nondiversifiable risk. Some are likely highly correlated. Others might represent idiosyncratic or redundant factors. Directly using these high-dimensional exposures to estimate ERP might lead to overfitting problem. Dimension reduction via PCA mitigates the problem.

grounded in the log-linear present-value relationship between a firm's current stock price, an accounting valuation anchor, and its expected future growth, as developed by Chattopadhyay et al. (2022). Similar approaches are also employed in the studies by Lyle and Wang (2015) and Lee et al. (2021). Specifically, we utilize a version that estimates expected simple returns within the LPV framework:

$$R_{i,t+1} = \beta_{i,0} + \beta_{i,1} BM_{i,t} + \beta_{i,2} ROE_{i,t} + \beta_{i,3} RVOL_{i,t} + \varepsilon_{i,t+1}$$

where $R_{i,t+1}$ is the simple return over *t*+1; $BM_{i,t}$ is the book-to-market ratio as of month *t*; $ROE_{i,t}$ is the most recent annual return on equity (Income before extraordinary items/average common equity), available before the end of month *t*; RVOL is the standard deviation of daily stock returns in month *t*. We estimate this cross-sectional regression on a monthly basis for all months in our sample period. $E_t[R_{i,t+1}]$ of LPV is computed with the mean coefficients over the 36 months prior to month *t* and the value of characteristics (i.e. BM, ROE, and RVOL) of month *t*.[14]

We estimate the second benchmark proxy, JLR, following the approach used by Lee et al. (2021). This model, originally introduced by Lewellen (2015), is based on three firm characteristics: SIZE (the logarithm of market capitalization), the book-to-market ratio, and cumulative stock returns from 12 months to 2 months prior to the estimation date, representing momentum. In line with Lee et al. (2021), we also create CER, the third benchmark proxy, by taking the average of LPV and JLR.

[14] In untabulated analyses, we use log(BM) and log(1+ROE) as in Chattopadhyay et al. (2022); we also use mean coefficients for all available months prior to month t to compute LPV_ER. The results are qualitatively similar.

Finally, considering the widespread adoption of the Fama and French five-factor model (Fama and French, 2015), we introduce the fourth characteristics-based ERP benchmark, FFC. FFC is estimated using the same procedure as described previously, but it incorporates the following characteristics: SIZE (the logarithm of market capitalization), BM (book-to-market ratio), OP (operating profitability, calculated as (revenue - cost of goods sold - selling, general, and administrative expenses - interest expense) / average book value of common equity), and INV (the rate of growth in total assets).

3.4. ERP comparison framework

We compare the performance RB_ERP with the benchmark ERPs along several dimensions. First, following prior literature (e.g. Easton and Monahan 2005; Chattopadhyay et al. 2022), we examine whether the ERPs are significantly associated with realized stock returns in the cross-section. Specifically, we study their cross-sectional correlations with future realized return, the slope coefficients from the cross-sectional regression of future realized return on ERPs and standardized ERPs, and the R^2 from the cross-sectional regression. Furthermore, we measure the future realized return spread between extreme deciles of ERPs to ascertain whether the ERPs help investors distinguish economically significant cross-sectional variation in realized returns. Last, Lee et al. (2021) argue that when ERPs are used to study treatment effects, one should use ERP's measurement error variance to select the most appropriate proxy. For this reason, we compare the cross-sectional measurement error variance of RB_ERP with that of the benchmark ERPs.

4. Empirical results

4.1. Sample selection and descriptive of disclosure-based risk factors

Our sample firms include all publicly-traded firms in the U.S. with available risk factor disclosure and required financial statement and stock returns data. We start with 9,242 firms with available 10-Ks from Edgar from 2006, the year that the SEC required firms to disclose risk factors, to 2022. We are able to extract Item 1A from 92,665 of the 10-Ks (9,227 unique firms), ranging from 4,010 in 2006 to 4,738 in 2022. We drop firm-year observations without required data from Compustat and CRSP. Our final sample includes 47,931 risk factor descriptions of 6,170 firms.

As discussed in Section 3.1, we conduct the topic modeling analysis at the paragraph level. That is, we define each paragraph in an Item 1A in 10-K as a document. We choose to model topics at the paragraph level instead of using an entire risk factor discussion of a firm-year or a sentence in the risk factor at the unit (Bao and Datta 2014; Grundy and Petry 2022) because firms usually discuss one risk factor in each paragraph, thus facilitating us discovering economically meaningful topics from risk factors. Our sample consists of 6,298,175 paragraphs. Each paragraph has an average of 66.28 words and 2.09 sentences.

As detailed in Section 3.3, to prevent look-ahead bias, i.e., BERTopic using information from subsequent textual disclosure to assign topics, we use a rolling window in constructing the training sample for BERTopic. Specifically, we train BERTopic at the end of June of every year, using risk factors included in 10-Ks filed within the previous 12 months, i.e., from July of the previous year. On average, each training sample includes 393,636 paragraphs. Initially, BERTopic assigns a topic to half of the paragraphs (e.g. 53.84% in

2022), with the remaining labelled as outliers, i.e., not belonging to any topics. We further assign each unassigned paragraph to a topic based on its tokens' most likely topic.[15]

Figure 1 shows the number of topics in each year's topic modeling result. We observe that the number of topics each year increases steadily from 126 in 2007 to 218 in 2020, but jumps to 283 and 356 topics in the last two years. Table 1 shows that, on average, each risk factor disclosure includes 47 topics, also steadily increasing over time from 30 in 2007 to 72 in 2022. Each topic covers an average of 2.61 paragraphs (2.66 in 2007 to 2.53 in 2022). The standard deviation of topic length is 0.95 paragraphs, suggesting that topic lengths have substantial variations.

[Figure 1 about here.]

[Table 1 about here.]

Figure 2 presents a word cloud generated from 72 of the 363 topics produced by BERTopic, based on the Item 1A paragraphs in 10-K filings from July 2021 to June 2022. The size of each word within the cloud corresponds to its c-TF-IDF value, indicating the word's importance to the respective topic. This figure effectively demonstrates BERTopic's capability to identify distinct and economically significant risk factors that firms encounter. These risks range from those common across industries—such as macroeconomic risk (Topic 64), climate risk (Topic 6), dependency on suppliers (Topic 14), reliance on major customers (Topic 71), competitive pressures (Topic 23), currency fluctuations (Topic 21), retention of

[15] Specifically, we use the default in BERTopic, which is a sliding window with a size of four tokens and a stride of one. That is, the algorithm calculates the topic distribution of each four consecutive tokens based on the comparison of the c-TF-IDF of the four tokens and that of each topic. It then aggregates the topic distribution of all four-token sets of a paragraph, and assign the paragraph to the topic with the highest probability. After this step, fewer than 1% of paragraphs remained unassigned to any topic.

key personnel (Topic 7), and weaknesses in internal controls (Topic 17)—to those specific to particular sectors, such as clinical trials in the pharmaceutical industry (Topic 38) and banking regulations (Topic 61). Additionally, the analysis highlights emerging risk factors like the COVID-19 pandemic (Topic 0), digital assets (Topic 41), and various geopolitical tensions, including issues related to China (Topic 47) and the Ukraine-Russia conflict (Topic 48). However, some redundancy is observed, for instance, Topics 38 and 69 both address clinical trials, and Topics 12 and 42 concern tax enforcement. These overlaps indicate that certain topics may capture similar economic risks. To address this, we apply PCA to the risk factor data, aiming to mitigate the redundancy and reduce the dimensions before estimating the proxy for expected returns.

[Figure 2 about here.]

4.2. Risk factor exposures and comovement

As outlined in Section 3.2, we utilize BERTopic to categorize the topics within the risk factor disclosures of each firm-year. We then represent the risk factor exposure of each firm-year as a vector, the elements of which correspond to the proportion of risk factor disclosures. Subsequently, for each firm, we identify its RBPs annually as the ten firms exhibiting the highest cosine similarity in terms of risk factor exposure.

We present the summary statistics of RBPs and other peer firms in Table 2. As detailed in Panel A, the average cosine similarity for the RBPs is 0.785, which is significantly higher than the overall firm-pair average of 0.211 and also exceeds the average cosine similarity of the top ten firms with the highest cosine similarity in risk factor exposure within the same industry: 0.420 for firms sharing a two-digit GICS code, 0.554 for those with a three-digit SIC code, and 0.505 for those with a three-digit NAICS code. Notably, the highest average cosine similarity within the same industry (based on a three-digit NAICS code) is only 0.829.

For additional reference points, the average cosine similarity between a firm and the ten firms with the most similar risk factor exposures within the same two-digit GICS, three-digit SIC, and three-digit NAICS codes are 0.756, 0.734, and 0.718, respectively. These statistics suggest that firms with the most similar risk factor exposures often do not share the same narrowly defined industry.

Panel B of Table 2 further illustrates this point: only about 73.2% of the average firm's ten RBPs share the same two-digit GICS code, while only 43.6% and 50.8% share the same three-digit SIC and NAICS codes, respectively. Additionally, the standard deviation of cosine similarity among firm-pairs within the same two-digit GICS (three-digit SIC, three-digit NAICS) is 0.153 (0.184 and 0.166). These findings underscore significant variations in risk factor exposure similarity that are not accounted for by industry classifications.

[Table 2 about here.]

Next, we investigate whether firms that report similar risk factors indeed face comparable economic risks. To assess this, we analyze the correlation between the returns and operating performance of firms and their RBPs, using industry peers defined by a six-digit GICS code as a benchmark. For each focal firm, we randomly select ten industry peers with the same six-digit GICS code, matching the number of RBPs.

We first examine the relationship between the stock returns of the focal firm and those of its peers. For the returns of industry peers, we calculate the average (equal-weighted) returns of the ten firms. For the returns of RBPs, we consider both the mean returns (equal-weighted) and the weighted average returns, using the cosine similarity of each peer's risk factor exposure to the focal firm as weights. We employ the Fama-MacBeth methodology to estimate cross-sectional returns on a monthly basis over a period spanning 186 months, from July 2007 to December 2022. The results are presented in Table 3, where Columns (1), (2), and (3) report the average R-squares for equal-weighted RBPs, weighted-average RBPs, and equal-weighted industry peers, respectively. Columns (4) and (5) detail the differences in the explanatory power of risk-based peers' returns relative to industry peers' returns.

Our findings reveal that both industry and risk-based peers' returns significantly explain the contemporaneous returns of focal firms, as indicated by positive and significant average R-squares in the first three columns. Notably, the returns of RBPs demonstrate substantially greater explanatory power, with positive and significant differences highlighted in the last two columns. Economically, the equal-weighted and weighted-average returns of RBPs outperform those of industry peers by 57% (0.0364/0.0232 − 1) and 59% (0.0369/0.0232 − 1), respectively, in explaining the returns of focal firms.

[Table 3 about here.]

To investigate if the return explainability advantage of RBPs over industry peers is influenced by the size of the focal firm, we analyze the return associations for both large and small firms, partitioned based on median market capitalization. Across both size categories, RBPs consistently outperform industry peers in terms of return explainability, with economic magnitudes of 60% for larger firms and 49% for smaller firms.

Additionally, we investigate the correlation between the financial metrics of focal firms and their peers, including valuation multiples, financial statement ratios, capital structure, and other financial indicators. Following Lee, Ma, and Wang (2015), we examine price-to-book multiple (pb), enterprise value-to-sales multiples (evs), and price-to-earnings multiple (pe), returns on net operating assets ($rnoa$), returns on equity (roe), asset turnover (at), profit margin (pm), leverage (lev), gross profitability (gpr), and research and development (R&D) expenses scaled by net sales ($rdpersales$). Detailed definitions on these variables are provided in Appendix.

Similar to the previous analysis, we adopt the Fama-MacBeth approach by estimating cross-sectional regressions of focal firms' variable of interest on that of the peer firms in each calendar quarter (using the most updated number as of the end of March, June, September and December). We estimate a total of 62 quarterly regressions from Q3 2007 to Q4 2022 and tabulate their average R-squares in Table 4.

[Table 4 about here.]

Across all measures, we observe that RBPs are more closely associated with focal firms than those of industry peers (differences significant at 1% except for $rdpersales$). For valuation multiples (pb, evs and pe), RBPs can explain between one fifth (evs) to a half (pb) more of the variation in focal firms than industry peers. The advantage of RBPs over industry peers in explaining financial statement ratios and other financial metrics vary greatly, with profit margin as high as more than doubling (0.1374/0.0515 = 2.66 times) and R&D per sales only improving 2% (0.5706 versus 0.5599). This evidence, combined with that from the return co-movements, confirms that firms disclosing similar risk factors are indeed exposed to similar economic risks.

4.3. Risk-based vs. characteristics-based ERPs

4.3.1. Main comparative results

In this sections below, we test the performance of expected return proxy estimated from the risk factor disclosure (RB_ERP) against the benchmark characteristics-based proxies. Table 5 presents the descriptive statistics for the ERPs alongside the one-month ahead realized stock returns (RET1). Panel A reveals that the risk-based ERP (RB_ERP) shares a similar mean value with the realized return, whereas the characteristics-based ERPs generally show a downward bias. Additionally, all ERPs significantly underestimate the cross-sectional variation observed in future realized stock returns. The cross-sectional standard deviation of RET1 averages about 0.151, while the standard deviations of the ERPs are mostly less than one-tenth of this value, with the RB_ERP exhibiting the highest variation among them at 0.0174. The inter-quartile range results are consistent, displaying an average of approximately 12.2% for RET1, but only 1.94% for the RB_ERP, and even lower for the characteristics-based ERPs. This underestimation highlights the limitations of ERPs in capturing the full extent of variability in future stock returns.

[Table 5 about here.]

In Table 6, we report the performance metrics of risk- and characteristics-based ERPs. Column (1) of Panel A reports the mean monthly correlation coefficient between ERPs and RET1. RB_ERP has the highest correlation of 0.0201 with a t-statistic of 2.28, which is followed by LPV at 0.0184 with a t-statistic of 1.69. The cross-sectional correlations for the other characteristics-based ERPs are statistically insignificant for our sample period. Column (2) of Panel A reports the mean slope coefficients of the monthly cross-sectional regression of RET1 on each of the expected return proxy. The results show that the slope coefficient for RB_ERP is significantly positive with a mean of approximately 0.486 and a t-statistic of 2.97. The coefficient for LPV is also positive, at 0.659, but only significant at the 10% level. Among the rest of the characteristics-based ERPs, only CER, which is an equally weighted average of LPV and JLR, also has a marginally significant positive coefficient.

Note that all ERPs exhibit relatively low cross-sectional variation (STD). Ceteris paribus, proxies with lower variance tend to exhibit relatively higher slope coefficients in cross-sectional regressions, as the slope coefficient in a univariate regression is simply calculated as cov(x,y)/var(x). Therefore, we also perform the regression using standardized ERPs, ensuring that all proxies have the same cross-sectional standard deviation of one. The results, reported in Column (3) of Panel A, show that the coefficient for this alternative specification is statistically significant only for the risk-based ERPs.

Column (4) of Panel A reports the mean R-square of the regression. The R-squares from the cross-sectional regressions are generally low, ranging from 0.75% for FFC to 1.84% for LPV. However, we refrain from drawing strong inferences from these comparative results, as regression slope coefficients are often negative (for approximately 20% of the months for all proxies). A higher R-square, particularly if it derives from months with negative coefficients, does not necessarily indicate better explanatory power of the ERPs.

To determine if these ERPs help investors identify economically significant cross-sectional variations in expected returns, we sort firms into ten equally sized deciles based on ERPs at the end of the forecasting month and track the return spread between the two extreme deciles in the subsequent month. The return spread is calculated from a hedge portfolio that takes a long position in the decile with the highest expected returns and a short position in the decile with the lowest expected returns. Columns (5) and (6) report the equal- and value-weighted returns of the ERP-sorted hedge portfolios. The results indicate that RB_ERP generates monthly hedge returns of 1.15% and 0.9% for equal- and value-weighted portfolios, respectively, both of which are statistically significant. In contrast, none of the characteristics-based proxies produce statistically significant hedge returns. Among these, CER-sorted portfolios yield the highest mean monthly hedge returns of 0.43% and 0.37% for equal- and value-weighted portfolios, respectively, though these returns are statistically insignificant.

Finally, Column (7) of Table 6 reports the scaled MEV of various ERPs, as defined by Lee et al. (2021). According to Lee et al. (2021), a lower scaled MEV indicates smaller measurement errors in the cross-section and is therefore preferred. The results reveal that RB_ERP exhibits the smallest MEV at 0.0012, which is statistically insignificant from zero. In contrast, the MEV for JLR and FFC is statistically significant at the 5% level.

At first glance, there appear to be some discrepancies between our results and those of recent studies (e.g., Lee et al. 2021 and Chattopadhyay et al. 2022), which document a substantially stronger association between characteristics-based ERPs and realized returns, and find a statistically significant negative scaled MEV for the CER measures. Additional analysis suggests that these discrepancies are primarily driven by differences in the sample periods. Due to the limited availability of Item 1A disclosures, our sample period starts from 2008. In contrast, the sample period in Lee et al. (2021) covers 1977 to 2018, while Chattopadhyay et al. (2022) extends even earlier, covering all firms with available accounting and price data until the end of 2017. To reconcile with the findings of previous studies, we generated characteristics-based ERPs using the same procedures as our main analysis but applied to data prior to our sample period of 2008. The results reported in Panel B of Table 6 indicate that characteristics-based ERPs perform substantially better in this earlier sample period. For instance, take CER as an example: the mean cross-sectional correlation between CER and RET1 increases to 0.049, which more than triples the corresponding figure in Panel A of Table 6 (0.016). Furthermore, the mean slope coefficients on the raw and standardized

CER are 1.2386 and 0.006, respectively, both statistically significant at less than the 1% level, and their magnitudes are also substantially larger than their counterparts in Panel A.

In terms of hedge returns (HGRET), all characteristics-based ERPs generate statistically and economically significant returns, with equal- and value-weighted monthly hedge returns of 2.1% and 1.99%, respectively, for CER. Consistent with the findings of Lee et al. (2021), the scaled MEV for CER is -0.0022, significant at the 10% level. These results suggest that in recent years, the performance of characteristics-based ERPs has deteriorated considerably compared to the results from the older and longer sample periods presented in prior studies. In contrast, the risk-based ERP continues to perform reasonably well even when other ERPs suffer during this period.

[Table 6 about here.]

4.3.2. Subsample analyses

In Table 7, we investigate the subsample performance of risk- and characteristics-based ERPs. To conserve space, we only report the results for CER for comparison with RB_ERP. The performance of other characteristics-based ERPs is qualitatively similar and available upon request. The results for RB_ERP are presented in Panel A. Column (1) shows that the magnitudes of correlations with future realized returns are similar between large and small firms (partitioned on annual sample median), with a correlation coefficient of 0.0183 for the top half and 0.0207 for the bottom half of the sample. We also partition our sample into recession and boom periods according to the NBER business cycle classification. The recession period in our sample includes 16 months of the total 175 months, ranges from June 2008 to June 2009, and from February 2020 to April 2020. In this period, while the correlation coefficient is still positive (0.0036), it becomes statistically insignificant, likely due to low power as a result of small sample. The results for the boom period are slightly stronger than the full-sample period, with a correlation of 0.0217.

The results for slope coefficients and hedge portfolio returns exhibit a similar pattern, showing robust results for both the top and bottom half of the sample, with the results being slightly stronger for the bottom half. Perhaps due to the short sample period, the proxy struggles in the recession period, but the results for the boom period are generally stronger than those in the full-sample period. Column (7) shows that scaled MEV is statistically insignificant for all subsamples. Consistent with Lee et al. (2021), MEV is slightly lower for smaller firms than for large firms.

We report the results for CER in Panel B for comparison purposes. The results show that the association between CER and future stock returns appears to be stronger among small firms, with a statistically significant correlation of 0.0232 for the bottom half of the sample versus a statistically insignificant correlation of 0.0097 for the top half. CER also struggles in the short recession period, with a correlation of -0.0111 and a t-statistic of -0.22. The results for the slope coefficients present a similar pattern. In terms of hedge return and scaled MEV, all subsamples are statistically insignificant. Thus, the results in Table 7 confirm that risk-based ERPs perform at least as well as CER and often outperform it.

[Table 7 about here.]

4.3.3. Performance evaluation using longer term realized returns

The results so far rely on one-month ahead return to evaluate various ERPs. In this section, we examine the robustness of the results for longer-term returns. Specifically, we investigate whether ERPs are associated with monthly returns in t+2 (RET2) and t+3 (RET3). Panel A of Table 8 presents the results using RET2. We find that the results are slightly weaker, but all inferences remain unchanged. The correlation for RB_ERP decreases marginally to 0.0154, remaining statistically significant at the 10% level. In contrast, none of the characteristics-based ERPs exhibits a statistically significant correlation with RET2. The slope coefficients in the return-ERP regressions maintain positive and statistically significant values for RB_ERP, but are statistically insignificant for all characteristics-based ERPs.

Similar patterns emerge for hedge portfolio returns. We document that the return spread at month *t*+2 amounts to 0.94% and 0.66% on an equal- and value-weighted basis, respectively, for the risk-based ERP. Conversely, the hedge portfolio returns over month *t*+2 are all statistically insignificant and often negative for characteristics-based ERPs. When evaluating with RET2, the scaled MEV remains statistically insignificant for RB_ERP, but becomes significantly positive for all characteristics-based proxies. Panel B of Table 8 presents the results using RET3, which offer a similar picture, demonstrating that in terms of both cross-sectional association with future returns and cross-sectional measurement error variances, the risk-based ERP outperforms the characteristics-based proxies.

[Table 8 about here.]

4.3.4. Alternative specifications for risk-based ERPs

We further examine the robustness of the risk-based ERP using alternative construction specifications. In our primary RB_ERP, we average the monthly expected components of return over the past 12 months to mitigate the effect of random noise. Table 9 presents the performance of risk-based proxies over alternative moving average windows of 9 and 6 months. The results indicate that the inferences remain consistent across these specifications. Specifically, both RB_ERP_MA(9) and RB_ERP_MA(6) demonstrate a significant association with future realized returns in the cross-section. All correlations, slope coefficients, and hedge portfolio returns maintain positive and statistically significant values. Regarding measurement error variance, we find that it becomes statistically significant when using a 6-month moving average.

[Table 9 about here.]

In our primary specification, we use the first 5 principal components to estimate the expected return component. Table 10 reports the results of risk-based ERPs using different numbers of principal components: 3, 10, 20, and 30, respectively. We find that the results remain robust across these alternative specifications. Risk-based proxies generated from all four alternative specifications are significantly associated with future realized stock returns, as evidenced by their cross-sectional correlation with future stock returns, slope coefficients in the return-ERP regressions, and return spreads between extreme deciles of ERPs. All these performance metrics are statistically significant, with the exception of the value-weighted return to the hedge portfolio formed on the expected return proxy constructed with the first 30 principal components. Furthermore, we find that the scaled MEV remains statistically insignificant except for the proxy estimated with 30 PCs. These results suggest that our inference is generally robust to different numbers of PCs used in ERP estimation. However, there is some evidence that when we use too many PCs, the benefit of dimension reduction diminishes, potentially introducing more noise to the ERPs and leading to weaker performance.

[Table 10 about here.]

4.3.5. Incremental usefulness of risk factor exposures in expected return estimations

In the preceding comparison, all characteristics-based ERPs are constructed using a small number of firm characteristics. Chattopadhyay et al. (2022) conduct a comprehensive analysis by incorporating several sets of additional characteristics into their LPV-based ERPs. Specifically, they add: (i) ten additional characteristics that Green et al. (2017) find to be incrementally informative of expected returns in the US; (ii) five additional characteristics from Lewellen (2015); and (iii) a total of 74 additional characteristics, including the 15 characteristics from (i) and (ii), plus 59 other characteristics from Green et al. (2017). Their findings indicate that incorporating these additional characteristics does not systematically improve the performance of LPV-based ERPs, as measured by the change in the R-square of the cross-sectional return-ERPs regressions. In fact, they observe that doing so decreases the R-squares for most countries in their sample. For instance, adding 74 additional characteristics leads to a significant decrease in R-square of 0.003 for the U.S. (Table 10 of Chattopadhyay et al. 2022).

One potential explanation for the lack of incremental usefulness of the additional characteristics is the high correlation among many of these characteristics, which may capture similar underlying constructs. Indeed, Green et al. (2017) find that only 12 characteristics are reliably independent determinants of non-microcap US stock returns over the full sample period of 1980-2014. Moreover, in the more recent period after 2003, only two characteristics have been identified as independent determinants. In contrast, our risk factor exposures are constructed from textual disclosures using unsupervised learning algorithms. These exposures are likely to be less correlated with the various characteristics that have been thoroughly examined in the literature. Consequently, they have a higher potential to provide independent and distinct information regarding the cross-section of expected returns.

To investigate this conjecture, we first examine whether adding risk-based characteristics improves the R-square of the cross-sectional return-ERPs regressions, following the approach of Chattopadhyay et al. (2022). Specifically, we estimate ERPs using characteristics from LPV, JLR, CER (union of LPV and JLR characteristics), and FFC, employing the fitted value approach (i.e., the same procedures used to estimate RB_ERP). We then augment these characteristics with the first five principal components and re-estimate the ERPs. Columns (1) and (2) of Table 11 Panel A report the increase in R-square of the cross-sectional return-ERPs regressions resulting from the addition of risk-based characteristics (i.e., first five PCs of risk factor exposures). The results demonstrate that incorporating risk-based characteristics leads to a statistically significant increase in the R-square across all four benchmarks, ranging from 0.0048 to 0.0063.

Additionally, we employ a kitchen-sink approach by adding all individual risk factor exposures identified by the BERTopic algorithm (rather than their lower-dimensional representation, i.e., 5 PCs) to the characteristics used in the four characteristics-based ERPs. The results, reported in Columns (3) and (4), again demonstrate that risk-based measures improve the performance of characteristics-based ERPs. However, the magnitudes of improvement are sometimes smaller and less significant, possibly due to overfitting issues in regression models incorporating hundreds of individual risk factor exposures.

To further demonstrate the incremental usefulness of risk factor disclosures in ERP estimation, we decompose the ERPs estimated from the augmented model in Panel A into two components: one attributable to the original characteristics (moving average of sum(characteristics × estimated coefficients)) and another attributable to the risk exposures or their lower-dimensional representations (moving average of sum(risk exposure × estimated coefficients)). We then examine their respective associations with future realized stock returns in the same regression. The results of this analysis are presented in Panels B-E of Table 11.

Panel B of Table 11 reports the results for characteristics used in LPV. Columns (1) and (2) show that both the characteristics component and the (reduced dimensional) risk component of ERP are positively associated with future realized stock returns. Both coefficients are statistically significant at the 10% level. Columns (3) and (4) demonstrate that using the original high-dimensional risk factor exposures leads to an even stronger association between the risk component of ERP and future returns, with a t-statistic of 2.63.

Panel C of Table 11 presents the results using JLR characteristics, which show even stronger effects. We find that the ERP component attributable to the JLR characteristics is no longer significantly associated with future realized stock returns. In contrast, the component attributable to risk factors is positively associated with future stock returns, with coefficients statistically significant at the 5% and 1% levels for the reduced-dimensional and original high-dimensional risk exposures, respectively.

The results in Panels D and E for CER and FFC characteristics are qualitatively similar, both demonstrating that risk factor exposures estimated from Item 1A contain incremental information with respect to future cross-sectional expected returns.

[Table 11 about here.]

5. Conclusions

This paper examines the utility of mandatory risk factor disclosures in Item 1A of corporate 10-K filings for investors' assessment of firms' comovement risk and expected returns. Using BERTopic, a deep neural network-based topic modeling algorithm, we extract numerical risk factor exposure vectors from textual disclosures for 47,931 firm-year observations from 2007 to 2022. Our empirical analyses suggest that BERTopic not only extracts economically meaningful risk factors from corporate disclosures but also provides reasonable estimates of firms' exposures to these risks.

We find that the resulting risk factor exposures are useful for investors in estimating comovement (risk) between firms. Specifically, (i) firms with similar disclosure-based risk exposures are more likely to be from the same industries, and (ii) compared to conventional industry peers, risk-based peers (defined as the ten firms with the highest exposure cosine similarity to the focal firms) explain a substantially higher fraction of cross-sectional variation in stock returns, fundamental profitability, valuation, and other key financial ratios of the focal firms.

Further analyses indicate that disclosure-based risk factor exposures also improve estimates of firm-specific expected returns. We adopt a Principal Component Regression algorithm to estimate a risk-based expected return proxy (ERP) from high-dimensional risk factor exposure data. The resulting risk-based ERP performs at least as well as, and often better than, state-of-the-art characteristics-based ERPs developed in recent literature and estimated from financial characteristics.

In particular, the risk-based ERP demonstrates a stronger and more consistent association with future realized stock returns than characteristics-based proxies. Furthermore, its cross-sectional scaled-measurement error variance (MEV) is no greater than the best-performing characteristics proxies, suggesting that the risk-based ERP is also one of the best proxies for researchers interested in understanding the cross-sectional treatment effect on expected returns. Our study demonstrates that applying sophisticated unsupervised machine learning algorithms to corporate risk factor disclosures can help investors and academics better understand firms' comovement risk and expected returns.

References

Aggarwal, C.C., Hinneburg, A. and Keim, D.A., 2001. On the surprising behavior of distance metrics in high dimensional space. In *Database theory—ICDT 2001: 8th International Conference*, London, UK, January 4–6, 2001 proceedings 8 (420-434). Springer Berlin Heidelberg.

Allaoui, M., Kherfi, M.L. and Cheriet, A., 2020, June. Considerably improving clustering algorithms using UMAP dimensionality reduction technique: a comparative study. In *International Conference on Image and Signal Processing (317-325)*. Cham: Springer International Publishing.

Bao, Y. and Datta, A., 2014. Simultaneously discovering and quantifying risk types from textual risk disclosures. *Management Science* 60, 1371-1391.

Beatty, A., Cheng, L. and Zhang, H., 2015. Sometimes less is more: evidence from financial constraints risk factor disclosures. *Working paper*.

Beatty, A., Cheng, L. and Zhang, H., 2019. Are risk factor disclosures still relevant? Evidence from market reactions to risk factor disclosures before and after the financial crisis. *Contemporary Accounting Research* 36, 805-838.

Becht, E., McInnes, L., Healy, J., Dutertre, C.A., Kwok, I.W., Ng, L.G., Ginhoux, F. and Newell, E.W., 2019. Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology* 37, 38-44.

Beyer, K., Goldstein, J., Ramakrishnan, R. and Shaft, U., 1999. When is "nearest neighbor" meaningful?. In *Database Theory—ICDT'99: 7th International Conference*, Jerusalem, Israel, January 10–12, 1999 Proceedings 7 (217-235). Springer Berlin Heidelberg.

Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3, 993-1022.

Botosan, C.A., 1997. Disclosure level and the cost of equity capital. *Accounting Review*, 323-349.

Brown, N.C., Crowley, R.M. and Elliott, W.B., 2020. What are you saying? Using topic to detect financial misreporting. *Journal of Accounting Research* 58, 237-291.

Campbell, J.L., Chen, H., Dhaliwal, D.S., Lu, H.M. and Steele, L.B., 2014. The information content of mandatory risk factor disclosures in corporate filings. *Review of Accounting Studies* 19, 396-455.

Campello, R.J., Moulavi, D. and Sander, J., 2013, April. Density-based clustering based on hierarchical density estimates. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining (160-172)*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Cao, K. and You, H., 2024. Fundamental Analysis via Machine Learning. *Financial Analysts Journal* 80, 74-98.

Cazier, R.A., McMullin, J.L. and Treu, J.S., 2021. Are lengthy and boilerplate risk factor disclosures inadequate? An examination of judicial and regulatory assessments of risk factor language. *The Accounting Review* 96, 131-155.

Chan, L.K., Karceski, J. and Lakonishok, J., 2003. The level and persistence of growth rates. *The Journal of Finance* 58, 643-684.

Chattopadhyay, A., Lyle, M.R. and Wang, C.C., 2022. Expected stock returns worldwide: A log-linear present-value approach. *The Accounting Review* 97, 107-133.

Chen, X., Cho, Y.H., Dou, Y., Lev, B., 2022. Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research* 60, 467–515.

Chen, L., Pelger, M. and Zhu, J., 2024. Deep learning in asset pricing. Management Science 70, 714-750. Claus, J. and Thomas, J., 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *The Journal of Finance* 56, 1629-1666.

Chiu, T.T., Guan, Y. and Kim, J.B., 2018. The effect of risk factor disclosures on the pricing of credit default swaps. *Contemporary Accounting Research* 35, 2191-2224.

Chiu, T.T., Kim, J.B. and Wang, Z., 2019. Customers' risk factor disclosures and suppliers' investment efficiency. *Contemporary Accounting Research* 36, 773-804.

Cochrane, J.H., 2011. Presidential address: Discount rates. *The Journal of Finance* 66, 1047-1108.

Daniel, K. and Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. *The Journal of Finance* 52, 1-33.

Dai, L., Landsman, W.R. and Peng, Z., 2024. Private Loan Issuance and Risk Factor Disclosure. *The Accounting Review* 99, 169-196.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dyer, T., Lang, M. and Stice-Lawrence, L., 2017. The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64, 221-245.

Easton, P.D. and Monahan, S.J., 2005. An evaluation of accounting-based measures of expected returns. *The Accounting Review* 80, 501-538.

Egger, R. and Yu, J., 2022. A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in Sociology* 7, 886498.

Fama, E.F., 1976. Efficient capital markets: reply. *The Journal of Finance* 31, 143-145.

Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3-56.

Fama, E.F. and French, K.R., 1996. The CAPM is wanted, dead or alive. *The Journal of Finance* 51, 1947-1958.

Fama, E.F. and French, K.R., 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116, 1-22.

Fama, E.F. and French, K.R., 2020. Comparing cross-section and time-series factor models. *The Review of Financial Studies* 33, 1891-1926.

Filzen, J.J., 2015. The information content of risk factor disclosures in quarterly reports. *Accounting Horizons* 29, 887-916.

Gebhardt, W.R., Lee, C.M. and Swaminathan, B., 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39, 135-176.

Gode, D. and Mohanram, P., 2003. Inferring the cost of capital using the Ohlson–Juettner model. *Review of Accounting Studies* 8, 399-431.

Green, J., Hand, J.R. and Zhang, X.F., 2017. The characteristics that provide independent information about average US monthly stock returns. *The Review of Financial Studies* 30, 4389-4436.

Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Grundy, B.D. and Petry, S., 2022. Understanding Risk Disclosures and Exposures: Insights from a Novel Measure of Information Content. *Available at SSRN 4057636*.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. *The Review of Financial Studies* 33, 2223–2273.

Harvey, C.R., Liu, Y., Zhu, H., 2016. … and the cross-section of expected returns. *The Review of Financial Studies* 29, 5–68.

Heinle, M.S. and Smith, K.C., 2017. A theory of risk disclosure. *Review of Accounting Studies* 22, 1459-1491.

Heinle, M.S., Smith, K.C. and Verrecchia, R.E., 2018. Risk-factor disclosure and asset prices. *The Accounting Review* 93, 191-208.

Hope, O.K., Hu, D. and Lu, H., 2016. The benefits of specific risk-factor disclosures. *Review of Accounting Studies* 21, 1005-1045.

Hou, K., Van Dijk, M.A. and Zhang, Y., 2012. The implied cost of capital: A new approach. *Journal of Accounting and Economics* 53, 504-526.

Hou, K., Xue, C. and Zhang, L., 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies* 28, 650-705.

Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. *The Review of Financial Studies* 33, 2019–2133.

Huang, A.H., Lehavy, R., Zang, A.Y. and Zheng, R., 2018. Analyst information discovery and interpretation roles: A topic modeling approach. *Management Science* 64, 2833-2855.

Huang, A.H., Shen, J. and Zang, A.Y., 2022. The unintended benefit of the risk factor mandate of 2005. *Review of Accounting Studies* 27, 1319-1355.

Huang, A.H., Wang, H., Yang, Y., 2022. FinBERT: a deep learning approach to extracting textual information. *Working paper*.

IRRC Institute. 2016. The corporate risk factor disclosure landscape. *Available at: https:// www.weinb erg. udel.edu/IIRCi Resea rchDo cuments/2016/ 01/ FINAL- EY- Risk- Disclosure- Study.pdf.*

Kaustia, M. and Rantala, V., 2021. Common analysts: method for defining peer firms. *Journal of Financial and Quantitative Analysis* 56, 1505-1536.

Kelly, B.T., Pruitt, S., Su, Y., 2019. Characteristics are covariances: a unified model of risk and return. *Journal of Financial Economics* 134, 501–524.

Kim, J.B., Wang, C. and Wu, F., 2023. The real effects of risk disclosures: evidence from climate change reporting in 10-Ks. *Review of Accounting Studies* 28, 2271-2318.

Kozak, S., Nagel, S., Santosh, S., 2020. Shrinking the cross-section. *Journal of Financial Economics* 135, 271–292.

Kravet, T. and Muslu, V., 2013. Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies* 18, 1088-1122.

Lee, C.M., Ma, P. and Wang, C.C., 2015. Search-based peer firms: Aggregating investor perceptions through internet co-searches. *Journal of Financial Economics* 116, 410-431.

Lee, C.M., Sun, S.T., Wang, R. and Zhang, R., 2019. Technological links and predictable returns. *Journal of Financial Economics* 132, 76-96.

Lee, C.M., So, E.C. and Wang, C.C., 2021. Evaluating firm-level expected-return proxies: implications for estimating treatment effects. *The Review of Financial Studies* 34, 1907-1951.

Lee, C.M., Shi, T.T., Sun, S.T. and Zhang, R., 2024. Production complementarity and information transmission across industries. *Journal of Financial Economics* 155, 103812.

Lewellen, J., 2015. The cross section of expected stock returns. *Critical Finance Review* 4, 1–44.

Li, K.K. and Mohanram, P., 2014. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies* 19, 1152-1185.

Li, N., 2024. Labor market peer firms: understanding firms' labor market linkages through employees' internet "also viewed" firms. *Review of Accounting Studies*, 1-52.

Lyle, M.R., Riedl, E.J. and Siano, F., 2023. Changes in Risk Factor Disclosures and the Variance Risk Premium. *Accounting Review* 98, 327-352.

Lyle, M.R. and Wang, C.C., 2015. The cross section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics* 116, 505-525.

McInnes, L., Healy, J. and Melville, J., 2020. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

Menchero, J., 2010. The Characteristics of Factor Portfolios. *Journal of Performance Measurement*, vol. 15, no. 1 (Fall): 52-62.

Menchero, J., Orr, D.J. and Wang, J., 2011. The Barra US Equity Model (USE4) Methodology Notes, *MSCI Research Insight*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J., 2013. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26.

Nelson, K.K. and Pritchard, A.C., 2016. Carrot or stick? The shift from voluntary to mandatory disclosure of risk factors. *Journal of Empirical Legal Studies* 13, 266-297.

Securities and Exchange Commission (SEC), 2016. Concept release: Business and financial required by regulation S-K. Release No. 33–10064. *Available at https:// www.sec.gov/ rules/ concept/ 2016/ 33-10064. pdf.*

Smith, K., 2024. Climate risk disclosure and risk sharing in financial markets. *Available at SSRN 4552385*.

Smith, K.C. and So, E.C., 2022. Measuring risk information. *Journal of Accounting Research* 60, 375-426.

Appendix: Variable Definitions

Variable	Description and computation
Description of RBP returns	
EW RBP	Equal-weighted returns of risk-based peer firms. Computed as $\frac{1}{10}\sum_{i=1}^{10} R_{i,t}$, where $R_{i,t}$ is the return of top-*i* similar peer firm in month *t*
CosW RBP	Weighted-average returns of risk-based peer firms (weighted by cosine similarity). Computed as $\frac{\sum_{i=1}^{10} w_{i,t} R_{i,t}}{\sum_{i=1}^{10} w_{i,t}}$, where $w_{i,t}$ is the cosine similarity between the focal firm and its top-*i* similar peer firm in month *t*
Financial ratios used in testing explainability of RBPs	
pb	Price-to-book multiple. Computed as market capitalization / total common equity available before the end of month *t*
evs	Enterprise value-to-sales multiples. Computed as (market capitalization + long-term debt) / net sales available before the end of month *t*
pe	Price-to-earnings multiple. Computed as market capitalization / net income before extraordinary items available before the end of month *t*
rnoa	Returns on net operating assets. Computed as net operating income after depreciation / (property, plant, and equipment + current assets - current liabilities) available before the end of month *t*
roe	Returns on equity. Computed as net income before extraordinary items / total common equity available before the end of month *t*
at	(Inverse of) Asset turnover. Computed as total assets / net sales available before the end of month *t*
pm	Profit margin. Computed as net operating income after depreciation / net sales available before the end of month *t*
lev	Leverage. Computed as long-term debt / total stockholder's equity available before the end of month *t*
gpr	Gross profitability. Computed as the sum of (sales in the latest fiscal quarter available before the end of month *t* + sales in the last three fiscal quarters) – sum of (cost of goods sold in the latest fiscal quarter available before the end of month *t* + cost of goods sold in the last three fiscal quarters) / the average of (total assets in the latest fiscal quarter available before the end of month *t* + total assets in the last four fiscal quarters)
rdpersales	Research and development expenses scaled by net sales. Computed as R&D expense / net sales available before the end of month *t*
Expected return proxies	
RB_PCA	Primary risk-based ERP. Computed as 12-month moving average of the fitted values obtained from monthly cross-sectional regressions of *RET1* on five (beginning-of-the-month) standardized principle components
LPV	Characteristics-based ERP based on the log-linear present-value relation in Chattopadhyay et al. (2022). Computed as $\mathbb{E}[R_{i,t+1}] = \beta_{i,0} + \beta_{i,1} BM_{i,t} + \beta_{i,2} ROE_{i,t} + \beta_{i,3} RVOL_{i,t} + \varepsilon_{i,t+1}$

JLR	Characteristics-based ERP based on SIZE, BM, and MOMENTUM in Lewellen (2015). Computed as $\mathbb{E}[R_{i,t+1}] = \beta_{i,0} + \beta_{i,1}SIZE_{i,t} + \beta_{i,2}BM_{i,t} + \beta_{i,3}MOMENTUM_{i,t} + \varepsilon_{i,t+1}$
CER	A composite measure that takes the equal-weighted average of LPV and JLR in Lee et al. (2021). Computed as *CER* = (*LPV* + *JLR*) / 2
FFC	Characteristics-based ERP based on SIZE, BM, OP, and INV in Fama and French (2015). Computed as $\mathbb{E}[R_{i,t+1}] = \beta_{i,0} + \beta_{i,1}SIZE_{i,t} + \beta_{i,2}BM_{i,t} + \beta_{i,3}OP_{i,t} + \beta_{i,4}INV_{i,t} + \varepsilon_{i,t+1}$
Characteristics to estimate ERPs	
RET1	Stock return in month t+1
RET2	Stock return in month t+2
RET3	Stock return in month t+3
BM	Most recent annual book-to-market ratio. Computed as common equity / market capitalization available before the end of month t
ROE	Most recent annual return on equity. Computed as income before extraordinary items / average common equity available before the end of month t
RVOL	Return volatility. Denoted by the standard deviation of daily stock returns in month t
SIZE	The natural logarithm of market capitalization in month t
MOMENTUM	The cumulative stock return from 12 months to 2 months prior to the estimation date
OP	Operating profitability. Computed as revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity available before the end of month t
INV	Investment. Denote by the rate of growth of total assets, $\ln(A_{t-1}/A_{t-2})$, from the last fiscal yearend in year $t-2$ to the last fiscal yearend in t-1

Figure 1: The number of topics across years

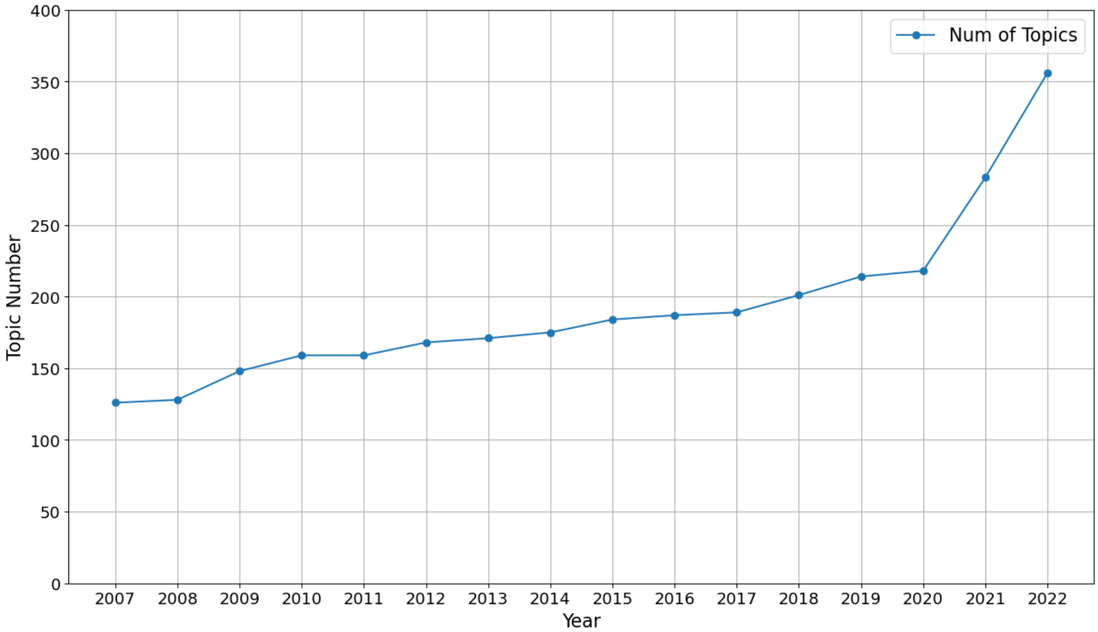

Figure 1 shows the number of topics in each year's topic modeling result. It can be observed that the number of topics each year increases steadily from 126 in 2007 to 218 in 2020, but jumps to 283 and 356 topics in the last two years.

Figure 2: Sample topics for Item 1A in year 2021-22

Figure 2 illustrates the word cloud of the 72 out of 363 topics that BERTopic produces based on the Item 1A paragraphs in 10-Ks filed between July 2021 and June 2022. Each word's size in the word cloud is based on the importance of the word to that topic, i.e., c-TF-IDF of the word.

Table 1: Sample paragraph and topic distribution by year

Year	(1) # Para.	(2) # topics	(3) # Para. per topic	(4) # Words per topic	(5) SD
2007	92.5	30.2	2.7	154.8	1.26
2008	86.9	31.4	2.5	158.3	1.07
2009	94.2	33.8	2.6	169.8	0.87
2010	99.3	37.7	2.5	168.0	1.36
2011	102.9	38.1	2.6	176.1	1.42
2012	110.6	41.1	2.5	171.8	0.78
2013	115.6	42.8	2.6	174.9	1.01
2014	121.9	44.8	2.6	176.5	0.76
2015	128.4	46.4	2.6	180.7	0.78
2016	135.9	48.6	2.6	184.8	0.78
2017	147.3	50.4	2.7	193.2	0.84
2018	156.9	55.2	2.7	190.1	0.83
2019	166.4	57.4	2.7	194.6	0.77
2020	173.4	61.2	2.7	194.1	0.79
2021	189.4	66.4	2.7	195.5	0.74
2022	195.3	72.6	2.5	184.8	0.62
Overall	131.4	47.1	2.6	178.9	0.95

Table 1 presents sample paragraph and topic distribution by year. Column (1) and (2) show the average number of paragraphs and topics covered in Item 1A across years. Column (3) provides the average number of paragraphs within each topic. Column (4) reports the average number of words within each topic. Column (5) reports the standard deviation of the number of paragraphs within each topic.

Table 2: Average cosine similarity in risk factor exposures and industry correspondence of peer firms ranked by cosine similarity

Panel A: Average cosine similarity across all industries and within each industry classification

Peer firm rank	(1) Average Similarity	(2) Same GICS6	(3) Same GICS2	(4) Same SIC4	(5) Same SIC3	(6) Same NAICS	(7) Same NAICS3
1	0.832	0.798	0.816	0.827	0.813	0.829	0.799
2	0.812	0.767	0.792	0.799	0.784	0.803	0.769
3	0.800	0.748	0.777	0.781	0.765	0.784	0.749
4	0.791	0.733	0.765	0.765	0.750	0.769	0.733
5	0.784	0.720	0.755	0.752	0.736	0.756	0.719
6	0.777	0.708	0.746	0.741	0.723	0.743	0.706
7	0.771	0.697	0.739	0.729	0.711	0.731	0.694
8	0.765	0.687	0.732	0.718	0.699	0.718	0.682
9	0.760	0.676	0.725	0.704	0.685	0.704	0.670
10	0.756	0.665	0.718	0.689	0.671	0.688	0.658
Average	0.785	0.720	0.756	0.751	0.734	0.752	0.718

Panel B: Average cosine similarity and industry correspondence

Peer firm rank	(1) Average Similarity	(2) Same GICS6	(3) Same GICS2	(4) Same SIC4	(5) Same SIC3	(6) Same NAICS	(7) Same NAICS3
1	0.832	0.536	0.772	0.398	0.496	0.319	0.560
2	0.812	0.501	0.750	0.375	0.469	0.300	0.537
3	0.800	0.486	0.741	0.360	0.453	0.287	0.523
4	0.791	0.475	0.736	0.350	0.440	0.279	0.511
5	0.784	0.465	0.731	0.343	0.433	0.273	0.506
6	0.777	0.455	0.725	0.336	0.426	0.267	0.500
7	0.771	0.450	0.722	0.332	0.421	0.262	0.495
8	0.765	0.445	0.719	0.326	0.413	0.257	0.488
9	0.760	0.435	0.713	0.321	0.408	0.254	0.483
10	0.756	0.433	0.711	0.315	0.402	0.250	0.478
Average	0.785	0.468	0.732	0.346	0.436	0.275	0.508

Table 2 Panel A reports the average cosine similarity of the RBPs and the top 10 peer firms based on cosine similarity of risk factor exposure within the industry. Column (1) reports the average cosine similarity of RBPs. Columns (2)-(7) report the average cosine similarity of top 10 peer firms based on cosine similarity of risk factor exposure within the same GICS6, GICS2, SIC4, SIC3, NAICS, and NAICS3, respectively.
Table 2 Panel B reports the average cosine similarity and the industry correspondence for the RBPs, ranked by their cosine similarity. Column (1) reports the cosine similarity of RBPs. Columns (2)-(7) report the fraction of RBPs that share the same industry classification based on GICS6, GICS2, SIC4, SIC3, NAICS, and NAICS3, respectively.

Table 3: Average R^2 from monthly cross-sectional regressions

	(1) **EW RBP**	(2) **CosW RBP**	(3) **GICS6**	(4) **(1) – (3)**	(5) **(2) – (3)**
All firms **(N = 1,866)**	0.0364	0.0369	0.0232	0.0131*** (7.91)	0.0136*** (8.05)
Top half **(N = 933)**	0.0734	0.0745	0.0466	0.0268*** (10.33)	0.0280*** (10.38)
Bottom half **(N = 933)**	0.0240	0.0243	0.0163	0.0078*** (5.64)	0.0081*** (5.78)
Before 31/12/2014 **(N = 1,938)**	0.0259	0.0262	0.0164	0.0095*** (5.59)	0.0098*** (5.73)
After 31/12/2014 **(N = 1,799)**	0.0462	0.0469	0.0296	0.0166*** (6.52)	0.0173*** (6.67)
Recession Period **(N = 1,972)**	0.0407	0.0410	0.0336	0.0071** (2.15)	0.0074** (2.20)
Boom Period **(N = 1,852)**	0.0358	0.0363	0.0218	0.0140*** (7.97)	0.0145*** (8.12)
Number of months	186	186	186	186	186

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 3 presents the time-series average R^2 from monthly cross-sectional regressions of focal-firm returns on those of the peer firms in each calendar month. Columns (1), (2), and (3) report the results for the equal-weighted RBPs, weighted-average RBPs, and equal-weighted industry peers, respectively, and Columns (4) and (5) report the differences between the focal-firm return explainability of RBPs with that of industry peers. *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 4: Average R^2 from quarterly cross-sectional regressions

	(1) GICS6	(2) EW RBP	(3) CosW RBP	(4) (2) – (1)	(5) (3) – (1)
pb **(N = 1,274)**	0.0983	0.1467	0.1481	0.0484*** (11.44)	0.0498*** (11.51)
evs **(N = 1,274)**	0.3106	0.3718	0.3739	0.0613*** (7.66)	0.0633*** (8.00)
pe **(N = 938)**	0.1029	0.1284	0.1298	0.0255*** (5.41)	0.0269*** (5.77)
rnoa **(N = 1,273)**	0.0671	0.1474	0.1487	0.0803*** (11.82)	0.0816*** (11.95)
roe **(N = 1,274)**	0.0434	0.0940	0.0955	0.0506*** (8.33)	0.0521*** (8.58)
at **(N = 1,274)**	0.2804	0.4008	0.4041	0.1204*** (23.81)	0.1237*** (24.36)
pm **(N = 1,274)**	0.0515	0.1374	0.1394	0.0859*** (8.95)	0.0879*** (9.11)
lev **(N = 1,274)**	0.0844	0.1755	0.1769	0.0911*** (12.09)	0.0925*** (12.39)
gpr **(N = 1,274)**	0.3272	0.3643	0.3676	0.0371*** (6.74)	0.0404*** (7.33)
rdpersales **(N = 570)**	0.5599	0.5706	0.5723	0.0107 (0.98)	0.0124 (1.13)
Number of quarters	62	62	62	62	62

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 4 presents the time-series average R^2 from quarterly cross-sectional regressions of focal-firm financial ratios of interest on those of the peer firms in each calendar quarter (using the most updated ratios as of the end of March, June, September and December). Columns (1), (2), and (3) report the results for the equal-weighted industry peers, equal-weighted RBPs, and weighted-average RBPs, respectively, and Columns (4) and (5) report the differences between the focal-firm ratio explainability of RBPs with that of industry peers. *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 5: Descriptive statistics for actual and expected return prxoies

	Mean	SD	P1	P25	Median	P75	P99
RET1	0.0108	0.1510	-0.3624	-0.0606	0.0059	0.0723	0.4767
RB_ERP	0.0100	0.0174	-0.0418	0.0010	0.0114	0.0204	0.0515
LPV	0.0061	0.0111	-0.0336	0.0015	0.0083	0.0132	0.0243
JLR	0.0051	0.0116	-0.0259	-0.0011	0.0069	0.0130	0.0283
CER	0.0056	0.0103	-0.0266	0.0004	0.0074	0.0131	0.0227
FFC	0.0074	0.0081	-0.0166	0.0026	0.0087	0.0131	0.0246

Table 5 provides the descriptive statistics for the expected return proxies (ERPs) together with the one-month ahead realized stock return (RET1), using data from a sample of 175 calendar months during our 2008-2022 sample period. RB_ERP is the 12-month moving average of the fitted values obtained from monthly cross-sectional regressions of stock returns on five (beginning-of-the-month) standardized PCs, which is our primary risk-based ERP. Our first benchmark, LPV, is a version of the characteristics-based ERPs based on the log-linear present-value relation between a firm's current stock price and an accounting valuation anchor and its expected future growth, as developed by Chattopadhyay et al. (2022). We construct our second benchmark proxy, JLR, similar to Lee et al. (2021), which is based on three firm characteristics: SIZE (the logarithm of market capitalization), book-to-market ratio, and cumulative stock return from 12 months to 2 months prior to the estimation date (momentum). CER is a composite measure that takes the equal-weighted average of LPV and JLR. The fourth characteristics-based ERP benchmark, FFC, is estimated using the following characteristics: SIZE (the logarithm of market capitalization), BM (book-to-market ratio), OP (operating profitability=(revenue – cost of goods sold – selling, general and administrative expenses – interest expense)/average book value of common equity)), and INV (the rate of growth in total assets).

Table 6: Comparison between disclosure- and characteristics-based ERPs

Panel A: Performance comparison between risk- and characteristics-based ERPs

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
RB_ERP	0.0201**	0.4859***	0.0027**	0.0110***	0.0115***	0.0090**	0.0012
	(2.28)	(2.97)	(2.34)	(5.36)	(2.58)	(2.57)	(0.62)
LPV	0.0184*	0.6594*	0.0009	0.0184***	0.0023	0.0034	0.0057
	(1.69)	(1.94)	(0.52)	(8.49)	(0.38)	(0.54)	(1.60)
JLR	0.0116	0.3462	0.0009	0.0108***	0.0036	-0.0015	0.0032**
	(1.54)	(1.48)	(0.83)	(6.87)	(0.87)	(-0.25)	(2.21)
CER	0.0161	0.5841*	0.0009	0.0160***	0.0043	0.0037	0.0033
	(1.56)	(1.66)	(0.55)	(7.34)	(0.77)	(0.55)	(1.49)
FFC	0.0056	0.1939	-0.0001	0.0075***	0.0008	-0.0055	0.0022**
	(0.73)	(0.46)	(-0.04)	(6.80)	(0.19)	(-1.11)	(2.05)

Panel B: Performance of characteristics-based ERPs before 2008 (sample period: 197610-200712)

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
LPV	0.0314***	0.5461*	0.0029**	0.0139***	0.0095**	0.0092*	0.0012
	(4.30)	(1.77)	(2.14)	(8.57)	(2.21)	(1.92)	(0.67)
JLR	0.0472***	0.9107***	0.0063***	0.0126***	0.0234***	0.0188***	-0.0012
	(9.40)	(10.40)	(9.09)	(7.23)	(9.08)	(5.88)	(-0.62)
CER	0.0490***	1.2386***	0.0060***	0.0157***	0.0214***	0.0199***	-0.0022*
	(8.16)	(7.79)	(6.78)	(7.57)	(6.98)	(5.48)	(-1.65)
FFC	0.0384***	0.5973***	0.0044***	0.0096***	0.0147***	0.0099***	-0.0021*
	(6.65)	(3.44)	(4.80)	(9.04)	(4.56)	(2.86)	(-1.74)

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 6 presents the performance metrics and *t*-statistics of risk- and characteristics-based ERPs from monthly cross-sectional regressions of RET1 on these ERPs. Column (1) reports the time-series average monthly correlation coefficient between ERPs and RET1. Column (2) and (3) report the raw and standardized time-series average slope coefficients of the monthly cross-sectional regressions. Column (4) reports the time-series average R^2 of the regressions. Column (5) and (6) report the equal- and value-weighted returns to the ERP-sorted hedge portfolios. Column (7) reports the scaled measurement error variance (MEV) of various ERPs as defined by Lee et al. (2021). Panel A reports results in our sample from 2008 to 2022. To reconcile with previous studies, we generate the characteristics-based ERPs following the same procedures as our main analysis using data before our sample period of 2008. The results reported in Panel B show that characteristics-based ERPs perform substantially better in this sample period. *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 7: Subsample performance

Panel A: Subsample performance of risk-based ERP

	(1) **Corr**	(2) **Slope^{Raw}**	(3) **Slope^{Stdz}**	(4) **R^2**	(5) **Hgret^{EW}**	(6) **Hgret^{VW}**	(7) **MEV**
All	0.0201**	0.4859***	0.0027**	0.0110***	0.0115***	0.0090**	0.0012
(N = 175)	(2.28)	(2.97)	(2.34)	(5.36)	(2.58)	(2.57)	(0.62)
Top half	0.0183*	0.4087**	0.0017	0.0197***	0.0075*	0.0083**	0.0028
(N = 175)	(1.73)	(2.25)	(1.54)	(7.16)	(1.87)	(2.35)	(1.61)
Bottom half	0.0207**	0.5145***	0.0033***	0.0096***	0.0140***	0.0101**	-0.0000
(N = 175)	(2.43)	(2.98)	(2.61)	(4.75)	(2.82)	(2.08)	(-0.01)
Recession	0.0036	0.2370	0.0000	0.0121***	0.0016	0.0012	0.0109
(N = 16)	(0.15)	(0.43)	(0.00)	(3.64)	(0.11)	(0.10)	(1.52)
Booming	0.0217**	0.5110***	0.0030**	0.0109***	0.0125***	0.0098***	0.0002
(N = 159)	(2.33)	(3.06)	(2.51)	(4.89)	(2.70)	(2.70)	(0.13)

Panel B: Subsample performance of CER

	(1) **Corr**	(2) **Slope^{Raw}**	(3) **Slope^{Stdz}**	(4) **R^2**	(5) **Hgret^{EW}**	(6) **Hgret^{VW}**	(7) **MEV**
All	0.0161	0.5841*	0.0009	0.0160***	0.0043	0.0037	0.0033
(N = 175)	(1.56)	(1.66)	(0.55)	(7.34)	(0.77)	(0.55)	(1.49)
Top half	0.0097	0.5622	0.0009	0.0271***	0.0031	0.0020	0.0017
(N = 175)	(0.89)	(1.42)	(0.44)	(8.05)	(0.68)	(0.36)	(1.48)
Bottom half	0.0232***	0.8830**	0.0023	0.0127***	0.0077	0.0036	0.0020
(N = 175)	(2.58)	(2.45)	(1.34)	(6.76)	(1.37)	(0.58)	(0.90)
Recession	-0.0111	-0.2346	-0.0060	0.0316***	-0.0165	-0.0320	0.0214
(N = 16)	(-0.22)	(-0.16)	(-0.57)	(5.56)	(-0.47)	(-0.80)	(1.23)
Booming	0.0189*	0.6665*	0.0016	0.0145***	0.0063	0.0073	0.0015
(N = 159)	(1.90)	(1.92)	(1.16)	(6.59)	(1.36)	(1.19)	(1.13)

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.

Table 7 presents the subsample performance metrics and *t*-statistics of risk- and characteristics-based ERPs from monthly cross-sectional regressions of RET1 on these ERPs. Panel A reports the results for our characteristics-based ERP. Column (1) reports the time-series average monthly correlation coefficient between ERPs and RET1. Column (2) and (3) report the raw and standardized time-series average slope coefficients of the monthly cross-sectional regressions. Column (4) reports the time-series average R^2 of the regressions. Column (5) and (6) report the equal- and value-weighted returns to the ERP-sorted hedge portfolios. Column (7) reports the scaled measurement error variance (MEV) of various ERPs as defined by Lee et al. (2021). Top and bottom half represent large and small firms (partitioned on annual median). We also try to partition our sample into recession and booming period according to NBER business cycle classification. Panel B report the results for CER for comparison. *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 8: Performance evaluation using longer-term returns

Panel A: ERPs evaluation using RET2

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
RB_ERP	0.0154*	0.3826**	0.0023*	0.0116***	0.0094**	0.0066*	0.0018
	(1.73)	(2.38)	(1.95)	(5.58)	(2.16)	(1.74)	(0.91)
LPV	0.0169	0.5126	0.0002	0.0170***	-0.0010	-0.0003	0.0059*
	(1.56)	(1.40)	(0.09)	(9.45)	(-0.15)	(-0.05)	(1.78)
JLR	0.0069	0.1935	-0.0002	0.0105***	0.0004	-0.0025	0.0049***
	(0.83)	(0.74)	(-0.15)	(7.09)	(0.09)	(-0.35)	(2.60)
CER	0.0121	0.3563	-0.0002	0.0154***	-0.0000	0.0013	0.0043*
	(1.11)	(0.89)	(-0.13)	(7.83)	(-0.00)	(0.18)	(1.85)
FFC	0.0048	0.1056	-0.0006	0.0070***	-0.0028	-0.0053	0.0027**
	(0.63)	(0.26)	(-0.41)	(7.05)	(-0.62)	(-1.22)	(2.26)

Panel B: ERPs evaluation using RET3

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
RB_ERP	0.0156*	0.3516**	0.0021*	0.0114***	0.0100**	0.0081**	0.0018
	(1.80)	(2.19)	(1.86)	(5.67)	(2.35)	(2.20)	(0.89)
LPV	0.0108	0.3633	-0.0006	0.0164***	-0.0032	-0.0028	0.0075**
	(0.98)	(0.97)	(-0.31)	(8.39)	(-0.46)	(-0.42)	(1.98)
JLR	0.0037	0.0768	-0.0006	0.0098***	-0.0013	-0.0027	0.0050***
	(0.49)	(0.31)	(-0.44)	(6.99)	(-0.29)	(-0.38)	(2.86)
CER	0.0066	0.1923	-0.0009	0.0147***	-0.0027	-0.0022	0.0051**
	(0.63)	(0.50)	(-0.50)	(7.38)	(-0.43)	(-0.31)	(2.10)
FFC	0.0059	0.1689	-0.0003	0.0066***	-0.0021	-0.0046	0.0026**
	(0.83)	(0.46)	(-0.25)	(6.92)	(-0.46)	(-0.94)	(2.07)

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.

Table 8 presents the robustness of the results to longer-term returns. Panel A and B report the performance metrics and *t*-statistics of risk- and characteristics-based ERPs towards realized return in t+2 (RET2) and t+3 (RET3). Column (1) in reports the time-series average monthly correlation coefficient between ERPs and RET2 in Panel A (RET3 in Panel B). Column (2) and (3) report the raw and standardized time-series average slope coefficients of the monthly cross-sectional regressions of RET2 and RET3 on these ERPs. Column (4) reports the time-series average R^2 of the regressions. Column (5) and (6) report the equal- and value-weighted returns to the ERP-sorted hedge portfolios. Column (7) reports the scaled measurement error variance (MEV) of various ERPs as defined by Lee et al. (2021). *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 9: Risk-based ERP using alternative moving average windows

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
RB_ERP_MA(9)	0.0184**	0.4376***	0.0025**	0.0116***	0.0103**	0.0100***	0.0033
	(2.07)	(2.60)	(2.01)	(5.92)	(2.35)	(3.26)	(1.51)
RB_ERP_MA(6)	0.0165*	0.2880*	0.0022*	0.0106***	0.0092**	0.0070**	0.0048*
	(1.85)	(1.68)	(1.78)	(6.19)	(2.07)	(2.14)	(1.95)

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 9 presents the performance of risk-based ERP over alternative moving average windows of 9 and 6 months. Column (1) reports the time-series average monthly correlation coefficient between ERPs and RET1. Column (2) and (3) report the raw and standardized time-series average slope coefficients of the monthly cross-sectional regressions of RET1 on these ERPs. Column (4) reports the time-series average R^2 of the regressions. Column (5) and (6) report the equal- and value-weighted returns to the ERP-sorted hedge portfolios. Column (7) reports the scaled measurement error variance (MEV) of various ERPs as defined by Lee et al. (2021). *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 10: Risk-based ERPs using different numbers of principal components

	(1) Corr	(2) $Slope^{Raw}$	(3) $Slope^{Stdz}$	(4) R^2	(5) $Hgret^{EW}$	(6) $Hgret^{VW}$	(7) MEV
RB_ERP_3	0.0142**	0.6328***	0.0020**	0.0086***	0.0061*	0.0053**	0.0017
	(2.01)	(3.12)	(2.10)	(7.13)	(1.70)	(2.12)	(1.27)
RB_ERP_10	0.0219**	0.4071***	0.0029**	0.0123***	0.0119***	0.0076**	0.0022
	(2.53)	(3.04)	(2.54)	(5.67)	(2.68)	(2.02)	(1.02)
RB_ERP_20	0.0222**	0.3311***	0.0029**	0.0131***	0.0110**	0.0092**	0.0038
	(2.53)	(2.78)	(2.49)	(5.57)	(2.48)	(2.10)	(1.48)
RB_ERP_30	0.0227**	0.3066***	0.0029**	0.0134***	0.0101**	0.0068	0.0052*
	(2.55)	(2.65)	(2.44)	(5.80)	(2.27)	(1.38)	(1.89)

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 10 presents the performance of risk-based ERP using different numbers of principal components of 3, 10, 20, and 30 respectively. Column (1) reports the time-series average monthly correlation coefficient between ERPs and RET1. Column (2) and (3) report the raw and standardized time-series average slope coefficients of the monthly cross-sectional regressions of RET1 on these ERPs. Column (4) reports the time-series average R^2 of the regressions. Column (5) and (6) report the equal- and value-weighted returns to the ERP-sorted hedge portfolios. Column (7) reports the scaled measurement error variance (MEV) of various ERPs as defined by Lee et al. (2021). *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.

Table 11: Incremental Usefulness of Disclosure-Based Risk Factor Exposures

Panel A: Impact of adding risk-based characteristics on characteristics based ERPs

	Adding first five PCs		Adding all individual risk exposures	
	(1) **Change in R-square**	**(2)** ***t*-stat.**	**(3)** **Change in R-square**	**(4)** ***t*-stat.**
LPV	0.0048***	(3.28)	0.0031*	(1.79)
JLR	0.0049***	(3.03)	0.0048**	(2.46)
CER	0.0048***	(3.03)	0.0032*	(1.79)
FFC	0.0063***	(3.87)	0.0063***	(3.24)

Panel B: Regression of realized return on LPV characteristics and risk components of ERPs

	Risk: First five PCs		Risk: all individual risk factor exposures	
	(1) **Coeff.**	**(2)** ***t*-stat.**	**(3)** **Coeff.**	**(4)** ***t*-stat.**
Intercept	0.0105***	(2.81)	0.0085**	(2.01)
Characteristics component	0.3095*	(1.84)	0.3820*	(1.89)
Risk component	0.3478*	(1.74)	0.1959***	(2.63)
Avg R-square	0.0217***	(8.02)	0.0196***	(7.54)
Avg N	2,009		2,009	

Panel C: Regression of realized return on JLR characteristics and risk components of ERPs

	Risk: First five PCs		Risk: all individual risk factor exposures	
	(1) **Coeff.**	**(2)** ***t*-stat.**	**(3)** **Coeff.**	**(4)** ***t*-stat.**
Intercept	0.0089	(1.57)	0.0064	(1.13)
Characteristics component	0.1732	(1.09)	0.1159	(0.61)
Risk component	0.4247**	(2.34)	0.2205***	(3.20)
Avg R-square	0.0196***	(8.07)	0.0187***	(7.68)
Avg N	2,009		2,009	

Panel D: Regression of realized return on CER characteristics and risk components of ERPs

	Risk: First five PCs		**Risk: all individual risk factor exposures**	
	(1)	**(2)**	**(3)**	**(4)**
	Coeff.	***t*-stat.**	**Coeff.**	***t*-stat.**
Intercept	0.0098**	(2.02)	0.0077	(1.50)
Characteristics component	0.3288	(1.61)	0.2208	(0.93)
Risk component	0.3923**	(2.05)	0.2086***	(2.91)
Avg R-square	0.0217**	(8.42)	0.0198***	(7.95)
Avg N	2,009		2,009	

Panel E: Regression of realized return on FFC characteristics and risk components of ERPs

	Risk: First five PCs		**Risk: all individual risk factor exposures**	
	(1)	**(2)**	**(3)**	**(4)**
	Coeff.	***t*-stat.**	**Coeff.**	***t*-stat.**
Intercept	0.0091	(1.46)	0.0068	(1.16)
Characteristics component	0.3641**	(1.96)	0.2616	(1.12)
Risk component	0.4378**	(2.55)	0.2002***	(2.72)
Avg R-square	0.0189***	(8.11)	0.0178***	(7.72)
Avg N	2,009		2,009	

*, **, *** Indicate significance level of 10 percent, 5 percent, and 1 percent, respectively.
Table 11 Panel A presents the effect of adding risk-based characteristics on improving the R^2 of the cross-sectional return-ERPs regressions as in Chattopadhyay (et al. 2022). Column (1) and (2) of Panel A report the increase in R^2 of the cross-sectional return-ERPs regressions resulting from adding the risk-based characteristics (i.e. first five PCs of risk factor exposures). Column (3) and (4) of Panel A report the increase in R^2 when simply adding all the individual risk factor exposures (instead of their lower dimensional representation, i.e. 5 PCs) to the four characteristics-based ERPs. Panel B-E presents the association of decomposed LPV, JLR, CER, and FFC with future realized stock return in the same regression. Column (1) and (2) report the association of characteristics components and the (reduced dimensional) risk component of ERP with future realized stock return. Column (3) and (4) report the association of characteristics components and the original high dimensional risk component of ERP with future realized stock return. *t*-statistics in brackets are estimated using Newey-West-corrected standard errors.