Conditional Asset Pricing with Text-Managed Portfolios*

Jian Feng[†] Jiantao Huang[‡] Shiyang Huang[§] Ran Shi[¶]

March 2026

Abstract

We investigate the asset pricing implications of soft, qualitative information extracted from earnings conference calls. Augmenting standard firm characteristics with text-based signals substantially enhances mean-variance efficiency but adds little to stock-level return predictability. Our evidence suggests a covariance channel: earnings calls provide incremental information about return covariances beyond what is captured by firm characteristics alone. Text explains roughly one-third of common variation in stock returns, particularly among high-growth and intangible-intensive firms where hard information is least informative. Textual topics account for a growing share of the conditional mean-variance-efficient portfolio weights, surpassing the role of conventional firm characteristics in recent years.

Keywords: factor models, mean-variance efficiency, conditioning information, text data.

JEL Classification Codes: C11, G11, G12.

* Any errors or omissions are the responsibility of the authors. For helpful comments, discussions, and suggestions, we thank Antoine Didisheim, Charles Martineau, Jack Strauss, Yinan Su and seminar and conference participants at University of Colorado Boulder, Renmin University of China, Front Range Finance Seminar, Finance Down Under 2025, 7th Future of Financial Information Conference, 17th Annual SoFiE Conference, and Five-Star Workshop in Finance.

† HSBC Business School, Peking University; jianfeng95@phbs.pku.edu.cn.

‡ Faculty of Business and Economics, The University of Hong Kong; huangjt@hku.hk.

§ Faculty of Business and Economics, The University of Hong Kong; huangsy@hku.hk.

¶ University of Colorado Boulder; ran.shi@colorado.edu.

Empirical asset pricing literature has documented a broad set of firm characteristics that can predict stock returns and their dynamic risk exposures. These characteristics are often used to construct *managed portfolios* that dynamically tilt toward stocks expected to improve mean-variance efficiency,[1] and thus serve as a benchmark information set in conditional asset pricing models. Yet these conventional signals primarily capture hard information embedded in realized accounting outcomes and past market behavior, and may therefore miss *softer* signals that are inherently subjective, qualitative, and forward-looking, but nonetheless crucial for firm valuation. This gap highlights a vital, yet underexplored, role for soft and nuanced information in conditional asset pricing.

In this paper, we investigate a prominent source of such soft information—text from earnings conference calls—to construct *text-managed portfolios* and study their implications for conditional asset pricing.[2] Conference call text provides a unique window into evolving firm fundamentals. Prepared remarks convey qualitative insights into strategic priorities, demand trends, and governance, while unscripted manager–analyst interactions reveal beliefs about risks and future outlooks. Together, these features make conference call text a natural source of soft information that is incremental to standard firm-characteristic data.

We show that the soft information embedded in conference calls is not spanned by standard characteristics and hence provides incremental value for conditional asset pricing. First, augmenting the conditioning set with conference call text materially improves mean-variance efficiency, despite having little effect on expected-return forecasts. This improvement operates mainly through a covariance channel: relative to conditioning on characteristics alone, text allows us to better capture time-varying return comovement. Second, the soft information extracted from conference calls plays an increasingly pivotal role in the mean-variance-efficient portfolio, surpassing conventional firm characteristics in the most recent subsample. Third, text explains roughly one-third of the common variation in stock returns. Consistent

[1]For example, Fama and French (1993, 2015) and Hou et al. (2015) sort stocks on their characteristics to create factors. Rosenberg (1974) and Fama and French (2020) use cross-sectional return regressions on characteristics to construct least square slope factors.

[2]Our use of text to capture soft, qualitative information for firm valuation is motivated by previous literature demonstrating its incremental value in, for example, measuring linguistic tone to forecast firms' financial outcomes (Tetlock et al., 2008), creating new industry classification systems (Hoberg and Phillips, 2016), quantifying firm-specific exposures to political risk (Hassan et al., 2019), and capturing unpatented corporate innovations that standard R&D metrics fail to measure (Bellstam et al., 2021).

with text conveying qualitative and forward-looking signals, its explanatory power is most pronounced precisely where hard information is least informative—such as in industries with high intangible capital or rapid technological change, and among high-growth firms.

Our exercise is challenging for two reasons. The first challenge is to construct a conditional asset pricing model—using a rich set of characteristics, including those extracted from earnings conference calls—that can span the universe of individual stocks. Our solution builds on the theoretical results in Kozak and Nagel (2024). When linear combinations of characteristics are sufficient to explain both expected returns and conditional covariances, the instrumented principal component analysis (IPCA) of Kelly et al. (2019) can identify factors that summarize the investment opportunity set of individual stocks (at the population level). The insights from Kelly et al. (2019) allow us to interpret the factors as the leading principal components of managed portfolios constructed from both the conventional characteristics and our text-based ones.

Another implementation challenge is to construct firm-level quantitative signals from text that are both low-dimensional and economically interpretable. We seek a mapping from a firm's conference-call transcript $\mathcal{D}$ to a numeric vector $\boldsymbol{d}$. At one end of the spectrum, modern neural-network-based language models produce compact document embeddings, but the elements of $\boldsymbol{d}$ are typically hard to interpret, and pre-training on large, time-aggregated corpora can introduce look-ahead bias (Glasserman and Lin, 2023; Sarkar and Vafa, 2024).[3] At the other end, conventional bag-of-words encodings based on word or *multi*-gram frequencies are transparent but ultra-high dimensional. We adopt an intermediate approach based on the topic model of Blei et al. (2003): topic modeling reduces dimensionality while preserving interpretability, as each element of $\boldsymbol{d}$ captures the firm's loading on a latent topic.

We estimate topics separately for the scripted presentation and the more interactive Q&A sessions. We optimally select 32 topics from each corpus, yielding a total of 64 conference call topics. Most topics are highly interpretable and can be categorized into three main themes: general topics (e.g., "Economic Uncertainty" and "Supply Chain"), industries (e.g., "Pharmaceutics" and "Oil"), and topics related to technology vintages (e.g., "Cloud Service"

[3] This issue can be addressed if researchers (pre-)train their own language models using time-stamped expanding text corpus instead of using the off-the-shelf ones (see, e.g., Engelberg et al. (2025) and He et al. (2025)), but the resulting $\boldsymbol{d}$ vectors can be more difficult to interpret.

and "Renewable Energy"). The presentation and Q&A sessions share common topics such as "Inflation" and various industries. However, some distinctive topics emerge in the Q&A sessions, such as "Forecasting and Estimation," "Growth Outlook," and "Optimism," while the presentation sessions uniquely emphasize topics like "Profits," "Organization Management," and "Business Operation." From a narrative perspective, these topics capture dimensions of the firm that are distinct from conventional characteristics.

Given that conference call text captures sentiment-rich information—such as the "Optimism" topic in Q&A—we examine whether these topic loadings contain cross-sectional information about exposure to investor sentiment. In the spirit of Baker and Wurgler (2006), harder-to-value firms should be more vulnerable to sentiment-driven mispricing than others. Consistent with this view, we find that, in high-sentiment periods, firms speaking more optimistically in Q&A become temporarily more overvalued and experience stronger subsequent negative price reversals. Conversely, firms providing more extensive fundamental discussion and clearer forward guidance—indicated by higher loadings on topics such as "Business Operation," "Forecasting and Estimation," and "Growth Outlook"—experience significantly smaller sentiment-driven reversals.

Firms' topic loadings exhibit significant cross-sectional variation, which we use as conditioning variables for dynamic factor loadings in the IPCA model. To evaluate the asset pricing value of this textual data, we append these 64 topic loadings to the 35 firm characteristics studied by Kelly et al. (2019), and examine the out-of-sample performance of the IPCA model conditional on this expanded set of cross-sectional information. Our full sample spans 2003 to 2022. To ensure that the initial in-sample window contains at least 60 monthly observations, our out-of-sample period begins in 2008. This analysis delivers four main implications for conditional asset pricing models.

First, conference call textual information significantly improves mean-variance efficiency, although it does not offer much help in explaining time-varying expected stock returns. In the out-of-sample period, the IPCA models conditional on the 32 presentation, 32 Q&A, or the 64 combined topic loadings can explain around 24%/0.14% of the variation in realized/expected returns. In comparison, the numbers are 27% and 0.15% for the 35 characteristics from Kelly et al. (2019) (KPS). Adding the conference call topic loadings to the firm characteristics does not lead to significantly higher predictive power for future stock returns. Next, we construct tangency portfolios of the IPCA factors and report the out-of-sample Sharpe ratio (SR). The tangency portfolios conditional on either Q&A or presentation topic loadings achieve annualized out-of-sample SRs of around 1.0–1.2. Combining both sources of topic loadings increases the out-of-sample SR to 1.64, comparable to that of the MVE portfolio conditional on the 35 KPS characteristics (SR $\approx$ 1.7). Most strikingly, adding the 64 topic loadings to the 35 KPS characteristics further improves the out-of-sample SR to over 2.6.

Second, we show that the efficiency gains from textual information operate primarily through the second-moment (covariance) channel, rather than the first-moment (expected return) channel. To test this mechanism, we compare the performance of long-short (5–1) quintile portfolios sorted on IPCA-implied expected returns with those sorted on conditional MVE portfolio weights. Augmenting standard firm characteristics with textual topic loadings yields only modest improvements in the first-moment channel, increasing the Sharpe ratio of the expected-return-sorted 5–1 portfolio from 0.55 to 0.88. By contrast, sorting on MVE weights—which further exploit the conditional covariance matrix to maximize diversification—generates a substantial performance increase, raising the Sharpe ratio of the 5–1 portfolio from 1.11 to 1.74.

Consistent with this covariance channel, the textual topics extract information largely orthogonal to traditional characteristics. Specifically, the correlation between MVE portfolios managed by textual topics and those managed by characteristics is strikingly low, ranging from 0.24 to 0.31 for Q&A topics and from -0.04 to 0.17 for presentation topics. These findings align with the theoretical insights from Kozak and Nagel (2024). Expanding the conditioning information set with conference call text leads to higher mean-variance efficiency, even though the text adds little incremental predictive power for expected returns. This suggests that conference call text captures variation in return covariances that is overlooked by traditional characteristics, highlighting the informational benefits of analyzing large-scale unstructured financial text data beyond its value for return prediction.

Third, we assess the relative importance of firm characteristics versus conference call text by decomposing the common return component captured by our conditional asset pricing model. We first estimate separate IPCA models using only characteristics and only text, generate two return forecasts, and then regress realized returns on both forecasts jointly. This yields a decomposition of the model's explanatory power into characteristic and textual components. Overall, the common component explains about 28% of the variation in stock returns, and text accounts for roughly one-third of this explained variation.

The contribution of text varies systematically across firms. It is largest in industries where fundamentals are more intangible, forward-looking, or difficult to summarize with standard characteristics, such as technology, healthcare, and business services, but much smaller in mature or regulated industries such as utilities. A similar pattern holds across firm styles: text plays a larger role for high-growth firms, as well as firms with strong recent returns or profitability. For these firms, the key question for investors is not simply whether current performance is strong, but whether it will persist or revert, which makes the qualitative and forward-looking information in conference calls especially valuable.

Fourth, conference call text is a key determinant of the conditional MVE portfolio, even after controlling for standard firm characteristics. In the decomposition of MVE portfolio weights, text explains about 40% of the cross-sectional variation in weights, compared with about 48% for characteristics. Several textual topics also stand out as especially important in the weight regressions, including "Business Operation," "Organizational Management," "International Finance," "Human Capital," and "Capital Structure." These topics capture nuanced, qualitative, and forward-looking disclosures that are difficult to measure using conventional accounting- or return-based characteristics, helping explain why text provides incremental information beyond those characteristics. Moreover, the role of text is not only economically large on average but also has been growing over time. In recent years, the share of MVE weight variation explained by conference-call text has risen above 50%, suggesting that soft information extracted from conference calls has become even more important than firm characteristics for portfolio allocation.

Concerns may be raised as our topic model estimation is performed on the whole corpus of earnings conference calls in our sample. To address these concerns, we use an online version of topic modeling to estimate real-time topic information in an expanding-window fashion. We confirm that the asset pricing implications—e.g., model fit, mean-variance efficiency, and the covariance channel—are virtually unchanged when using these online topics.

Related Literature. Our paper contributes to the literature on textual analysis in financial economics, especially in asset pricing. As a specific class of "big data" in finance (Goldstein et al., 2021), text has become an increasingly prevalent data source for empirical work, as surveyed in Gentzkow et al. (2019), Das et al. (2014) and Loughran and McDonald (2020). In the asset pricing literature, text data have been used to capture expected returns (Tetlock, 2007; Tetlock et al., 2008; García, 2013; Chen et al., 2022; Ke et al., 2022), create novel risk measures (Manela and Moreira, 2017; Hassan et al., 2019; Hanley and Hoberg, 2019; Giglio et al., 2024), conduct risk management (Engle et al., 2020), and recover asset pricing factors (Bybee et al., 2023; Lopez-Lira, 2023; Aleti and Bollerslev, 2024).

Closest to our work are Lopez-Lira (2023) and Bybee et al. (2023), which also exploit information in loadings estimated from topic models to instrument for the factor exposures. Lopez-Lira (2023) studies the implications of risk disclosures in annual reports (form 10-K) on identifying priced factors. Bybee et al. (2023) construct narrative factors based on firms' return covariances with respect to loadings on news topics. Quarterly earnings conference calls are more timely than annual 10-K filings and contain richer soft, qualitative, and forward-looking information, owing to their conversational format and interactive setting. These features make conference call text an ideal source of qualitative information that adds incremental value beyond conventional firm characteristics based on accounting variables and historical market prices.

The literature on conditional asset pricing dates back at least to the seminal theoretical work of Hansen and Richard (1987). Empirical studies include factor pricing that conditions on macroeconomic variables (Ferson and Harvey, 1999) and firm-specific characteristics (Kelly et al., 2019), the efficient use of conditioning information for mean-variance analysis (Ferson and Siegel, 2001), model testing and evaluation (Lewellen and Nagel, 2006; Nagel and Singleton, 2011), and conditional mean-variance spanning (Kozak and Nagel, 2024). We contribute to this literature by characterizing the incremental value of soft, qualitative information—contained in earnings conference calls—and by studying its implications for return prediction and mean-variance analysis.

Our use of text-managed portfolios to forecast aggregate outcomes echoes the works of Lamont (2001), Kelly and Pruitt (2013), and Bryzgalova et al. (2024), although our portfolios are not constructed to track any economic outcomes by design. The practice of building managed portfolios as traded proxies to economic "state variables" is related to the mimicking portfolio approach to estimating factor risk premiums (Huberman et al., 1987; Balduzzi and Robotti, 2008; Giglio and Xiu, 2021). Textual analysis of news data has also been proven useful for forecasting market returns and economic activities (Larsen and Thorsrud, 2019; Bybee et al., 2024; van Binsbergen et al., 2024; Chen et al., 2025). Different from our exercise, forecasting variables in this literature are not tradable asset returns.

The rest of the paper is organized as follows. Section 1 briefly summarizes the theoretical insights of Kelly et al. (2019) and Kozak and Nagel (2024) to clarify the impacts of adding text data to characteristics on factor spanning. Our estimation method and its theory foundations are then discussed. Section 2 introduces our data sources and sample construction. Section 3 describes the common topics in earnings conference calls. Section 4 presents our results on using the topic loadings to measure expected returns and construct spanning factors. Section 5 documents the forecasting power of these factors for macroeconomic outcomes. Section 6 investigates the role of qualitative textual information in the conditional mean-variance-efficient portfolio. Section 7 concludes.

1 Theory and Method

We review in this section the implications of enriching the conditioning information set (with, for example, firm-specific text documents in addition to the standard characteristics) for portfolio analysis. Special attention is paid to spanning the mean-variance efficient frontier with a relatively small number of factors. We then present our methodology for incorporating firm-specific text information to identify and estimate these factors.

To fix ideas, consider a panel of n realized stock returns in excess of the risk-free rate $\boldsymbol{r}_{t+1} = [r_{1,t+1}, \ldots, r_{n,t+1}]'$. Assume that stock i's characteristics and other text-based information (e.g., disclosure or news coverage), observed by the econometrician at time t, are concatenated into a vector $\boldsymbol{x}_{i,t}$ of dimension m. Define the $n \times m$ time-varying matrix $\boldsymbol{X}_t = [\boldsymbol{x}_{1,t}, \cdots, \boldsymbol{x}_{n,t}]'$ and denote by $\boldsymbol{\mu}_t = \mathbb{E}[\boldsymbol{r}_{t+1} \mid \boldsymbol{X}_t]$ the conditional expected returns and by $\boldsymbol{\Sigma}_t = \mathrm{var}[\boldsymbol{r}_{t+1} \mid \boldsymbol{X}_t]$ the conditional covariance matrix.

We are interested in the possibility of finding p ($p < n$) factors $\boldsymbol{f}_{t+1} \in \mathbb{R}^p$ that span the same investment opportunity set as the n stocks. That is, by investing the p assets defined by the factors, an investor can achieve the same maximal conditional Sharpe ratio as that attainable from the n individual stocks. We follow the exposition of Kozak and Nagel (2024) and assume that the conditional expected returns are linear in the vector of firm-specific information.

Assumption 1. *$\boldsymbol{\mu}_t = \boldsymbol{X}_t \boldsymbol{b}_t$ for some vector $\boldsymbol{b}_t \in \mathbb{R}^m$.*

The ideal benchmark. Under Assumption 1, one can always construct the factors that span the investment opportunities offered by the individual stocks. Define the factors as $\boldsymbol{f}_{t+1} = \boldsymbol{X}_t' \boldsymbol{\Sigma}_t^{-1} \boldsymbol{r}_{t+1}$. Since

$$\mathbb{E}[\boldsymbol{r}_{t+1} \mid \boldsymbol{X}_t] = \text{cov}[\boldsymbol{r}_{t+1},\ \boldsymbol{f}_{t+1} \mid \boldsymbol{X}_t] \times \boldsymbol{b}_t,$$

the factors $\boldsymbol{f}_t$ must achieve conditional mean-variance efficiency, and the number of factors equals the dimension of $\boldsymbol{x}_{i,t}$, i.e., $p = m$.

Managed portfolios. In practice, factors such as those in Fama and French (2015) and the MSCI Barra or Axioma factors are constructed as simple managed portfolios $\boldsymbol{f}_{t+1} = \boldsymbol{X}_t' \boldsymbol{r}_{t+1}$— and as transformations of these portfolios based on functions of $\boldsymbol{X}_t$. These portfolios, unlike the ideal benchmarks, can be constructed without resorting to the large covariance matrix $\boldsymbol{\Sigma}_t$, which is often difficult to estimate in empirical studies. The managed portfolios are not mean-variance efficient unless the following assumption about the covariance matrix holds (Kozak and Nagel, 2024).

Assumption 2. *$\boldsymbol{\Sigma}_t$ is such that $\boldsymbol{\Sigma}_t \boldsymbol{X}_t = \boldsymbol{X}_t \boldsymbol{B}_t$ for some nonsingular $m \times m$ matrix $\boldsymbol{B}_t$.*

A direct implication of Assumption 2 is that the $n \times m$ matrix of conditional covariances $\text{cov}[\boldsymbol{r}_{t+1},\ \boldsymbol{f}_{t+1} \mid \boldsymbol{X}_t]$ is a linear function of $\boldsymbol{X}_t$, when the factors are managed portfolios $\boldsymbol{f}_{t+1} = \boldsymbol{X}_t' \boldsymbol{r}_{t+1}$. Assumption 2 is also equivalent to a requirement that the stock returns' loadings on the factors that are not the managed portfolios—namely the omitted factors (Giglio and Xiu, 2021)—must be in the nullspace of $\boldsymbol{X}_t$ (Kozak and Nagel, 2024).

Assumptions 1 and 2 are more likely to hold when richer information is embedded into the vectors $\boldsymbol{x}_{i,t}$, $i = 1, \ldots, n$. As a result, the managed portfolios are more likely to span the investment opportunity set of the individual stocks when either the expected returns, the covariances, or both are better explained by $\boldsymbol{X}_t$. Firm characteristics are commonly used as explanatory variables in the empirical finance literature, and including more characteristics could help satisfy these two assumptions.

A concern, however, is that the newly added firm characteristics are likely to be highly correlated with the existing ones. Textual information, such as firm disclosures and news coverage, could be helpful by providing signals that are closer to orthogonal to existing characteristics. We focus on earnings conference calls in this paper. Before presenting our approach to incorporating earnings conference call information into $\boldsymbol{X}_t$, we discuss dimension reduction.

Dimension reduction. The managed portfolios still yield the same number of factors as the dimension of $\boldsymbol{x}_{i,t}$ ($p = m$). The following assumption ensures that the first p ($p \leq m$) principal components of the (transformed) managed portfolios could achieve the same mean-variance efficiency as investing directly in the n individual stocks.

Assumption 3. *The conditional mean* $\boldsymbol{\mu}_t$ *and covariance* $\boldsymbol{\Sigma}_t$ *are such that* $\boldsymbol{\mu}_t = \widetilde{\boldsymbol{X}}_t \boldsymbol{b}_t$ *and* $\boldsymbol{\Sigma}_t \widetilde{\boldsymbol{X}}_t = \widetilde{\boldsymbol{X}}_t \boldsymbol{B}_t$, *where* $\widetilde{\boldsymbol{X}}_t = \boldsymbol{X}_t \boldsymbol{\Gamma}$ *for some* $m \times p$ *matrix* $\boldsymbol{\Gamma}$.

If Assumption 3 holds, we can achieve dimension reduction by studying the following latent factor structure for individual stock returns:

$$\boldsymbol{r}_{t+1} = \boldsymbol{\beta}_t \boldsymbol{f}_{t+1} + \boldsymbol{\epsilon}_{t+1}, \quad \boldsymbol{\beta}_t = \boldsymbol{X}_t \boldsymbol{\Gamma}, \quad \text{cov}_t(\boldsymbol{\epsilon}_{t+1}) \boldsymbol{X}_t = \boldsymbol{0}, \tag{1}$$

where $\boldsymbol{\beta}_t$ denote the dynamic factor loadings instrumented by $\boldsymbol{X}_t$, and $\boldsymbol{\epsilon}_{t+1}$ are the error terms that are orthogonal to $\boldsymbol{X}_t$ and $\boldsymbol{f}_{t+1}$ and have zero conditional means. Once again, the condition $\text{cov}_t(\boldsymbol{\epsilon}_{t+1}) \boldsymbol{X}_t = \boldsymbol{0}$ is more likely to hold if more independent conditioning information $\boldsymbol{X}_t$ is included in the model (1).

We estimate the factors $\{\boldsymbol{f}_t\}_{t=1}^T$, as well as the matrix $\boldsymbol{\Gamma}$, using the instrumented principal component analysis (IPCA) estimator proposed by Kelly et al. (2019),

$$\{\widehat{\boldsymbol{\Gamma}},\ \widehat{\boldsymbol{f}}_1, \ldots, \widehat{\boldsymbol{f}}_T\} = \underset{\boldsymbol{\Gamma},\, \{\boldsymbol{f}_t\}_{t=1}^T}{\arg\min} \sum_{t=0}^{T-1} (\boldsymbol{r}_{t+1} - \boldsymbol{X}_t \boldsymbol{\Gamma} \boldsymbol{f}_{t+1})' (\boldsymbol{r}_{t+1} - \boldsymbol{X}_t \boldsymbol{\Gamma} \boldsymbol{f}_{t+1}). \tag{2}$$

The estimator can be interpreted as performing a principal component analysis on the (transformed) managed portfolios $(\boldsymbol{X}_t' \boldsymbol{X}_t)^{-1} \boldsymbol{X}_t' \boldsymbol{r}_{t+1}$, or the OLS "slope portfolios" from running a cross-sectional regression of the individual stock returns $\boldsymbol{r}_{t+1}$ on the conditional information matrix $\boldsymbol{X}_t$.

Summing up, expanding the information set beyond standard characteristics (e.g., market betas, valuation ratios, and technical signals) with alternative data such as text can help capture more variation in conditional expected returns and covariances. As a result, managed portfolios—and their transformations or principal components—are more likely to achieve conditional mean-variance efficiency in theory. Empirically, the IPCA estimator of Kelly et al. (2019) can be used to identify the factors derived from these managed portfolios.

As a remark, we stress that the statistical model described above cannot distinguish risk-based from mispricing-based explanations. Consistent with Kozak and Nagel (2024), comparisons among heuristic factor models mainly reveal limitations of factor models that ignore the covariance structure and, by themselves, do not resolve competing theories of risk premia versus mispricing.

1.1 Encoding the Text Information

Enriching the firm-specific information with text means working with vectors of the form

$$\boldsymbol{x}_{i,t} = [\text{characteristic}_{i,t},\ \text{text}_{i,t},\ \ldots],$$

for stock i at time t. It is straightforward to put firm characteristics into these vectors, but not so for text. Ideally, for any given text document $\mathcal{D}$, we would like to have a mapping $\mathcal{D} \mapsto \mathbb{R}^K$, from the document to a vector of dimension K. This is a process called *encoding*. In this section, we lay out our text encoding method for performing conditional asset pricing.

The simplest way to encode text is to count words: for a dictionary of W words, we obtain a W-dimension vector $\boldsymbol{d} = [d_1, \ldots, d_W]'$, where d_j is the count of word j in document $\mathcal{D}$. The dimension of $\boldsymbol{d}$ tends to be large (the same as the dictionary size), especially when bigrams, trigrams, or more sophisticated combinations of words are included in the dictionary.

To reduce the dimensionality of the encoding, we consider the following probabilistic specification: for the text document of firm i observed at time t, its word-count vector $\boldsymbol{d}_{i,t}$

is generated from a multinomial distribution

$$\boldsymbol{d}_{i,t} \sim \text{Multinomial}\left(L_{i,t}, \sum_{k=1}^{K} \phi_{i,t}^{(k)} \boldsymbol{\theta}_k\right), \tag{3}$$

where $L_{i,t} = \mathbf{1}'\boldsymbol{d}_{i,t}$ is the sum of word counts—the document length if we put only unigrams into the dictionary—in document $\mathcal{D}_{i,t}$. The vectors $\{\boldsymbol{\theta}_k\}_{k=1}^{K}$ denote different probability distributions over the dictionary, which can be interpreted as common *topics* shared across the document collection. Each topic represents a group of semantically related words that capture latent structures of the text data. For the text associated with firm i at time t, the *loading* on the k-th topic is given by $\phi_{i,t}^{(k)}$. In our applications, we aim to set $K \ll W$ and encode the document using the topic loading vector $\boldsymbol{\phi}_{i,t} = [\phi_{i,t}^{(1)}, \ldots, \phi_{i,t}^{(K)}]'$.

Summing up, instead of representing a document with a large vocabulary of words, we reduce the document to a low-dimensional vector of topics. A text document $\mathcal{D}_{i,t}$ for firm i observed at time t is first mapped to a sparse word count vector $\boldsymbol{d}_{i,t}$ and then to a low-dimensional vector $\boldsymbol{\phi}_{i,t}$. The elements in $\boldsymbol{\phi}_{i,t}$ measure the likelihood that document $\mathcal{D}_{i,t}$ is related to different topics.

Estimating the loadings $\boldsymbol{\phi}_{i,t}$ and the topic vectors $\{\boldsymbol{\theta}_k\}_{k=1}^{K}$ under equation (3) is analogous to performing a factor analysis of count data, because equation (3) implies

$$\mathbb{E}[\boldsymbol{d}_{i,t}] = \sum_{k=1}^{K} \left(L_{i,t}\phi_{i,t}^{(k)}\right) \boldsymbol{\theta}_k,$$

where the topic vectors serve the role of "common" factors, and the product terms in the parentheses are factor loadings. We adopt the well-known Bayesian approach of Blei et al. (2003) by assigning Dirichlet priors over $\boldsymbol{\phi}_{i,t}$ and performing variational inference. This implementation is also known as latent Dirichlet allocation (LDA). LDA is commonly used as a topic-modelling tool, but it can also be viewed as a form of dimensionality reduction that represents documents as mixtures of topics. This representation retains the main themes of the documents while filtering out noise from irrelevant or rare words that are unrelated to common topics. Focusing on common topics is natural also from an asset pricing perspective: topics that are pervasive across firms are candidates for common factors that drive asset returns, whereas topics that are specific to a single firm or a small set of firms primarily capture idiosyncratic rather than systematic risk.

We use the perplexity score to select the number of LDA topics. The traditional perplexity score, as described by Blei et al. (2003), is defined as

$$\text{Perplexity} = \exp\left\{ -\frac{\sum_{i,t} \log p(\boldsymbol{d}_{i,t})}{\sum_{i,t} L_{i,t}} \right\}, \tag{4}$$

where $L_{i,t}$ is the number of words in the document $\boldsymbol{d}_{i,t}$, and $p(\boldsymbol{d}_{i,t})$ is the likelihood of observing $\boldsymbol{d}_{i,t}$ given the LDA model. Overall, a lower perplexity score indicates a better fit of the LDA model.

2 Data Description

We obtain the conference call transcripts between 2003 and 2022 from Refinitiv StreetEvents, together with the exact date of the conference call and information on the conference speakers. A conference call typically consists of a presentation and a question-and-answer (Q&A) session. During the presentation segment, executives disclose information they wish to emphasize. The presentation segment is followed by a Q&A session in which outside investors can directly communicate with firm managers. Figure 2 displays examples from the presentation and Q&A sessions. Before estimating the LDA model, we first preprocess the text data. Details of the data-cleaning procedure are presented in Appendix A.

Next, we collect individual stock return data from the Global Factor Dataset provided by Jensen et al. (2023). We eliminate stocks whose market capitalizations fall below the first percentile of the NYSE market capitalization distribution. As explained in Section 1, we estimate the IPCA model of Kelly et al. (2019) using various sets of cross-sectional information as instruments. An important source of instruments is the collection of firm characteristics studied by Kelly et al. (2019). Since the sample used in Kelly et al. (2019) ends in May 2014, we reconstruct these firm characteristics and extend the sample through December 2022 using the Global Factor Data provided by Jensen et al. (2023). The list of 35 firm characteristics used in the empirical analysis is reported in Table A1, with exact matches in Panel A and close substitutes in Panel B. Moreover, we follow standard practice in the literature by normalizing firm characteristics cross-sectionally in each period (month). Specifically, we transform the raw characteristics into rank quantiles and scale them to lie

Table 1: Sample Description

Year	**Panel A.** Full sample			**Panel B.** Above 50th NYSE Firm Size		
	# of firms	# of firms holding conference calls	Coverage ratio	# of firms	# of firms holding conference calls	Coverage ratio
2003	3978	1573	0.40	959	590	0.62
2004	3777	1674	0.44	929	585	0.63
2005	3910	1843	0.47	929	625	0.67
2006	3807	1969	0.52	896	626	0.70
2007	3734	2041	0.55	874	641	0.73
2008	3660	2127	0.58	888	673	0.76
2009	3351	2113	0.63	853	664	0.78
2010	3216	2117	0.66	863	688	0.80
2011	3273	2225	0.68	871	727	0.83
2012	3117	2214	0.71	874	753	0.86
2013	3155	2183	0.69	911	796	0.87
2014	3221	2304	0.72	967	844	0.87
2015	3391	2450	0.72	995	887	0.89
2016	3101	2331	0.75	949	871	0.92
2017	3113	2398	0.77	952	884	0.93
2018	3189	2462	0.77	980	925	0.94
2019	3204	2474	0.77	999	952	0.95
2020	3374	2553	0.76	1090	1035	0.95
2021	3739	2720	0.73	1247	1172	0.94
2022	3674	2734	0.74	1186	1135	0.96

The table describes the sample of firms holding conference calls. We include all CRSP firms listed in the US on the NYSE, AMEX, or NASDAQ that have CRSP share codes of 10 or 11. Firms with market capitalizations below the first percentile of the NYSE market capitalization distribution are excluded from our analysis. Panel A reports the total number of firms, the number of firms holding at least one earnings conference call during the previous four quarters, and the corresponding coverage ratio in the last month of each year. Panel B repeats this analysis for the subsample of firms with market capitalizations above the 50th percentile of the NYSE market capitalization distribution.

in the interval [-0.5, 0.5].

After merging the conference call data with the sample of individual stocks, we obtain 185,184 unique earnings (conference) calls for the period 2003–2022. Panel A of Table 1 reports the total number of firms, the number of firms holding at least one earnings conference call within the previous four quarters, and the corresponding coverage ratio. The coverage ratios increase over time, from only 40% in December 2003 to above 70% during the last decade of the sample. Panel B of Table 1 additionally reports sample coverage for large firms whose market capitalizations are above the 50th percentile of the NYSE market capitalization distribution. Interestingly, large firms are more likely to hold earnings conference calls than small firms, with coverage in the most recent subsample reaching about 95%.

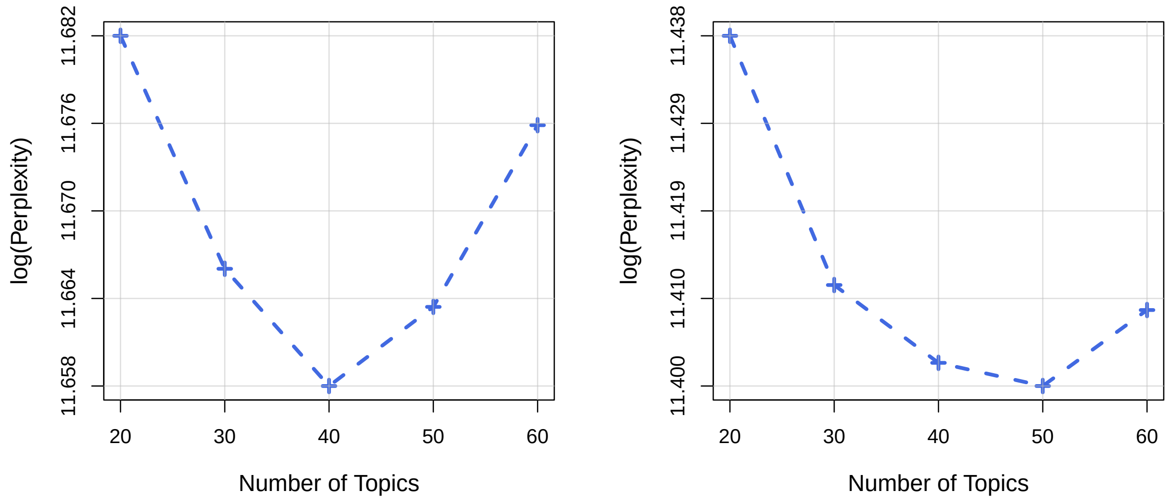
Finally, we use several aggregate economic variables. The sentiment index in Baker and Wurgler (2006) is obtained from the authors' websites. Other macroeconomic variables are taken from the FRED-MD dataset constructed by McCracken and Ng (2016).

3 Topics in Conference Calls

In this section, we summarize the results from the LDA estimation for both the Q&A and presentation sessions of conference calls. Specifically, we estimate the model using the "gensim" Python package, which requires choosing both the number of topics and the random seed for initializing the model parameters.

To determine the number of topics, we rely on the perplexity score defined in equation (4). We point out that the estimation results can differ across random seeds. Therefore, we repeat the LDA estimation ten times (using ten different random seeds) and report the average log-perplexity score for each number of topics. Figure 1 plots the average log-perplexity scores for different numbers of topics (20, 30, 40, 50, and 60). The figure shows that the perplexity score is lowest when the number of topics is 40 for the Q&A session and 50 for the presentation session. Therefore, in the LDA estimation we choose 40 and 50 topics for the Q&A and presentation sessions, respectively.

We now illustrate how we address the issue of random initialization in the LDA estimation. It is well known that the variational Bayesian method (as implemented in the "gensim" package) heavily relies on the random starting values. To ensure that our empirical results are not driven by a particular random seed, we repeat the LDA estimation using 10 different random seeds and identify the stable topics that appear in most estimates across initializations. We illustrate the intuition behind the algorithm for finding stable topics in the Q&A session and apply the same logic to the presentation session, with details provided in Appendix B. Starting with the 400 topics from ten LDA estimates (40 Q&A topics per random seed), we merge topics whose term distributions are highly correlated into a single topic. Moreover, we retain a topic only if it has been merged at least five times, and we refer to such topics as stable topics. For instance, if the topic "Capital Structure" is essential in the Q&A session, it should emerge as a dominant common theme even if different initial states



(a) Q&A topics

(b) Presentation topics

Figure 1: Log Perplexity Scores of LDA Model for Different Numbers of Topics

The figure plots the log-perplexity scores of the LDA model for different numbers of topics in the Q&A and presentation sessions of conference calls, shown in Panels (a) and (b), respectively.

are used in the estimation, and its estimated probability distributions over terms should be highly correlated across random seeds.[4]

Using the algorithm in Appendix B, we obtain 32 stable topics in both the Q&A and presentation sessions. Table A2 presents the topic names, the average loadings of firms' conference call texts on these topics, and the key terms in each topic based on the full-sample LDA estimation. Since the LDA estimation is entirely unsupervised, we manually assign a label to each topic based on our understanding of the key terms. Several observations are worth noting.

First, the estimated LDA topics are clearly interpretable and semantically meaningful. For example, one important topic in the Q&A session is labeled "Cloud Service," with the

[4]Prior literature has also acknowledged that LDA estimation is sensitive to random initial states (see the Online Technical Appendix of Hansen et al. (2018)). Specifically, Hansen et al. (2018) compute perplexity scores for five random states and select the estimate with the lowest perplexity score. However, as pointed out in their paper and confirmed by our empirical analysis, perplexity scores are rather similar across random seeds; hence, selecting one particular random seed entails high model uncertainty and is likely to omit important topics. Unlike their paper, we consider topic estimates across ten random seeds and propose an algorithm to identify stable topics (see Appendix B).

following key terms:

> *"cloud, subscription, billings, recurring revenue, gotomarket, automation, ecosystem, subscription revenue, intelligence, workflow, hardware, hybrid, professional services."*

This list of words is highly coherent, and all of the terms are closely related to the "Cloud Service" topic.

Second, the firms' loadings on the LDA topics are able to capture interpretable cross-sectional information. Figure 2 provides two such examples. In Panel (a), we show a short excerpt from the Q&A session of Broadwing Communications in Q4 2002. From this conversation between the analyst and the CFO, we infer that "Capital Structure" is the dominant topic, and the LDA model indeed captures high loadings on this theme. Panel (b) of Figure 2 presents another example from the presentation session, in which the COO of Dun & Bradstreet Corporation focuses heavily on issues such as leadership and team members; hence, the LDA model assigns the largest weight to the "Organizational Management" topic. Overall, the LDA model successfully translates the common themes that managers and analysts emphasize into simple quantitative measures—namely, the topic loadings. These topic loadings will later be used as a new set of cross-sectional information in conditional asset pricing models,[5] in addition to existing firm characteristic data.

Third, the Q&A and presentation sessions contain both session-specific and common information. Figure 3 displays the common (highlighted in red) and session-specific (highlighted in green) topics in the Q&A and presentation sessions, with font size corresponding to the average loadings of firms' conference call texts on these topics.[6] While "Forecasting and Estimation," "Economic Uncertainty," and "Capital Structure" are important topics in the Q&A sessions, the presentation sessions are dominated by topics such as "Organizational Management," "Profits," and "Business Operations." We label these topics as Q&A-specific

[5]Similar to the preprocessing applied to firm characteristic data, we apply a rank transformation to the topic loadings so that they lie in the interval $[-0.5, 0.5]$, which ensures that the transformed topic loadings capture purely cross-sectional information. Moreover, under this standardization, the univariate text-managed portfolios constructed conditional on firms' loadings on each LDA topic are zero-cost long-short portfolios.

[6]We define common topics based on the correlation between the term probability distributions of the corresponding Q&A and presentation topics. In particular, we require a pair of topics to be classified as common if their correlation coefficient is at least 0.7. We further verify that such highly correlated topic pairs share common information by inspecting their key terms.

Broadwing Communications, 2002Q4 Earnings call	
Bardone, Analyst (Question)	Hey, guys, congratulations on securing the restructuring. Just wanted to ask three quick questions. First, I think Tom, you were talking about some deleveraging targets through 2005, which looked pretty substantial. Could you be more specific about how the deleveraging is going to take place? Are you talking about operational cash flows being used to buy back debt early in the public markets, or are you talking about potentially an equity deal here and whether you've made commitments to the banks on that front? The second question I have was on the Oak Hill paper...
Schilling, CFO (Answer)	Hey, David. Tom. I'll take the first couple questions then probably turn it over to Jack for the last one. On the senior debt deleveraging I was talking about I was specifically talking about our senior debt which is our senior secured debt which is our bank credit facility and our capital leases, as well as the 7% notes. And that -- the deleveraging there is coming from the junior capital we're raising arranged by Goldman Sachs of $350m, along with the operational cash flows during the next three years that will be generated from the business. On the convert terms of Oak Hill...
Topic Loadings: Capital Structure, 0.403 (Highest)	

(a) Q&A example

Dun & Bradstreet Corp, 2004Q3 Earnings Call	
Alesio, COO	... Our team members' focused leadership drove our risk management results in this quarter ... Let me show an example of how our team members' leadership benefited our e-business results in the quarter. In this other example, we have a customer which is a large company in the high tech industry that has used D&B Solutions for many years. In some cases they needed access to deeper and timelier information than their D&B current solution offered. And they wanted to offer access to this business insight more broadly through their organization. So, members of our sales and marketing solutions team partnered with our Hoover's team to meet our customers' needs. Through teamwork and results leadership they showed how our customer could achieve deeper, more accurate, and more timely business insight by adding Hoover's information to their current D&B Solution. As a result, our customer purchased an enterprise wide license and a full subscription to Hoover's data. This is another great example of how our team members are leading with the interest of our customers first in order to drive superb results...
Topic Loadings: Organizational Management, 0.270 (Highest)	

(b) Presentation example

Figure 2: Conference Call Examples in Q&A and Presentation Sessions

The figure displays two examples of conference calls: a Q&A session in Panel (a) and a presentation session in Panel (b). For each example, we show the transcript and its loadings on the most important topics, with key terms highlighted in red.

(a) Q&A topics **(b)** Presentation topics

Figure 3: Common and Specific Topics in Q&A and Presentation Sessions

The figure plots the common (red) and session-specific (green) topics in the Q&A and presentation sessions. The font size of each topic reflects the average loading of firms' conference call texts on that topic.

or presentation-specific. Some topics are important in both sessions, such as "Inflation," which we refer to as common topics. Overall, there are 13 session-specific (Q&A or presentation) topics and 19 common topics. Furthermore, many common topics (e.g., "Health Care," "Pharmaceutics," "Steel," "Telecommunications") represent industry-specific exposures. Unlike topics such as "Economic Uncertainty" or "Macro Environment," which potentially impact all firms, most industry-specific topic loadings primarily affect only a relatively small subset of firms. Therefore, using the topic loadings as conditional instruments in the IPCA model implies that a relatively large number of latent factors may be required to span the mean–variance efficient frontier.

Fourth, the topics we recover from conference call transcripts—especially those in the Q&A session—shed light on how these calls differ from other sources of firm-specific textual information, such as 10-K filings. Unlike the one-way disclosure in the 10-K, conference calls are interactive, allowing analysts to probe, challenge, and redirect the discussion, which can reveal managerial beliefs and information not fully captured in written filings. This interactivity also allows conference calls to capture more soft, nuanced information—typically qualitative, subjective, and forward-looking. For example, "Growth Outlook" appears uniquely in the Q&A session, where managers and analysts discuss in depth their forward-looking expectations about the firm's future performance. This topic contrasts with related topics such as “Financial Reporting” in the presentation session or the audited financial statements with accompanying notes (Item 8) in the 10-K, which are more standardized and less interactive. Moreover, because conference call transcripts are more conversational, they exhibit richer variation in sentiment. As shown in Figure 3, we identify an “Optimism” topic in the Q&A sessions of conference calls, which supports the view that these calls provide incremental, sentiment-rich information relative to 10-K filings.

Given that conference calls convey sentiment-rich information, we next examine whether the returns on our text-managed portfolios are related to aggregate investor sentiment in the spirit of Baker and Wurgler (2006). In their framework, sentiment has stronger effects on stocks that are harder to arbitrage or more difficult to value. The underlying explanation is that arbitrageurs find these stocks more costly and risky to trade due to the higher degree of noise-trader risk. Consequently, following high sentiment states, the future returns on these stocks tend to experience larger subsequent declines than the returns on stocks that are relatively easier to value and arbitrage.

If the above logic holds, firms with larger loadings on topics like “Optimism” should exhibit stronger (and more adverse) return predictability with respect to aggregate investor sentiment than other firms. To test this hypothesis, we first construct the text-managed portfolio returns, $z_{j,t+1} = \boldsymbol{X}_{jt}^{\top}\boldsymbol{r}_{t+1}$, where $\boldsymbol{X}_{jt}$ is a vector of rank-transformed loadings on topic j. We further decompose $z_{j,t+1}$ into long and short legs, i.e., $z_{j,t+1} = z_{j,t+1}^{\text{long}} - z_{j,t+1}^{\text{short}}$, where $z_{j,t+1}^{\text{long}}$ and $z_{j,t+1}^{\text{short}}$ comprise the firms with positive and negative loadings on topic j, respectively. Next, following prior literature (e.g., Stambaugh et al. (2012)), we run the predictive regression

$$r_{t+1}^{\text{text}} = a + b \times \text{sentiment}_t + \text{error}_{t+1}, \tag{5}$$

where r_{t+1}^{text} is either the long-short text-managed portfolio, or its own long and short legs separately, and sentiment_t is the one-month lagged Baker and Wurgler (2006) sentiment measure that is orthogonal to several macro variables. The regression results are reported in Table 2, where we display the topics whose text-managed (long-short) portfolios have sentiment coefficients that are statistically significant at the 5% level.

First, the long-short portfolio associated with “Optimism” loads negatively on lagged

Table 2: Investor Sentiment and Text-Managed Portfolios

	Long-short			Long leg			Short leg		
Topic	b	t-stat.	R^2_{adj}	b	t-stat.	R^2_{adj}	b	t-stat.	R^2_{adj}
Optimism (QA-specific)	-0.09	-1.97	4.32	-0.31	-4.19	3.93	-0.22	-2.92	1.76
Electronics (QA-specific)	-0.09	-3.41	3.63	-0.30	-3.96	3.63	-0.21	-3.28	2.09
Litigation and Negotiation (QA-specific)	-0.10	-2.78	3.16	-0.31	-4.12	3.60	-0.21	-2.94	1.65
Organizational Management (PRE-specific)	-0.10	-3.50	7.39	-0.29	-3.75	3.27	-0.20	-3.15	1.59
Legal and Regulatory Matters (PRE-specific)	-0.16	-3.12	7.28	-0.34	-4.34	4.22	-0.17	-2.35	1.12
Financial Reporting (PRE-specific)	-0.16	-3.10	5.57	-0.34	-4.39	4.42	-0.18	-2.34	1.05
Entertainment (Common)	-0.11	-3.54	11.95	-0.29	-4.09	4.12	-0.18	-2.94	1.59
Telecommunication (Common)	-0.09	-4.84	11.18	-0.24	-4.44	4.93	-0.15	-3.27	1.97
Cloud Service (Common)	-0.14	-2.75	6.69	-0.33	-4.40	4.57	-0.19	-2.60	1.24
Education (Common)	-0.06	-1.99	4.32	-0.28	-3.88	3.06	-0.22	-3.12	1.95
Semiconductor and Electronics (Common)	-0.08	-2.72	2.29	-0.30	-3.75	3.11	-0.22	-3.27	2.11
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Economic Uncertainty (QA-specific)	0.07	3.16	4.11	-0.22	-3.04	1.77	-0.29	-3.96	3.37
Forecasting and Estimation (QA-specific)	0.10	2.08	3.71	-0.21	-2.73	1.62	-0.31	-4.08	3.78
Growth Outlook (QA-specific)	0.11	3.04	3.41	-0.20	-2.97	1.56	-0.31	-3.95	3.36
Transportation: Service (QA-specific)	0.07	2.93	2.28	-0.21	-2.91	1.43	-0.28	-3.82	3.24
Business Operation (PRE-specific)	0.15	3.59	5.76	-0.18	-2.89	1.39	-0.33	-4.04	3.77
Energy (PRE-specific)	0.11	4.44	4.98	-0.13	-1.95	0.45	-0.24	-4.11	3.73
Metals and Mining (PRE-specific)	0.09	3.53	4.73	-0.17	-2.51	1.09	-0.26	-3.94	3.38
Asset Management (PRE-specific)	0.07	2.92	3.75	-0.18	-2.75	1.67	-0.26	-3.72	3.02
International Finance (PRE-specific)	0.10	3.13	3.62	-0.20	-3.11	1.80	-0.30	-3.95	3.39
Freight (PRE-specific)	0.05	2.62	2.69	-0.20	-3.00	1.70	-0.25	-3.68	3.14
Renewable Energy (Common)	0.13	4.80	8.92	-0.16	-2.47	0.97	-0.29	-4.09	3.64
Inflation (Common)	0.11	2.46	3.22	-0.20	-2.61	1.23	-0.31	-4.31	3.93
Oil (Common)	0.08	3.40	2.96	-0.06	-1.38	-0.04	-0.14	-3.90	2.81
Insurance (Common)	0.04	2.11	1.38	-0.15	-3.00	2.03	-0.19	-3.50	2.82

The table reports the regression results in equation (5). The Q&A and presentation topics come from the full-sample LDA estimation. For common topics that appear in both sessions (e.g., "Cloud Service"), we compute the equal-weighted average of the returns on their text-managed portfolios. We use the Baker and Wurgler (2006) sentiment measure, which is constructed to be orthogonal to several macro variables. In the panel labeled "Long-short", we regress the text-managed portfolio return on the lagged sentiment measure. In the panels labeled "Long leg" and "Short leg", we regress the long and short legs of the text-managed portfolio returns on the lagged sentiment measure, respectively. We report the point estimates of the coefficient of sentiment, b, its t-statistic based on the Newey and West (1987) standard errors with 12 lags, and the adjusted R^2 (%). We display the topics whose text-managed portfolios have coefficients on sentiment that are statistically significant at the 5% level. The sample period ranges from January 2003 to December 2022.

aggregate sentiment. In high-sentiment periods, firms that speak more optimistically in Q&A are more exposed to sentiment-driven mispricing and become temporarily more overvalued than other firms, leading to stronger negative price reversals in the subsequent period.

Similar negative predictive relationships arise for the "Litigation and Negotiation" and "Legal and Regulatory Matters" topics. Because litigation-intensive stocks are riskier and harder to arbitrage, high-sentiment periods—when noise-trader risk is elevated—discourage arbitrageurs from aggressively shorting these companies, in line with a limits-to-arbitrage mechanism. Consequently, the returns on these "legal risk" stocks tend to be lower than those of other firms after sentiment spikes.

By contrast, many long-short portfolios tied to macro fundamentals or forward-looking operational guidance exhibit positive predictive coefficients on sentiment. Topics such as "Business Operation," "Economic Uncertainty," "Forecasting and Estimation," and "Growth Outlook" all show this pattern. It implies that firms providing more extensive fundamental discussion and clearer forward guidance experience smaller sentiment-driven reversals.

Moreover, the industry topics exhibit heterogeneous cross-sectional patterns in how sentiment drives returns. Text-managed portfolios tied to fast-evolving or intangible-intensive sectors—such as Telecommunication, Cloud Service, and Semiconductor and Electronics—show significantly negative coefficients on sentiment. In contrast, topics linked to tangible-asset-intensive sectors such as "Energy," "Metals and Mining," and "Oil" show opposite-signed exposure to sentiment. Overall, the soft and nuanced information in conference calls helps identify firms that are harder to value and more vulnerable to investor sentiment.

4 Asset Pricing Performance

In this section, we investigate whether the firms' loadings on Q&A or presentation topics, as uncovered in Section 3, contain useful information for conditional asset pricing models. In particular, our paper examines the out-of-sample (OOS) performance of the IPCA model conditional on various sources of cross-sectional information. For each month τ, we conduct an expanding-window exercise: using all available data up to month τ, we estimate the IPCA model and denote the resulting parameter estimates by $\hat{\mathbf{\Gamma}}_\tau$. This, in turn, implies that the dynamic betas in month τ are given by $\hat{\boldsymbol{\beta}}_\tau = \boldsymbol{X}_\tau \hat{\boldsymbol{\Gamma}}_\tau$. Next, we construct the OOS latent factors as $\hat{\boldsymbol{f}}_{\tau+1} = \left(\hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_\tau^\top \boldsymbol{X}_\tau \hat{\boldsymbol{\Gamma}}_\tau\right)^{-1} \hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_\tau^\top \boldsymbol{r}_{\tau+1}$. Since the constructed latent factors are tradable, we can estimate their risk premia based on time-series averages: $\hat{\boldsymbol{\lambda}}_{f\tau} = \frac{1}{\tau}\sum_{t=0}^{\tau-1}\left(\hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_t^\top \boldsymbol{X}_t \hat{\boldsymbol{\Gamma}}_\tau\right)^{-1} \hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_t^\top \boldsymbol{r}_{t+1}$. Note that the estimation of $\hat{\boldsymbol{\Gamma}}_\tau$, $\hat{\boldsymbol{\lambda}}_{f\tau}$, and the individual stock return weights, $\left(\hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_\tau^\top \boldsymbol{X}_\tau \hat{\boldsymbol{\Gamma}}_\tau\right)^{-1} \hat{\boldsymbol{\Gamma}}_\tau^\top \boldsymbol{X}_\tau^\top$, rely only on information up to time τ. Hence, these quantities can be used to evaluate the OOS performance.

Following Kelly et al. (2019), we report both total and predictive time-series R^2 measures for individual stock returns, defined as follows:

$$\text{total } R^2 = 1 - \frac{\sum_t (\boldsymbol{r}_{t+1} - \hat{\boldsymbol{\beta}}_t \hat{\boldsymbol{f}}_{t+1})^\top (\boldsymbol{r}_{t+1} - \hat{\boldsymbol{\beta}}_t \hat{\boldsymbol{f}}_{t+1})}{\sum_t \boldsymbol{r}_{t+1}^\top \boldsymbol{r}_{t+1}},$$

$$\text{predictive } R^2 = 1 - \frac{\sum_t (\boldsymbol{r}_{t+1} - \hat{\boldsymbol{\beta}}_t \hat{\boldsymbol{\lambda}}_{ft})^\top (\boldsymbol{r}_{t+1} - \hat{\boldsymbol{\beta}}_t \hat{\boldsymbol{\lambda}}_{ft})}{\sum_t \boldsymbol{r}_{t+1}^\top \boldsymbol{r}_{t+1}}.$$

We also examine the mean-variance efficiency of the IPCA model conditional on different sets of cross-sectional instruments. To do so, we construct the model-implied mean-variance efficient (MVE) portfolios in the out-of-sample. Using the same expanding window analysis, for each month τ, we estimate the sample mean and covariance matrix of $\hat{\boldsymbol{f}}$, denoted by $\hat{\boldsymbol{\lambda}}_{f\tau}$ and $\hat{\boldsymbol{\Sigma}}_{f\tau}$, respectively. The MVE portfolio is then defined as

$$\hat{\boldsymbol{f}}_{\tau+1}^{MVE} = \hat{\boldsymbol{\lambda}}_{f\tau}^\top \hat{\boldsymbol{\Sigma}}_{f\tau}^{-1} \hat{\boldsymbol{f}}_{\tau+1}. \tag{6}$$

Note that the MVE portfolios in equation (6) are constructed conditional on different cross-sectional information sets (e.g., Q&A and presentation topic loadings, 35 firm characteristics, and their combinations). Therefore, we refer to them as conditional MVE portfolios.

4.1 Full-Sample vs. Online LDA

In Section 3, we estimate the LDA model using the entire corpus—namely, all earnings call transcripts in the full sample. Yet, the full-sample LDA model is likely subject to the critique that the estimates of the probability distributions over terms, $\boldsymbol{\Theta}$, rely on full-sample information. Hence, the estimates of firm i's topic loadings in month t, $\boldsymbol{\phi}_{it}$, are also susceptible to the lookahead bias. Note that $\boldsymbol{\phi}_{it}$ is obtained directly from the time-t conference call transcript of firm i, so the only source of forward-looking bias originates from

$\mathbf{\Theta}$. In fact, lookahead bias is a general issue in pretrained language models (e.g., BERT and ChatGPT), as pointed out by Sarkar and Vafa (2024).

We extend our analysis to construct a purely real-time conditional asset pricing model based on online LDA, which was initially proposed by Hoffman et al. (2010) to overcome the computational issue in the LDA model. In particular, in each month τ, we estimate the LDA model using only the conference call transcripts from month 1 through τ and obtain the topic loading estimates, $\hat{\boldsymbol{\phi}}_{i\tau}$, which will later be used as instruments in the IPCA model. The primary issue underlying online LDA analysis is that the estimates of the probability distributions over terms, $\hat{\mathbf{\Theta}}_\tau$, evolve over time. Therefore, it is difficult to visualize the topics as in Figure 3 and Table A2 or study the topic importance as in Section 6. For interpretation, the analysis based on the full-sample LDA is more transparent and interpretable than that based on the online LDA.

As with the full-sample LDA, the online LDA model also heavily relies on its initial states. To handle this issue, we repeat the online LDA estimation (32 topics for Q&A and presentation sessions, respectively) using 10 different random seeds. For each seed, we construct both the OOS predicted stock returns ($\hat{\boldsymbol{\beta}}_\tau \hat{\boldsymbol{f}}_{\tau+1}$ and $\hat{\boldsymbol{\beta}}_\tau \hat{\boldsymbol{\lambda}}_{f\tau}$) and the MVE portfolio ($\hat{\boldsymbol{\lambda}}_{f\tau}^\top \hat{\mathbf{\Sigma}}_{f\tau}^{-1} \hat{\boldsymbol{f}}_{\tau+1}$), following the same expanding-window procedure as before. Finally, we take the simple averages of $\hat{\boldsymbol{\beta}}_\tau \hat{\boldsymbol{f}}_{\tau+1}$, $\hat{\boldsymbol{\beta}}_\tau \hat{\boldsymbol{\lambda}}_{f\tau}$, and $\hat{\boldsymbol{\lambda}}_{f\tau}^\top \hat{\mathbf{\Sigma}}_{f\tau}^{-1} \hat{\boldsymbol{f}}_{\tau+1}$ across 10 random seeds. The final step can be interpreted as integrating out the uncertainty induced by the initial random states in the LDA estimation.

Throughout this section, we report the OOS performance of the conditional asset pricing models using topic loadings estimated from both the full-sample and online LDA models.

4.2 Time-Series Fit

We first explore the time-series fit of the IPCA model using topic loadings and firm characteristics as instruments. Table 3 reports the OOS total and predictive R^2 of the IPCA model using as instruments: i) 32 Q&A topic loadings, ii) 32 presentation topic loadings, iii) 64 Q&A plus presentation topic loadings, iv) 35 firm characteristics, and v) both the 64 topic loadings and the 35 firm characteristics. The full sample runs from January 2003 to December 2022. To ensure that the initial in-sample window contains sufficient time-series

Table 3: Out-of-Sample Time-Series Fit

Cross-Sectional Information Sets:		$p = 5$	$p = 10$	$p = 15$	$p = 20$	$p = 25$	$p = 30$
		Panel A. Full-Sample LDA					
32 Q&A topics	total R^2	21.28	22.15	22.50	22.78	23.03	23.27
	pred R^2	0.14	0.14	0.13	0.13	0.14	0.14
32 PRE topics	total R^2	21.69	22.58	22.98	23.26	23.53	23.75
	pred R^2	0.14	0.13	0.14	0.14	0.14	0.14
64 Q&A + PRE topics	total R^2	22.05	22.99	23.41	23.72	23.99	24.24
	pred R^2	0.13	0.12	0.12	0.13	0.13	0.13
35 Firm Characteristics	total R^2	24.00	25.13	25.79	26.30	26.71	27.04
	pred R^2	0.14	0.13	0.14	0.15	0.15	0.14
64 Q&A + PRE topics + 35 Firm Chars	total R^2	24.46	25.86	26.58	27.03	27.42	27.76
	pred R^2	0.09	0.12	0.15	0.15	0.15	0.16
		Panel B. Online LDA					
32 Q&A topics	total R^2	21.31	22.23	22.73	23.10	23.44	23.76
	pred R^2	0.14	0.14	0.14	0.14	0.15	0.15
32 PRE topics	total R^2	21.82	22.80	23.30	23.68	24.03	24.35
	pred R^2	0.15	0.15	0.15	0.15	0.15	0.15
64 Q&A + PRE topics	total R^2	22.18	23.27	23.84	24.25	24.60	24.95
	pred R^2	0.14	0.13	0.14	0.14	0.15	0.15
64 Q&A + PRE topics + 35 Firm Chars	total R^2	24.48	25.98	26.80	27.39	27.89	28.33
	pred R^2	0.09	0.12	0.15	0.15	0.15	0.15

The table presents the out-of-sample time-series fit of the IPCA model, using as instruments: (i) 32 Q&A topic loadings, (ii) 32 presentation (PRE) topic loadings, (iii) 64 Q&A and presentation topic loadings, (iv) 35 firm characteristics, and (v) both the 64 topic loadings and the 35 firm characteristics. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. The number of latent factors in the IPCA estimation, denoted by p, is set to 5, 10, 15, 20, 25, and 30. The total and predictive R^2 are calculated across the cross-section of individual stock returns over the out-of-sample period from January 2008 to December 2022.

information to estimate the model parameters, the out-of-sample period begins in January 2008. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. We consider different numbers of latent factors, $p \in \{5, 10, 15, 20, 25, 30\}$.

Both Q&A and presentation topics contain substantial information for explaining the time-series variation in asset returns. For example, when the number of latent factors is $p = 25$, the total and predictive R^2 of the IPCA model that uses both Q&A and presentation topic loadings from the full-sample LDA are 24% and 0.13%, respectively, which are close to those of the model based on traditional firm characteristics (27% total and 0.15% predictive R^2). The same patterns hold when topics are estimated using online LDA. However, augmenting firm characteristics with textual information from conference calls only modestly increases the total R^2 and leaves the predictive R^2 essentially unchanged, suggesting limited additional gains in time-series fit from combining the two information sets.

4.3 Mean-Variance Efficiency

Table 4 reports the annualized Sharpe ratios of the MVE portfolios implied by the different IPCA specifications over 2008–2022. When we use Q&A or presentation topic loadings from full-sample LDA as the cross-sectional conditioning information (first two rows of Panel A), more than 20 latent factors are needed to approach the maximal OOS Sharpe ratio. For example, when the number of latent factors is 25, the OOS Sharpe ratios of the MVE portfolios conditional on Q&A or presentation topic loadings are 1.15 or 1.07, respectively, versus only 0.34 and 0.79 in the five-factor models. Using both sets of topic loadings further raises the OOS Sharpe ratio to 1.38 when $p = 25$, and up to 1.64 when $p = 30$ (see row 3 in Panel A). The lack of sparsity in the MVE portfolios is plausibly due to the fact that many topics capture industry-specific exposures, as discussed in Section 3, making a sparse latent factor structure unrealistic when conditioning on many industry-like signals. In the subsequent analysis, we therefore focus on specifications with $p = 25$ or $p = 30$.

The IPCA model that uses the 35 firm characteristics as instruments also performs well, delivering out-of-sample Sharpe ratios around 1.5–1.7 once $p \geq 20$ (fourth row of Panel A). More importantly, augmenting firm characteristics with conference call topics substantially improves mean-variance efficiency. When $p = 25$, the MVE portfolio conditional on both the 64 conference call topics and the 35 firm characteristics attains an out-of-sample Sharpe ratio of 2.53 (fifth row of Panel A), and 2.65 when $p = 30$, a large improvement relative to using either information set alone.

Panel B shows that the same qualitative patterns hold when topics are estimated using online LDA. The IPCA model remains non-sparse, and Sharpe ratios increase with the number of latent factors. When $p = 25$, the MVE portfolio based solely on online-LDA topic loadings achieves an out-of-sample Sharpe ratio of 1.46, very close to the 1.38 obtained with full-sample LDA topics. We also confirm that the conference call transcripts contain significant incremental, *real-time*, cross-sectional information beyond the traditional firm

Table 4: Out-of-Sample Sharpe Ratios of MVE Portfolios

Cross-Sectional Information Sets:	$p = 5$	$p = 10$	$p = 15$	$p = 20$	$p = 25$	$p = 30$
	Panel A. Full-Sample LDA					
32 Q&A topics	0.34	0.67	0.47	0.74	1.15	1.04
32 PRE topics	0.79	0.92	1.03	1.17	1.07	1.16
64 Q&A + PRE topics	0.55	0.65	0.95	1.29	1.38	1.64
35 Firm Characteristics	0.63	1.16	1.57	1.68	1.74	1.56
64 Q&A + PRE topics + 35 Firm Chars	0.23	1.49	2.30	2.33	2.53	2.65
	Panel B. Online LDA					
32 Q&A topics	0.20	0.02	0.73	0.91	1.03	1.00
32 PRE topics	0.52	0.69	0.95	1.05	1.08	1.16
64 Q&A + PRE topics	0.43	0.42	1.08	1.38	1.46	1.47
64 Q&A + PRE topics + 35 Firm Chars	0.21	1.27	1.99	2.14	2.19	2.30

The table reports the annualized out-of-sample Sharpe ratios of the MVE portfolios implied by the IPCA models using as instruments: (i) 32 Q&A topic loadings, (ii) 32 presentation (PRE) topic loadings, (iii) 64 Q&A and presentation topic loadings, (iv) 35 firm characteristics, and (v) both the 64 topic loadings and the 35 firm characteristics. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. The number of latent factors in the IPCA estimation, denoted by p, is set to 5, 10, 15, 20, 25, and 30. The out-of-sample period is from January 2008 to December 2022.

characteristics.

One may be concerned that the 2008 global financial crisis disproportionately drives our results. To address this concern, we repeat the analysis in Table 4 using an alternative OOS period—January 2013 to December 2022—and report the results in Table A3. The results are qualitatively unchanged: the textual information contained in conference call transcripts continues to provide substantial incremental cross-sectional conditioning information beyond traditional firm characteristics.

As a final exercise in this subsection, we test whether standard factor models can explain our text-managed MVE portfolios. We consider five benchmark factor models: i) Fama and French (1993) three-factor model (FF3), ii) Fama and French (2015) five-factor model (FF5), iii) five Fama-French factors plus momentum (FFC6), iv) Hou et al. (2015) four-factor model (Q4), and v) the characteristics-managed MVE portfolio (p equals 25 or 30 in the IPCA model). We regress the monthly returns of the text-managed MVE portfolios—constructed using 64 Q&A and presentation topic loadings—on each benchmark and report the intercepts (alphas) and Newey–West (Newey and West, 1987) t-statistics with 12 lags.

Table 5 shows that all alphas are positive, statistically significant, and larger than 1% per month.[7] Therefore, none of these models can explain the OOS performance of the text-managed MVE portfolios.

Table 5: Can Canonical Factor Models Explain Text-Managed MVE Portfolios?

	Panel A. Full-Sample LDA		**Panel B.** Online LDA	
	$p = 25$	$p = 30$	$p = 25$	$p = 30$
FF3	1.7%	2.1%	1.8%	1.8%
t-stat	(4.762)	(5.618)	(3.482)	(3.532)
FF5	1.6%	2.0%	1.5%	1.6%
t-stat	(4.584)	(5.885)	(3.070)	(3.165)
FFC6	1.6%	2.0%	1.5%	1.5%
t-stat	(5.031)	(6.064)	(3.498)	(3.612)
Q4	1.6%	2.0%	1.5%	1.5%
t-stat	(5.457)	(5.954)	(3.090)	(3.090)
CHAR-managed	1.8%	2.2%	1.4%	1.5%
t-stat	(3.788)	(5.467)	(2.528)	(2.699)

The table reports the estimated alphas of text-managed MVE portfolios constructed conditional on 64 Q&A and presentation topic loadings. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. The number of latent factors used to construct the MVE portfolios is either 25 or 30. We run the univariate time-series regression by regressing the text-managed MVE portfolio returns on several factor models: i) Fama and French (1993) three-factor model (FF3), ii) Fama and French (2015) five-factor model (FF5), iii) five Fama-French factors plus momentum (FFC6), iv) Hou et al. (2015) four-factor model (Q4), and v) the MVE portfolio conditional on 35 firm characteristics (based on the IPCA model, denoted as CHAR-based). We report the estimates of the intercepts (alphas) and the t-statistics based on Newey-West standard errors with 12 lags. The sample ranges from January 2008 to December 2022.

4.4 First- vs. Second-Moment Channels of Textual Information

But how does the textual information enhance the mean-variance efficiency? By construction, the conditional MVE portfolio weights over the universe of individual stocks are proportional to $\boldsymbol{\Sigma}_t^{-1}\mathbb{E}_t[\boldsymbol{r}_{t+1}]$, so additional information contained in topic loadings can help either by refining expected returns $\mathbb{E}_t[\boldsymbol{r}_{t+1}]$ (first-moment channel) or by improving estimates of the conditional covariance matrix $\boldsymbol{\Sigma}_t$ (second-moment channel). Table 6 quantifies these two channels by comparing portfolios sorted on model-implied expected returns, $\boldsymbol{\beta\lambda}$, with

[7] We normalize the text-managed MVE portfolio returns to match the volatility of the market portfolio (4.9% per month) over 2008–2022.

portfolios sorted on model-implied MVE portfolio weights, $\boldsymbol{\omega}_t^{\text{mve}}$, under three conditional information sets.

The first set of columns in each panel uses the model-implied expected excess return, $\boldsymbol{\beta\lambda}$, to sort stocks into quintiles. This directly evaluates the first-moment channel. With only 35 firm characteristics (Panel A), the IPCA model's predictive R^2 is just 0.15% (see Table 3), suggesting that expected returns are very noisy. Nevertheless, the economic magnitude of cross-sectional predictability is nontrivial: the long–short portfolio that buys the highest-predicted-return quintile and sells the lowest (row "5–1") earns an annualized mean excess return of 7.3% with a Sharpe ratio of 0.55. This divergence between small R^2 and sizable economic spreads is consistent with prior literature (e.g., Lewellen (2015)) emphasizing economic magnitudes over purely statistical fit.

When the IPCA model is estimated with topic loadings alone (Panels B and D), the portfolios sorted on mean-implied returns are noticeably weaker. For example, in Panel B (full-sample LDA topics), the 5–1 portfolio based on $\boldsymbol{\beta\lambda}$ delivers an annualized mean of only 2.9% and a Sharpe ratio of 0.32, much smaller than those in the characteristics-only IPCA model. Textual information by itself is therefore a relatively poor predictor of expected returns. However, when firm characteristics and topic loadings are combined (Panels C and E), the cross-sectional return spreads improve. In Panel C, the 5–1 portfolio sorted on $\boldsymbol{\beta\lambda}$ earns an annualized mean of 10.4% (with a t-statistic of 3.26) and a Sharpe ratio of 0.88. Thus, textual information does add incremental predictive content for expected returns when combined with standard characteristics, but the overall improvement in the mean-return channel remains modest.

The second set of columns in each panel uses the model-implied MVE weights, $\boldsymbol{\omega}_t^{\text{mve}}$, to form quintile portfolios and thereby incorporates the covariance channel. Two striking patterns emerge. First, for all information sets, the quintiles sorted on $\boldsymbol{\omega}_t^{\text{mve}}$ exhibit monotonically increasing mean excess returns and declining volatilities from portfolio 1 to 5. As a result, the Sharpe ratios increase monotonically across the quintiles. This contrasts sharply with the mean-return sorts, where volatilities rise with expected return rank. The intuition is that $\boldsymbol{\omega}_t^{\text{mve}}$ explicitly incorporates the conditional covariance matrix $\boldsymbol{\Sigma}_t$ and therefore constructs portfolios that better exploit diversification to maximize the portfolio Sharpe ratio.

Table 6: Sources of Incremental Information: Quintile Portfolios Sorted on IPCA-Implied Expected Returns vs. MVE Portfolio Weights

	Sorted by $\boldsymbol{\beta}_t\boldsymbol{\lambda}_t$			Sorted by $\boldsymbol{\omega}_t^{\text{mve}}$		
Portfolio rank	mean (%)	std. (%)	SR	mean (%)	std. (%)	SR
	Panel A. 35 Firm Char.					
1	7.8 (1.55)	20.2	0.38	7.9 (1.13)	25.0	0.32
2	10.3 (1.92)	20.2	0.51	10.1 (1.60)	23.7	0.43
3	10.8 (1.87)	22.0	0.49	11.5 (1.88)	22.6	0.51
4	12.6 (1.95)	24.0	0.53	12.2 (1.98)	21.9	0.56
5	15.1 (1.63)	29.3	0.52	14.9 (2.39)	21.5	0.69
5–1	7.3 (1.55)	13.4	0.55	7.0 (5.43)	6.3	1.11
	Panel B. 64 Topic Loadings in Full-Sample LDA					
1	9.5 (1.46)	23.2	0.41	7.6 (1.10)	24.2	0.32
2	11.0 (1.74)	23.0	0.48	10.4 (1.59)	23.3	0.45
3	11.8 (1.86)	22.7	0.52	12.5 (1.94)	23.1	0.54
4	12.0 (1.84)	22.9	0.52	12.6 (2.13)	22.3	0.57
5	12.4 (1.95)	23.7	0.52	13.3 (2.28)	21.7	0.62
5–1	2.9 (1.33)	9.1	0.32	5.7 (3.72)	5.0	1.14
	Panel C. 35 Firm Char. + 64 Topic Loadings in Full-Sample LDA					
1	6.6 (1.17)	20.1	0.33	6.3 (0.89)	24.9	0.25
2	9.4 (1.77)	20.9	0.45	9.4 (1.44)	23.5	0.40
3	11.4 (1.92)	22.1	0.51	11.4 (1.90)	22.4	0.51
4	12.3 (1.86)	24.0	0.51	12.9 (2.15)	22.5	0.57
5	16.9 (2.03)	28.2	0.60	16.6 (2.70)	21.3	0.78
5–1	10.4 (3.26)	11.8	0.88	10.3 (7.43)	5.9	1.74
	Panel D. 64 Topic Loadings in Online LDA					
1	10.5 (1.51)	23.7	0.44	8.0 (1.07)	25.0	0.32
2	11.3 (1.80)	22.8	0.49	10.6 (1.54)	23.8	0.45
3	11.4 (1.75)	23.0	0.49	11.8 (1.89)	23.2	0.51
4	12.0 (1.91)	22.8	0.53	12.4 (2.13)	22.3	0.56
5	11.5 (1.87)	23.4	0.49	13.7 (2.54)	20.4	0.67
5–1	1.0 (0.35)	10.2	0.10	5.6 (1.98)	6.9	0.82
	Panel E. 35 Firm Char. + 64 Topic Loadings in Online LDA					
1	6.8 (1.22)	20.1	0.34	5.9 (0.80)	25.5	0.23
2	10.1 (1.88)	20.5	0.49	9.6 (1.46)	23.3	0.41
3	11.1 (1.94)	21.7	0.51	12.0 (2.07)	22.6	0.53
4	12.6 (1.90)	24.0	0.52	13.1 (2.19)	21.9	0.60
5	16.1 (1.87)	29.1	0.55	16.0 (2.62)	21.4	0.75
5–1	9.3 (2.57)	12.8	0.73	10.1 (5.98)	6.6	1.54

The table reports the performance of quintile portfolios sorted by the model-implied mean returns of individual stocks ($\boldsymbol{\beta}_t\boldsymbol{\lambda}_t$) and by the mean–variance efficient portfolio weights ($\boldsymbol{\omega}_t^{\text{mve}}$) implied by the IPCA model. The number of latent factors in the IPCA estimation is set to 30. At the beginning of each calendar month, all individual stocks are ranked in ascending order based on either their model-predicted mean excess returns or their weights in the model-implied MVE portfolio. The ranked stocks are then assigned to one of five quintile portfolios. All stocks are equally weighted within each portfolio, and the portfolios are rebalanced each month to maintain equal weights. We report the annualized mean, standard deviation, and Sharpe ratio of each quintile portfolio. For the portfolio mean returns, we also report t-statistics based on Newey and West (1987) standard errors using 12 lags within the parentheses. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panels B and C and from the online LDA estimation in Panels D and E. The sample period spans January 2008 to December 2022.

Second, and more importantly, the improvement from adding textual information is much larger for the MVE-sorted portfolios than for the mean-return-sorted portfolios. In the characteristics-only model (Panel A), the 5–1 portfolio based on $\boldsymbol{\omega}_t^{\text{mve}}$ achieves a Sharpe ratio of 1.11. In the model using only topic loadings (Panel B), the corresponding 5–1 portfolio has a similar Sharpe ratio of 1.14, even though the mean-return sorts based on topics alone are much weaker. This indicates that topics already help span important dimensions of $\boldsymbol{\Sigma}_t$ even when they carry little stand-alone information about $\mathbb{E}_t[\boldsymbol{r}_{t+1}]$. When characteristics and topics are combined (Panel C), the 5–1 Sharpe ratio for the $\boldsymbol{\omega}_t^{\text{mve}}$-sorted portfolios rises to 1.74, compared to 1.11 with characteristics alone. By contrast, the Sharpe ratio of the mean-return-sorted 5–1 portfolio rises only from 0.55 (Panel A) to 0.88 (Panel C). Hence, the performance gain from adding text is much more pronounced for the MVE-sorted portfolios, clearly pointing to a second-moment channel.

Taken together, these patterns suggest that the primary benefit of incorporating conference call transcripts into the IPCA framework is a better approximation of the conditional covariance matrix, rather than a substantial improvement in expected return forecasts. The textual topics extract information that is largely distinct from standard firm characteristics, and this heterogeneity is confirmed empirically in Table 7, where MVE portfolios built on characteristics versus topics are only weakly correlated. This echoes the insight in Kozak and Nagel (2024): additional characteristics can be valuable even if they add little to expected returns, as long as they help capture the main sources of covariance among stock returns. In our setting, conference call topics play exactly this role, explaining why the largest improvements in mean-variance efficiency arise through the covariance (second-moment) channel.

4.5 Textual vs. Characteristic Information: Return Decomposition

Previous empirical results indicate that both firm characteristics and conference call transcripts are key components of the cross-sectional information set in conditional asset pricing models. But how important is each information source for explaining the cross-section of stock returns? To answer this question, we first use the IPCA models to project returns onto each information set separately, producing total return forecasts $\widehat{R}_{i,t+1}^{char}$ and $\widehat{R}_{i,t+1}^{text}$. Here, $\widehat{R}_{i,t+1}^{char}$ and $\widehat{R}_{i,t+1}^{text}$ denote the IPCA forecasts of $R_{i,t+1}$ conditional on only firm characteristics

Table 7: Correlation Coefficients among MVE Portfolios: 32 Q&A topics, 32 Presentation Topics, and 35 Firm Characteristics

	$p = 25$			$p = 30$		
	32 Q&A topics	32 PRE topics	35 Firm Chars	32 Q&A topics	32 PRE topics	35 Firm Chars
			Panel A. Full-Sample LDA			
32 Q&A topics	1.00	0.22	0.24	1.00	0.17	0.25
32 PRE topics	0.22	1.00	-0.04	0.17	1.00	0.07
35 Firm Chars	0.24	-0.04	1.00	0.25	0.07	1.00
			Panel B. Online LDA			
32 Q&A topics	1.00	0.62	0.31	1.00	0.57	0.31
32 PRE topics	0.62	1.00	0.15	0.57	1.00	0.17
35 Firm Chars	0.31	0.15	1.00	0.31	0.17	1.00

The table reports the correlation coefficients among the MVE portfolios implied by the IPCA model using as instruments the i) 32 Q&A topic loadings, ii) 32 presentation (PRE) topic loadings, and iii) 35 firm characteristics. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. The number of latent factors in the IPCA estimation, denoted by p, equals 25 and 30.

and textual information, respectively. We interpret these forecasts as summary statistics for the information contained in characteristics and text, respectively.

Next, to assess the relative importance of these two information sources in driving return dynamics, we regress realized returns on both forecasts simultaneously:

$$R_{i,t+1} = b_0 + b_1 \widehat{R}^{char}_{i,t+1} + b_2 \widehat{R}^{text}_{i,t+1} + e_{i,t+1}.$$

This regression yields the best linear combination of the characteristic- and text-based signals for explaining realized returns. We then decompose the variation in returns explained by this combined signal into parts attributable to each forecast.

Given the estimates of (b_1, b_2), the share of return variance associated with the fitted component is

$$\underbrace{\frac{\text{Cov}(R_{i,t+1}, b_1 \widehat{R}^{char}_{i,t+1} + b_2 \widehat{R}^{text}_{i,t+1})}{\text{Var}(R_{i,t+1})}}_{R^2_{total}} = \underbrace{b_1 \frac{\text{Cov}(R_{i,t+1}, \widehat{R}^{char}_{i,t+1})}{\text{Var}(R_{i,t+1})}}_{R^2_{char}} + \underbrace{b_2 \frac{\text{Cov}(R_{i,t+1}, \widehat{R}^{text}_{i,t+1})}{\text{Var}(R_{i,t+1})}}_{R^2_{text}}.$$

This variance decomposition quantifies the share of return variation captured by our model that is associated with characteristic information versus textual information, after controlling for their overlap.

Table 8 reports the results of the return decomposition described above. Over the universe of all individual stocks, the IPCA forecasts explain about 28% of the return variation, with textual information accounting for 33–35% of this explained component (depending on whether we use full-sample or online LDA). Thus, conference call text contributes a sizable share of priced return variation.

A natural next step is to examine how the importance of textual information varies across different types of firms. Ex ante, one would expect conference call texts—which capture soft, qualitative information—to be more informative in sectors where fundamentals are harder to summarize using standard balance-sheet and market-based characteristics. This includes R&D-intensive or rapidly evolving industries (such as technology, healthcare, or energy exploration), as well as cyclical, demand-sensitive sectors (such as consumer discretionary and business services). By contrast, in more mature or heavily regulated industries (such as utilities or certain consumer staples), where demand and investment opportunities are relatively stable and well reflected in conventional characteristics, textual disclosures may add little incremental information. These considerations motivate our industry-level heterogeneity analysis of the return decomposition.

Across industries,[8] the decomposition results line up closely with the ex ante hypotheses. Textual information is the most important in sectors where fundamentals are hard to summarize with standard characteristics: Measuring & Control Equipment, Mines, Chips, Oil, Business Services, Telecommunication, Medical Equipment, Computer Hardware, and Healthcare all exhibit relatively high text shares, with R^2_{text}/R^2_{total} around 37–47% in Panel A (full-sample LDA), well above the full-sample benchmark of about one-third. Many cyclical, demand-sensitive industries—such as Restaurants/Hotels, Construction Materials, and Machinery—also feature elevated text contributions. By contrast, more mature or regulated sectors display much lower incremental value of text. Utilities is the most extreme case, where textual information explains only about 4% of the explained variation in Panel A and essentially none in Panel B. Staple-like industries such as Consumer Goods and Retail also have text shares well below the full-sample average.

[8] We start from the Fama-French 49-industry classification but merge several small industries to ensure sufficient firm counts for our analysis. Specifically, we group Autos, Aerospace, Ships, and Guns into a single "Autos" industry and combine Gold, Coal, and Mines into "Mines." Paper and Boxes are merged into a "Paper/Boxes" industry. Financial industries—Banks, Insurance, Real Estate, and Other Finance—are combined into a single "Finance" group. Several small consumer and manufacturing industries (Soda, Beer, Agric, Textiles, Fabricated Products, and Tobacco/Smoke) are pooled into an "Other" category. Finally, Chemicals and Rubber are combined into a single "Chemicals/Rubber" industry.

Table 8: Decomposing Return Variation into Characteristic and Textual Components

Industry	N_{firms}	**Panel A.** Full-Sample LDA				**Panel B.** Online LDA			
		R^2_{total} (%)	R^2_{char} (%)	R^2_{text} (%)	$\frac{R^2_{text}}{R^2_{total}}$ (%)	R^2_{total} (%)	R^2_{char} (%)	R^2_{text} (%)	$\frac{R^2_{text}}{R^2_{total}}$ (%)
All Firms	2338	27.80	18.67	9.14	32.87	27.74	17.92	9.82	35.41
Measuring and Control Equipment	45	31.00	16.32	14.68	47.36	31.67	12.61	19.06	60.18
Mines	19	29.96	15.76	14.20	47.40	29.39	16.29	13.11	44.59
Chips	132	28.80	16.59	12.21	42.40	28.81	15.19	13.62	47.28
Computer Hardware	43	19.27	11.63	7.64	39.63	19.40	10.22	9.19	47.35
Medical Equipment	80	20.06	12.18	7.88	39.29	20.21	10.63	9.58	47.41
Restaurants, Hotels, Motels	46	37.81	22.28	15.53	41.08	37.76	20.78	16.98	44.96
Business Services	123	28.39	17.01	11.38	40.07	28.34	15.80	12.54	44.25
Machinery	76	39.37	24.27	15.10	38.36	39.54	21.60	17.94	45.37
Construction Materials	36	39.79	24.78	15.01	37.73	39.86	22.21	17.65	44.28
Wholesale	73	32.01	20.18	11.82	36.94	32.07	18.19	13.89	43.30
Healthcare	41	27.38	17.15	10.23	37.35	27.22	15.73	11.49	42.22
Pharmaceutical Products	196	16.85	10.79	6.05	35.94	17.02	9.60	7.42	43.58
Telecommunication	60	23.03	13.61	9.42	40.91	22.61	14.08	8.53	37.73
Apparel	29	35.89	22.36	13.53	37.70	35.68	21.93	13.76	38.55
Printing and Publishing	14	29.07	18.17	10.90	37.50	28.73	17.87	10.86	37.82
Paper and Boxes	32	37.98	25.19	12.79	33.67	38.26	22.46	15.80	41.29
Oil	84	48.82	28.13	20.69	42.38	47.79	33.75	14.04	29.38
Auto	70	35.00	23.52	11.47	32.78	35.12	21.45	13.68	38.94
Food Products	37	18.06	12.19	5.87	32.50	18.26	11.12	7.14	39.08
Transportation	49	31.45	21.04	10.41	33.10	31.33	20.03	11.31	36.08
Computer Software	209	26.03	17.88	8.15	31.33	26.15	16.45	9.70	37.10
Construction	36	38.97	26.35	12.62	32.40	38.88	25.00	13.89	35.71
Recreation	13	26.53	18.38	8.15	30.72	26.56	17.12	9.44	35.54
Finance	320	37.79	26.10	11.69	30.94	37.78	24.72	13.07	34.59
Other	69	17.12	11.98	5.14	30.02	17.14	11.33	5.81	33.89
Steel	25	42.41	30.78	11.62	27.41	42.64	27.40	15.24	35.74
Electrical Equipment	32	26.97	19.86	7.11	26.37	27.17	17.72	9.45	34.79
Consumer Goods	32	30.83	23.82	7.01	22.74	31.07	21.40	9.68	31.15
Chemicals	64	32.95	24.35	8.60	26.11	32.85	23.81	9.04	27.51
Retail	124	23.42	17.77	5.65	24.13	23.45	16.86	6.58	28.08
Personal Services	31	26.33	20.76	5.57	21.15	26.17	21.06	5.11	19.53
Entertainment	29	34.30	29.05	5.25	15.31	34.28	28.44	5.85	17.06
Utilities	69	29.90	28.77	1.13	3.79	29.92	31.44	-1.52	-5.08

The table reports a decomposition of stock return variation into characteristic and textual components. For each panel, we regress individual stock returns on model-predicted returns from IPCA models estimated separately on 35 firm characteristics and on 64 conference call topic loadings, respectively. The number of latent factors in the IPCA estimation is set to 30. We report the average number of firms over the out-of-sample period (January 2008 to December 2022), the total return variation explained by both characteristic and textual information (R^2_{total}), the corresponding contributions of characteristic-only and text-only information (R^2_{char} and R^2_{text}), and the relative importance of textual information, measured as R^2_{text}/R^2_{total}. The first row reports the decomposition for the full sample of firms over the out-of-sample period. The remaining rows report similar decomposition results for the different industries. In Panel A, the textual information is obtained from a full-sample LDA model, whereas Panel B uses an online LDA model.

In addition to examining industry-level heterogeneity, we exploit cross-sectional variation in firm "styles" to test whether the value of textual information varies systematically across firms. Ex ante, fast-growing firms tend to derive much of their value from intangible assets and growth options that are not well captured by standard accounting characteristics. In this dimension, we consider indicators for sales and asset growth (high sale_gr1, at_gr1) and valuation ratios (at_me, at_be, be_me). Likewise, firms with strong recent performance

Table 9: Decomposition of Stock Return Variation and the Heterogeneous Role of Textual Information Across Firm Styles

Style	High characteristic values				Low characteristic values				
	R^2_{total} (%)	R^2_{char} (%)	R^2_{text} (%)	$\frac{R^2_{text}}{R^2_{total}}$ (%)	R^2_{total} (%)	R^2_{char} (%)	R^2_{text} (%)	$\frac{R^2_{text}}{R^2_{total}}$ (%)	$\Delta\frac{R^2_{text}}{R^2_{total}}$ (%)
	Panel A. Full-Sample LDA								
	Measures of Growth								
sale_gr1	27.93	17.48	10.45	37.43	27.67	19.78	7.89	28.51	8.91
at_gr1	29.46	18.85	10.60	36.00	26.43	18.50	7.93	29.99	6.01
at_me	31.21	22.01	9.20	29.47	23.90	14.81	9.08	38.01	8.54
be_me	30.36	21.37	8.99	29.61	24.79	15.53	9.26	37.36	7.75
at_be	28.60	20.05	8.55	29.89	26.95	17.05	9.90	36.74	6.85
	Measures of Recent Performance								
ret_60_12	28.48	18.19	10.29	36.13	27.21	19.14	8.07	29.67	6.46
ret_12_7	28.56	18.17	10.39	36.38	27.34	19.01	8.33	30.47	5.91
niq_at	33.25	21.04	12.21	36.72	25.58	17.46	8.12	31.75	4.98
niq_be	34.78	22.15	12.63	36.32	25.12	17.12	8.00	31.86	4.46
	Panel B. Online LDA								
	Measures of Growth								
sale_gr1	27.96	16.18	11.78	42.14	27.56	19.55	8.01	29.06	13.09
at_gr1	29.45	17.71	11.75	39.88	26.33	18.10	8.23	31.26	8.62
at_me	31.06	21.77	9.29	29.89	23.96	13.66	10.30	42.98	13.09
be_me	30.22	21.08	9.14	30.24	24.84	14.38	10.46	42.11	11.86
at_be	28.49	19.68	8.81	30.92	26.95	15.93	11.02	40.88	9.96
	Measures of Recent Performance								
ret_60_12	28.47	17.09	11.39	39.99	27.11	18.76	8.35	30.81	9.18
ret_12_7	28.54	17.13	11.41	39.97	27.26	18.45	8.81	32.31	7.66
niq_at	33.23	19.78	13.45	40.48	25.52	16.85	8.67	33.96	6.52
niq_be	34.80	20.66	14.14	40.63	25.04	16.61	8.44	33.69	6.94

The table reports a decomposition of stock return variation into characteristic and textual components over the out-of-sample period (January 2008 to December 2022). For each panel, we regress individual stock returns on model-predicted returns from IPCA models estimated separately on 35 firm characteristics and on 64 conference call topic loadings, respectively. The number of latent factors in the IPCA estimation is set to 30. We report the total return variation explained by both characteristic and textual information (R^2_{total}), the corresponding contributions of characteristic-only and text-only information (R^2_{char} and R^2_{text}), and the relative importance of textual information, measured as R^2_{text}/R^2_{total}. For each firm characteristic, we equally split firms into two groups: those with high and those with low rank-transformed characteristic values. In Panel A, the textual information is obtained from a full-sample LDA model, whereas Panel B uses an online LDA model.

are expected to be more heavily influenced by qualitative and forward-looking information. For these firms, the key question for investors is not whether current numbers are good, but whether current performance will persist or revert. We use both lagged cumulative returns (ret_60_12, ret_12_7) and measures of recent profitability (niq_at, niq_be) to quantify recent performance.

Table 9 shows that the relative importance of textual information is heterogeneous across firm "styles." For firms with high sales and asset growth, the fraction of explained variation attributable to text, R^2_{text}/R^2_{total}, is substantially higher than for low-growth firms (e.g., 36–42% vs. 28–31%). For valuation ratios the pattern reverses: firms with low assets-to-market, book-to-market, and assets-to-book ratios (i.e., growth-type firms) exhibit much larger text shares than high-ratio (value-type) firms (roughly 37–43% vs. 29–31%). Similarly, firms with strong recent performance—high past returns and high profitability—show a consistently higher contribution of text than their low-performance counterparts, with differences of about 4–9 percentage points.

Taken together, these industry and style tests indicate that the value of conference call text is highly concentrated in settings where fundamentals are especially forward-looking, intangible, and uncertain. In such cases, the "hard" information embedded in conventional characteristics is insufficient to summarize the return dynamics, making the "soft," qualitative information transmitted in conference calls particularly valuable.

5 Macro Information in Text-Managed MVE Portfolios

Past literature has shown that the textual information embedded in newspapers contains a forward-looking component that can forecast macro quantities (see Bybee et al. (2024), Bybee et al. (2023), and van Binsbergen et al. (2024)). Motivated by these papers, we go on to study whether the MVE portfolios conditional on Q&A and presentation sessions of conference calls can predict several important macro state variables. In particular, following Bybee et al. (2023), we study eight economic variables from FRED-MD: 1) real industrial production, 2) real personal consumption expenditures, 3) nonfarm employees, 4) total housing starts, 5) consumer price index, 6) unemployment rate, 7) BAA minus FED funds rate, and 8) 10-year treasury yield. For each variable Y_t , we regress its future one-month or one-year changes on the text-managed MVE portfolio returns. For the first five variables, we compute their log growth rates, i.e., $\log(Y_{t+1}/Y_t)$ or $\log(Y_{t+12}/Y_t)$. For the remaining three variables, we compute their first-order differences, i.e., $Y_{t+1} - Y_t$ or $Y_{t+12} - Y_t$.

Table 10 reports predictive regressions that use the text-managed MVE portfolios constructed from both the full-sample LDA (Panels A and B) and the online LDA (Panels C and D). At the one-month horizon (Panels A and C), the strongest and most robust result is for inflation. In the full-sample LDA (Panel A), the Q&A-managed portfolio significantly predicts next-month CPI growth, with a large coefficient and an adjusted R^2 of about 17%, whereas the presentation-managed portfolio exhibits weaker predictive content. The online LDA results (Panel C) confirm this pattern: both portfolios forecast CPI growth, but the Q&A portfolio is substantially stronger. This observation echoes our previous finding that "Inflation" is a common topic in both Q&A and presentation sessions. Both text-managed portfolios also negatively predict the BAA–Fed funds spread in both LDA specifications, with the Q&A portfolio again displaying the stronger and more robust effect. The consistently strong predictive power of the Q&A-managed portfolio for the BAA–Fed funds spread is in line with the fact that the "Capital Structure" topic appears only in the Q&A segment, and it suggests that manager-analyst discussions about financing and balance-sheet conditions contain information about credit markets in the short term.

At the 12-month horizon (Panels B and D), the Q&A-managed portfolio is particularly informative about the labor market and long-term interest rates. Using the full-sample LDA (Panel B), higher Q&A portfolio returns predict faster future growth in nonfarm employment and a lower unemployment rate over the next year, while the presentation-managed portfolio shows no comparable predictive power for these variables. The online LDA results (Panel D) reinforce this conclusion: the Q&A-managed portfolio significantly forecasts higher employment growth and lower long-term yields, whereas the presentation-based portfolio again does not predict labor market outcomes and only helps forecast the 10-year yield. The strong link between Q&A-based returns and future employment outcomes is consistent with our earlier finding that "Human Capital" is a distinctive Q&A topic.

Overall, text-managed portfolios—especially those based on Q&A topic loadings—contain economically and statistically significant forward-looking information about future inflation, labor market conditions, and financing conditions. This evidence complements the literature using newspaper text by showing that the qualitative information revealed in conference calls also anticipates key macroeconomic outcomes.

Table 10: Predicting Macro Variables

	Industrial Production	Personal Consumption	Non-farm Employment	Housing Starts	CPI	Unemploy Rate	BAA minus Fed Rate	10-year Yield
	Panel A. Predicting One-Month Changes, Full-Sample LDA							
	Predictor = Q&A-Managed MVE Portfolio							
Coefficient	0.085	0.078	0.153	0.055	0.425	-0.192	-0.302	-0.156
t-stat	(0.902)	(0.690)	(1.449)	(0.838)	(3.956)	(-1.502)	(-2.425)	(-1.675)
R^2	0.16%	0.05%	1.80%	-0.26%	17.57%	3.15%	8.64%	1.90%
	Predictor = Presentation-Managed MVE Portfolio							
Coefficient	0.077	0.010	0.027	0.111	0.171	-0.002	-0.224	-0.148
t-stat	(1.131)	(0.224)	(1.216)	(1.328)	(1.584)	(-0.065)	(-1.668)	(-1.404)
R^2	0.03%	-0.55%	-0.49%	0.67%	2.39%	-0.56%	4.50%	1.63%
	Panel B. Predicting 12-Month Changes, Full-Sample LDA							
	Predictor = Q&A-Managed MVE Portfolio							
Coefficient	0.109	0.076	0.210	0.097	0.202	-0.196	-0.102	-0.156
t-stat	(1.601)	(1.536)	(2.337)	(0.878)	(1.589)	(-2.267)	(-1.340)	(-1.815)
R^2	0.64%	0.02%	3.89%	0.39%	3.54%	3.30%	0.48%	1.90%
	Predictor = Presentation-Managed MVE Portfolio							
Coefficient	-0.060	-0.100	0.047	0.085	-0.098	-0.002	0.004	-0.157
t-stat	(-0.862)	(-0.735)	(0.305)	(1.035)	(-0.660)	(-0.010)	(0.046)	(-3.381)
R^2	-0.20%	0.45%	-0.34%	0.17%	0.41%	-0.56%	-0.56%	1.92%
	Panel C. Predicting One-Month Changes, Online LDA							
	Predictor = Q&A-Managed MVE Portfolio							
Coefficient	0.011	0.033	0.102	0.085	0.345	-0.101	-0.386	-0.207
t-stat	(0.200)	(0.557)	(1.710)	(1.370)	(2.439)	(-1.480)	(-2.713)	(-2.067)
R^2	-0.55%	-0.46%	0.48%	0.16%	11.42%	0.46%	14.42%	3.76%
	Predictor = Presentation-Managed MVE Portfolio							
Coefficient	0.047	0.046	0.077	0.101	0.241	-0.048	-0.308	-0.204
t-stat	(0.840)	(0.953)	(1.642)	(1.705)	(2.001)	(-1.049)	(-2.051)	(-1.781)
R^2	-0.34%	-0.35%	0.03%	0.46%	5.29%	-0.33%	8.98%	3.62%
	Panel D. Predicting 12-Month Changes, Online LDA							
	Predictor = Q&A-Managed MVE Portfolio							
Coefficient	0.043	-0.006	0.164	0.050	0.081	-0.114	-0.073	-0.194
t-stat	(0.601)	(-0.069)	(2.755)	(0.518)	(1.145)	(-1.652)	(-1.006)	(-3.740)
R^2	-0.38%	-0.56%	2.15%	-0.31%	0.11%	0.74%	-0.03%	3.22%
	Predictor = Presentation-Managed MVE Portfolio							
Coefficient	-0.094	-0.149	0.020	0.054	-0.094	0.044	-0.064	-0.265
t-stat	(-1.104)	(-0.874)	(0.115)	(0.458)	(-0.644)	(0.251)	(-0.447)	(-2.046)
R^2	0.32%	1.66%	-0.52%	-0.27%	0.33%	-0.37%	-0.15%	6.52%

The table reports the regression results in which the text-managed MVE portfolio returns are used to forecast the future one-month or one-year changes in eight economic variables. In the first five columns, we predict the log growth rates of the economic variables, i.e., $\log(Y_{t+1}/Y_t)$ or $\log(Y_{t+12}/Y_t)$. In columns 6–8, we forecast the first-order differences of economic variables, i.e., $Y_{t+1} - Y_t$ or $Y_{t+12} - Y_t$. We report the point estimates of the coefficient on the text-managed portfolio return, its t-statistic based on the Newey and West (1987) standard errors with 12 lags, and the adjusted R^2. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panels A and B and from the online LDA estimation in Panels C and D. The number of latent factors in the IPCA estimation is 25. The sample period ranges from January 2008 to December 2022.

6 Text vs. Characteristics in MVE Portfolios

This section investigates which dimensions of conference call text are most important for portfolio formation and how they compare with traditional firm characteristics. For interpretation, the analysis in this section is based on the full-sample LDA, which delivers stable, transparent, and interpretable topics.

First, to quantify the relative importance of the Q&A- and presentation-specific topics, we report in Table 11 the OOS Sharpe ratios of the MVE portfolios under two specifications: (i) conditioning on 19 Q&A common topic loadings and 32 presentation topic loadings (i.e., dropping 13 Q&A-specific topics), and (ii) conditioning on 19 presentation common topic loadings and 32 Q&A topic loadings (i.e., dropping 13 presentation-specific topics). For the dense conditional asset pricing models ($p \in \{25, 30\}$), removing either Q&A- or presentation-specific topics from the cross-sectional information set leads to lower out-of-sample Sharpe ratios compared with using all 64 topics. Hence, both Q&A and presentation sessions of conference calls contain common and session-specific information that is useful for building MVE portfolios.

Table 11: Common vs Specific Topics

Cross-Sectional Information Sets:	$p = 5$	$p = 10$	$p = 15$	$p = 20$	$p = 25$	$p = 30$
Drop 13 Q&A specific topics	0.72	0.67	0.91	1.16	1.20	1.20
(19 Q&A common topics plus 32 presentation topics)						
Drop 13 PRE specific topics	0.27	0.49	0.53	0.80	0.87	1.25
(19 presentation common topics plus 32 Q&A topics)						
64 Q&A + PRE topics	0.55	0.65	0.95	1.29	1.38	1.64

The table reports the annualized out-of-sample Sharpe ratio of the MVE portfolios implied by the IPCA model using as instruments the i) 19 Q&A common topic loadings plus 32 presentation topic loadings (see the row "Drop 13 Q&A specific topics") and ii) 19 presentation common topic loadings plus 32 Q&A topic loadings (see the row "Drop 13 PRE specific topics"). As a comparison, we also display the estimation results based on all the 64 Q&A plus presentation topic loadings. The Q&A and presentation topics come from the full-sample LDA estimation. The number of latent factors in the IPCA estimation, denoted by p, equals 5, 10, 15, 20, 25, and 30. The out-of-sample ranges from January 2008 to December 2022.

But which individual topics are important? To address this question, we conduct a topic-level leave-one-out analysis. Starting from the baseline IPCA specification with all 64 Q&A and presentation topics as instruments and 30 latent factors, we sequentially remove each topic and recompute the conditional MVE portfolio. For Q&A-specific and presentation-specific topics, we drop one topic at a time. For topics that are common across the two sessions (e.g., "Cloud Service"), we remove both the Q&A and presentation instances of that topic simultaneously. If a topic carries important information for mean-variance efficiency, excluding it should lead to a noticeable decline in the OOS Sharpe ratio.

Table 12 summarizes the most influential topics identified in this exercise. We report topics for which eliminating that topic reduces the annualized OOS Sharpe ratio by at least 0.10.[9] The benchmark MVE portfolio constructed from all 64 topics attains an out-of-sample Sharpe ratio of 1.64. Excluding the Q&A-specific "Capital Structure" topic produces the largest reduction, lowering the Sharpe ratio to 1.21 (a drop of 0.43). Among the more general topics, removing the presentation-specific topics "Business Operation" and "Organizational Management" also generates sizeable losses in performance. Within industry-related topics, removing "Cloud Service" (a common topic across sessions) reduces the Sharpe ratio to 1.41 (a drop of 0.23), while excluding "Steel" (presentation-specific) and "Semiconductor" (common) also meaningfully worsens performance. Overall, these results indicate that both general corporate themes (e.g., capital structure and operations) and industry-oriented discussions (e.g., cloud services and semiconductors) are key drivers that provide independently incremental economic value extracted from conference call text.

To complement the leave-one-out exercise, which asks how much out-of-sample performance deteriorates when a given instrument is removed, we also study how strongly the MVE portfolio loads on each instrument in each month. Concretely, for every month t in the out-of-sample period, we run a cross-sectional regression of the MVE portfolio weights on the contemporaneous conditional instruments. This yields a time series of regression coefficients and R^2 statistics.[10] Table 13 reports, for each instrument, the average regression coefficient and its 5th–95th percentile range across months. Figure 4 reports, over time, the fraction of cross-sectional variation in MVE weights explained by different groups of instruments. This regression-based decomposition allows us to quantify the stability and relative importance

[9]Dropping some topics can occasionally lead to higher OOS Sharpe ratios. Unlike in-sample fit measures, the OOS Sharpe ratio does not mechanically increase with the inclusion of additional instruments.

[10]To mitigate the issue of extreme collinearity among regressors, we address topics that appear in both sessions (e.g., "Cloud Service") by computing the equally weighted average of their topic loadings and using this average as a single regressor. Similarly, for firm characteristics exhibiting high pairwise correlations (above 0.7), we group these correlated variables into the same category and construct their average to serve as a representative regressor.

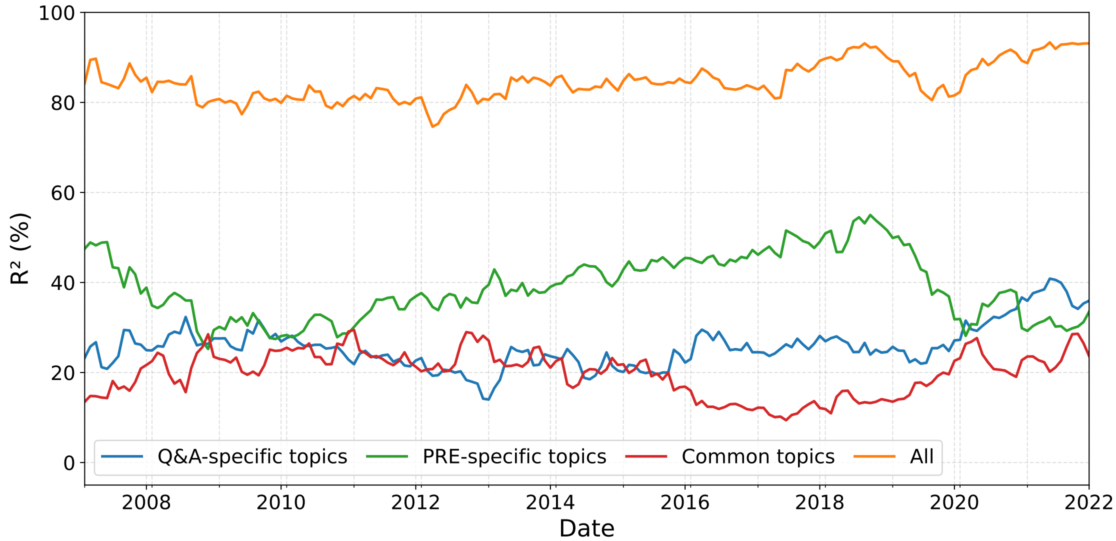
Table 12: Topic Importance: Leave-one-out Analysis

64 Q&A + PRE topics, excluding:	SR_{oos}	ΔSR_{oos}
	1.64	
Panel A. General topics		
Capital Structure (Q&A-specific)	1.21	-0.43
Business Operation (PRE-specific)	1.46	-0.18
Organizational Management (PRE-specific)	1.46	-0.18
Panel B. Industry		
Cloud Service (Common)	1.41	-0.23
Steel (PRE-specific)	1.51	-0.13
Semiconductor (Common)	1.54	-0.10

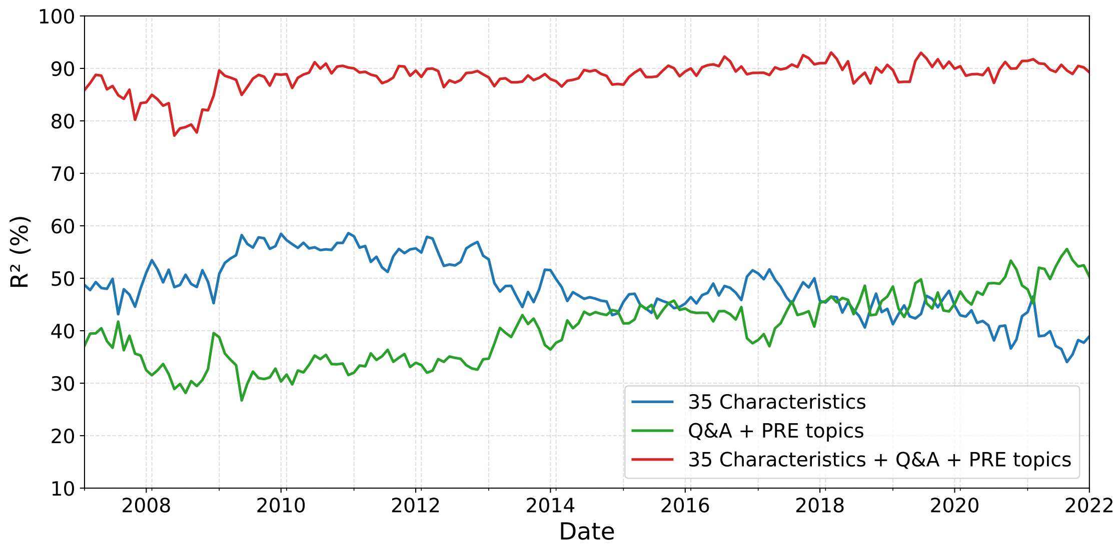
The table reports the topic importance in conference calls. For the Q&A-specific and presentation-specific topics, we sequentially exclude each topic one at a time. For topics that are common to both sessions, we exclude the two session-specific instances of that topic (e.g., the two "Cloud Service" topics in the Q&A and presentation) from the IPCA estimation. In column SR_{oos}, we report the annualized Sharpe ratio of the out-of-sample MVE portfolio based on the remaining 62 topics (when excluding two common topics) or 63 topics (when excluding one session-specific topic), as well as the drop in the Sharpe ratio relative to the benchmark analysis in Table 4 (see column ΔSR_{oos}). We list the top six sources of textual information that generate the largest drops in the out-of-sample Sharpe ratio. The Q&A and presentation topics come from the full-sample LDA estimation. The number of latent factors in the IPCA estimation is 30. The out-of-sample ranges from January 2008 to December 2022.

of specific topics and characteristics in the MVE portfolios.

When conditioning only on textual information (Panel A of Table 13 and Panel A of Figure 4), a small set of topics emerges as especially influential in the MVE portfolio. The Q&A-specific "Capital Structure" topic has the largest average coefficient, while presentation-specific topics "Business Operation," "Steel," and "Organizational Management," together with the common "Semiconductor and Electronics" and "Cloud Service" topics, also exhibit sizable coefficients with relatively tight 5th–95th percentile bands. This indicates that both the sign and magnitude of their impact on portfolio weights are stable over time. Aggregating across topics, Panel A of the figure shows that Q&A-specific topics explain on average about 26% of the cross-sectional variation in MVE weights, presentation-specific topics explain about 39%, and common topics explain about 20%, so that all topics together account for roughly 85% of the variation. Thus, both session-specific and common themes play persistently important roles in the MVE portfolio, consistent with the earlier leave-one-out evidence that these same topics are crucial for out-of-sample performance.



(a) MVE portfolio conditional on textual information



(b) MVE portfolio conditional on textual and characteristic information

Figure 4: Decomposition of the MVE Portfolio Weights

The figure shows the decomposition of the MVE portfolio weights implied by the IPCA model over the out-of-sample period from January 2008 to December 2022. The number of latent factors in the IPCA estimation is 30. In each month t, we regress the MVE portfolio weights on the contemporaneous conditional instruments. All instruments are standardized each month to have the same cross-sectional standard deviation of 1%. In Panel A, we consider the IPCA model with both Q&A and presentation topic information from the full-sample LDA estimation. In the regression analysis, we treat the loadings on the Q&A-specific and presentation-specific topics as individual regressors. For topics that appear in both sessions (e.g., "Cloud Service"), we compute the equally weighted average of their topic loadings and use this as a regressor. We report the percentages of cross-sectional variation in MVE portfolio weights explained by Q&A-specific, presentation-specific and common topics. In Panel B, we consider the IPCA model with both textual and characteristic information. We group firm characteristics with pairwise correlation coefficients above 0.7 into the same category and construct the equally weighted average of the characteristics in each category. We display the percentages of cross-sectional variation in MVE portfolio weights explained by textual and characteristic information, respectively.

Table 13: Importance of Conditional Information in MVE Portfolio Weights

Panel A. Q&A + PRE topics

	mean	5th	95th
Capital Structure (QA-specific)	0.59	0.51	0.74
Business Operation (PRE-specific)	0.46	0.18	0.67
Steel (PRE-specific)	0.43	0.30	0.50
Organizational Management (PRE-specific)	-0.37	-0.60	-0.23
Semiconductor and Electronics (Common)	-0.37	-0.52	-0.20
Cloud Service (Common)	0.36	0.14	0.59
Financial Reporting (PRE-specific)	0.30	0.20	0.42
International Finance (PRE-specific)	-0.29	-0.40	-0.17
Entertainment (Common)	-0.23	-0.36	0.10
Energy (PRE-specific)	-0.22	-0.33	0.04
Metals and Mining (PRE-specific)	-0.19	-0.31	0.03
Fiscal Periods (PRE-specific)	0.17	0.04	0.28
Inflation (Common)	-0.16	-0.32	0.04
Oil (Common)	0.17	-0.05	0.32
Banking and Lending (Common)	-0.15	-0.26	-0.04
Freight (PRE-specific)	0.16	0.05	0.24
Human Capital (QA-specific)	-0.11	-0.38	0.04
Asset Management (PRE-specific)	-0.14	-0.24	-0.02
Medical Devices (Common)	0.10	0.00	0.32
Renewable Energy (Common)	-0.05	-0.21	0.18

Panel B. Q&A + PRE topics + Characteristics

	mean	5th	95th
(prc, dolvol_126d, market_equity)	-0.47	-0.65	-0.32
Business Operation (PRE-specific)	0.48	0.24	0.64
(ebit_bev, ni_be, niq_be, niq_at)	0.40	0.19	0.74
(at_me, be_me)	0.41	0.21	0.66
gp_at	0.40	0.28	0.54
ret_1_0	-0.38	-0.55	-0.26
prc_highprc_252d	0.33	0.12	0.55
Organizational Management (PRE-specific)	-0.35	-0.45	-0.16
ni_me	-0.30	-0.42	-0.17
International Finance (PRE-specific)	-0.31	-0.41	-0.12
Financial Reporting (PRE-specific)	0.24	0.17	0.37
Semiconductor and Electronics (Common)	-0.24	-0.36	-0.13
ret_12_1	0.15	-0.27	0.44
Pharmaceutics (Common)	0.21	0.10	0.32
Oil (Common)	0.22	0.02	0.35
ret_12_7	-0.16	-0.40	-0.01
Human Capital (QA-specific)	-0.17	-0.33	-0.05
ret_6_1	-0.18	-0.34	0.05
ivol_ff3_21d	-0.18	-0.30	-0.06
Capital Structure (QA-specific)	0.15	0.08	0.27

The table reports the importance of each conditional instrument in the MVE portfolio implied by the IPCA model. The number of latent factors in the IPCA estimation is 30. In each month t, we regress the MVE portfolio weights on the contemporaneous conditional instruments. We report the top 20 conditional instruments with the largest average regression coefficients in absolute value over the out-of-sample period from January 2008 to December 2022. For each reported instrument, we display the average regression coefficient and the 5th–95th percentile range across the 180 out-of-sample months. In Panel A, we consider the IPCA model with both Q&A and presentation topic information from the full-sample LDA estimation. In the regression analysis, we treat the loadings on the Q&A-specific and presentation-specific topics as individual regressors. For topics that appear in both sessions (e.g., "Cloud Service"), we compute the equally weighted average of their topic loadings and use this as a regressor. In Panel B, we consider the IPCA model with both textual and characteristic information. We group firm characteristics with pairwise correlation coefficients above 0.7 into the same category and construct the equally weighted average of the characteristics in each category. All variables are standardized each month to have the same cross-sectional standard deviation of 1%.

After adding firm characteristics, standard predictors naturally absorb some of the explanatory power, but several topics remain important (Panel B of Table 13 and Panel B of Figure 4). In particular, the presentation-specific topics "Business Operation," "Organizational Management," "International Finance," and "Financial Reporting," as well as the common topics "Semiconductor and Electronics" and "Oil" and the Q&A-specific topics "Human Capital" and "Capital Structure," still appear among the top instruments, with economically meaningful coefficients that are stable over time. Panel B of the figure quantifies these contributions: on average, characteristics alone explain about 48% of the cross-sectional variation in MVE weights, text alone explains about 40%, and characteristics and text together explain about 88%. Moreover, the share explained by text rises over time—above 50% in recent years—relative to that explained by characteristics, indicating that soft information extracted from conference calls becomes increasingly important.

7 Conclusion

This paper studies the role of soft, qualitative information in conditional asset pricing using the text of earnings conference calls. We apply topic models to firms' scripted presentations and Q&A sessions to construct interpretable, low-dimensional textual signals and embed them in an IPCA framework. We show that this qualitative information is as crucial as standard firm characteristics and provides incremental value for conditional asset pricing.

First, conference call text substantially improves mean-variance efficiency, even though it adds little incremental power for forecasting expected returns. When used as conditioning information in IPCA, topic loadings from conference call text deliver Sharpe ratios comparable to those from traditional firm characteristics. Combining text with characteristics further generates large out-of-sample gains. These gains operate through the covariance channel: augmenting the conditioning set with textual information captures the conditional covariance matrix better, while the improvement in expected-return forecasts is modest.

Second, conference call text explains an economically meaningful share of the common variation in stock returns. A decomposition of the fitted common component shows that text accounts for roughly one-third of the variation explained by the conditional asset pricing model. Its contribution is especially large for firms and industries whose fundamentals are intangible, rapidly evolving, or hard to summarize with standard accounting- or return-based variables—such as high-growth or innovation-intensive sectors. For these harder-to-value firms, qualitative and forward-looking disclosures are particularly valuable.

Third, conference call text has become a central determinant of optimal portfolio allocation. In the cross-section of conditional MVE portfolio weights, textual topics explain almost as much variation as traditional characteristics on average, and their importance has grown over time. In recent years, the contribution of text to explaining weight variation even surpasses that of standard characteristics, indicating that soft information from conference calls now plays a salient role in portfolio formation.

Overall, our evidence indicates that soft, qualitative information plays a first-order and distinct role in conditional asset pricing. Future research can extend this approach to other qualitative sources and assess their value for modeling conditional asset pricing.

References

Aleti, Saketh, and Tim Bollerslev, 2024, News and Asset Pricing: A High-Frequency Anatomy of the SDF, *Review of Financial Studies* .

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Balduzzi, Pierluigi, and Cesare Robotti, 2008, Mimicking portfolios, economic risk premia, and tests of multi-beta models, *Journal of Business & Economic Statistics* 26, 354–368.

Bellstam, Gustaf, Sanjai Bhagat, and J Anthony Cookson, 2021, A text-based analysis of corporate innovation, *Management Science* 67, 4004–4031.

Blei, David M, Andrew Y Ng, and Michael I Jordan, 2003, Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1022.

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard, 2024, Macro strikes back: Term structure of risk premia, *Available at SSRN 4752696* .

Bybee, Leland, Bryan Kelly, and Yinan Su, 2023, Narrative asset pricing: Interpretable systematic risk factors from news text, *Review of Financial Studies* 36, 4759–4787.

Bybee, Leland, Bryan T Kelly, Asaf Manela, and Dacheng Xiu, 2024, Business news and business cycles, *Journal of Finance, forthcoming* .

Chen, AJ, Gerard Hoberg, and Miao Ben Zhang, 2025, Haven't we seen this before? return predictions from 200 years of news, *Available at SSRN 5380343* .

Chen, Yifei, Bryan T Kelly, and Dacheng Xiu, 2022, Expected returns and large language models, *Available at SSRN 4416687* .

Das, Sanjiv Ranjan, et al., 2014, Text and context: Language analytics in finance, *Foundations and Trends® in Finance* 8, 145–261.

Engelberg, Joseph, Asaf Manela, William Mullins, and Luka Vulicevic, 2025, Entity neutering, *Available at SSRN 5182756* .

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2020, Hedging Climate Change News, *Review of Financial Studies* 33, 1184–1216.

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

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

Fama, Eugene F, and Kenneth R French, 2020, Comparing Cross-Section and Time-Series Factor Models, *The Review of Financial Studies* 33, 1891–1926.

Ferson, Wayne E, and Campbell R Harvey, 1999, Conditioning variables and the cross section of stock returns, *Journal of Finance* 54, 1325–1360.

Ferson, Wayne E, and Andrew F Siegel, 2001, The efficient use of conditioning information in portfolios, *Journal of Finance* 56, 967–982.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics nonparametrically, *Review of Financial Studies* 33, 2326–2377.

García, Diego, 2013, Sentiment during recessions, *Journal of Finance* 68, 1267–1300.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, 2019, Text as data, *Journal of Economic Literature* 57, 535–574.

Giglio, Stefano, Theresa Kuchler, Johannes Stroebel, and Xuran Zeng, 2024, Biodiversity risk, *Available at SSRN 4410107* .

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129, 1947–1990.

Glasserman, Paul, and Caden Lin, 2023, Assessing look-ahead bias in stock return predictions generated by gpt sentiment analysis, *arXiv preprint arXiv:2309.17322* .

Goldstein, Itay, Chester S Spatt, and Mao Ye, 2021, Big data in finance, *Review of Financial Studies* 34, 3213–3225.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2019, Dynamic interpretation of emerging risks in the financial sector, *Review of Financial Studies* 32, 4543–4603.

Hansen, Lars Peter, and Scott F Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 587–613.

Hansen, Stephen, Michael McMahon, and Andrea Prat, 2018, Transparency and deliberation within the fomc: A computational linguistics approach, *Quarterly Journal of Economics* 133, 801–870.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, Markus Schwedeler, and Ahmed Tahoun, 2023, Firm-level exposure to epidemic diseases: Covid-19, sars, and h1n1, *Review of Financial Studies* 36, 4919–4964.

Hassan, Tarek A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *The Quarterly Journal of Economics* 134, 2135–2202.

He, Songrun, Linying Lv, Asaf Manela, and Jimmy Wu, 2025, Chronologically consistent large language models, *arXiv preprint arXiv:2502.21206* .

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Hoffman, Matthew, Francis Bach, and David Blei, 2010, Online learning for latent dirichlet allocation, *Advances in neural information processing systems* 23.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Huang, Allen H, Reuven Lehavy, Amy Y Zang, and Rong Zheng, 2018, Analyst information discovery and interpretation roles: A topic modeling approach, *Management Science* 64, 2833–2855.

Huberman, Gur, Shmuel Kandel, and Robert F Stambaugh, 1987, Mimicking portfolios and exact arbitrage pricing, *The Journal of Finance* 42, 1–9.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen, 2023, Is there a replication crisis in finance?, *Journal of Finance* 78, 2465–2518.

Ke, Tracy Z, Bryan T Kelly, and Dacheng Xiu, 2022, Predicting returns with text data, *Available at SSRN 4416687* .

Kelly, Bryan, and Seth Pruitt, 2013, Market expectations in the cross-section of present values, *Journal of Finance* 68, 1721–1756.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Kozak, Serhiy, and Stefan Nagel, 2024, When do cross-sectional asset pricing factors span the stochastic discount factor?, *Available at SSRN 4443643* .

Lamont, Owen A, 2001, Economic tracking portfolios, *Journal of Econometrics* 105, 161–184.

Larsen, Vegard H., and Leif A. Thorsrud, 2019, The value of news for economic developments, *Journal of Econometrics* 210, 203–218.

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

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional capm does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289–314.

Lopez-Lira, Alejandro, 2023, Risk factors that matter: Textual analysis of risk disclosures for the cross-section of returns, *Available at SSRN 3313663* .

Loughran, Tim, and Bill McDonald, 2020, Textual analysis in finance, *Annual Review of Financial Economics* 12, 357–375.

Manela, Asaf, and Alan Moreira, 2017, News implied volatility and disaster concerns, *Journal of Financial Economics* 123, 137–162.

McCracken, Michael W, and Serena Ng, 2016, Fred-md: A monthly database for macroeconomic research, *Journal of Business & Economic Statistics* 34, 574–589.

Nagel, Stefan, and Kenneth J Singleton, 2011, Estimation and evaluation of conditional asset pricing models, *Journal of Finance* 66, 873–909.

Newey, Whitney K, and Kenneth D West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Rosenberg, Barr, 1974, Extra-market components of covariance in security returns, *Journal of Financial and Quantitative Analysis* 9, 263–274.

Sarkar, Suproteem K, and Keyon Vafa, 2024, Lookahead bias in pretrained language models, *Available at SSRN 4754678* .

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

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

Tetlock, Paul C, Maytal Saar-Tsechansky, and Sofus Macskassy, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437–1467.

van Binsbergen, Jules H, Svetlana Bryzgalova, Mayukh Mukhopadhyay, and Varun Sharma, 2024, (almost) 200 years of news-based economic sentiment, *Available at SSRN 4261249* .

Appendices

Appendix A Preprocessing Conference Call Data

We select the conference call transcripts of the US publicly listed firms from the full sample provided by Refinitiv StreetEvents. That is, we keep 371,659 conference call transcripts among the original 542,472 unique transcripts during the sample period 2003–2022. We further follow Hassan et al. (2023) to keep earnings conference calls since they typically contain both a presentation and a Q&A session, the main focus of our paper. Using this filter, we obtain 240,723 earnings conference call transcripts.

We conduct standard procedures to preprocess the conference call data following the previous literature (see, e.g., Hansen et al. (2018); Huang et al. (2018); Bybee et al. (2023, 2024)). Specifically, we remove non-alphabetic characters and stop words. This step is conducted using the Python package "NLTK". Next, we remove narrative words (e.g., would, would like to), person names, greeting words, and common narratives (such as "hello", "slide", and "please") from conference call texts. Moreover, to account for grammatical variations of English words, we conduct a standard light lemmatizing of derivative words. The rules below are applied in the order given:

(a) Replace trailing "sses" with "ss".

(b) Replace trailing "ies" with "y".

(c) Remove trailing "s", "ly", and "ed" sequentially, and replace remaining "ed" with "e".

(d) Replace trailing "ing" with "e". For remaining trailing "ing" that follow a pair of identical consonants, remove "ing" and one consonant. Remove remaining trailing "ing".

(e) Remove words with less than 3 letters.

It is well-known in textual analysis that bi-grams (a combination of adjacent single words) can capture important semantics beyond uni-grams. Therefore, we generate a set of bi-grams as all pairs of (ordered) adjacent uni-grams and include them in the bag of words. Finally, as too frequent or infrequent words generally contain no information for understanding the topic structure underlying the textual data, we remove words that appear in more than 10% (or less than 0.1%) of the entire sample of conference call transcripts.

Appendix B Algorithm for Finding Stable Topics

This section describes the algorithm for finding stable topics in the LDA estimation across different random seeds that initialize the model parameters in the variational Bayesian estimation. Throughout this section, we demonstrate the procedure using the Q&A topics, but the same algorithm is applied to the presentation topics.

As Figure 1 indicates, the data likelihood, averaged across 10 random seeds, is the highest (equivalently, the perplexity score is the lowest) when the number of Q&A topics is 40. We, therefore, collect all 400 Q&A topics (40 topics across 10 random seeds) and try to find stable topics among them. Note that if a topic is salient and stable, it should show up in most of the 10 random seeds. Conversely, the inconsequential topics, which are driven mainly by estimation noise, should appear in a small proportion of random seeds. Motivated by this logic, we design the following algorithm to find stable topics.

Step 1. We collect all the LDA topics across 10 random seeds and denote the set of topics by $\mathcal{L}$. Let $|\mathcal{L}|$ denote the total number of candidate topics (e.g., 400 in the Q&A session). In this step, we keep the estimates of probability distributions over terms, $\boldsymbol{\Theta}$, as well as the average loadings on all $|\mathcal{L}|$ topics, $\boldsymbol{\phi}$, with $\sum_l \phi_l = 1$. Note that $\boldsymbol{\Theta}$ is a $|\mathcal{L}| \times |V|$ matrix, with the l-th row denoting the probability distribution over terms in topic l.

Step 2. We search for stable topics based on the pairwise correlation between two topics' probability distributions over terms. Specifically, we compute $\text{corr}(\boldsymbol{\theta}_p, \boldsymbol{\theta}_q)$, where $\boldsymbol{\theta}_l$ is the l-th row of $\boldsymbol{\Theta}$. Note that $\boldsymbol{\theta}_l$ tends to be a sparse vector; hence, we compute the pairwise correlations using only the top 1% most frequent words implied by the LDA model.

Step 3. For each iteration, we pick the pairs $\boldsymbol{\theta}_{p^\star}$ and $\boldsymbol{\theta}_{q^\star}$ that have the highest correlation, and combine them into one new topic r, as follows:

$$\boldsymbol{\theta}_r = \frac{\boldsymbol{\theta}_{p^\star}\phi_{p^\star} + \boldsymbol{\theta}_{q^\star}\phi_{q^\star}}{\phi_{p^\star} + \phi_{q^\star}}, \quad \phi_r = \phi_{p^\star} + \phi_{q^\star}. \tag{A1}$$

Merging two topics as in equation (A1) ensures the coherent probabilistic interpretation of the model estimates; that is, $\boldsymbol{\theta}_r$ remains as a probability distribution over terms in the corpus. After the successful combination of two topics, we will remove $\boldsymbol{\theta}_{p^\star}$ and $\boldsymbol{\theta}_{q^\star}$ from $\boldsymbol{\Theta}$ but incorporate $\boldsymbol{\theta}_r$ as a new row in $\boldsymbol{\Theta}$. This step will be repeated until $\text{corr}(\boldsymbol{\theta}_p, \boldsymbol{\theta}_q)$ is less than a certain threshold κ (which is currently chosen to be 0.6).

Step 4. We define stable topics as the rows in $\boldsymbol{\Theta}$ that are merged no less than five times in Step 3. This step requires a stable topic to appear at least five times across ten random seeds.

Appendix C Additional Tables

Table A1: List of 35 Firm Characteristics

Characteristic Code	Characteristic Name in Jensen et al. (2023)	Matched Characteristic Name in Kelly et al. (2019)
Panel A. Exactly matched characteristics		
at_me	assets-to-market	assets-to-market
bidaskhl_21d	the high-low bid-ask spread	bid-ask spread
be_me	book-to-market equity	book-to-market
ni_me	earnings-to-price	earnings-to-price
ivol_ff3_21d	idiosyncratic volatility from FF3	idiosyncratic volatility with respect to FF3
at_gr1	asset growth	investment
at_be	book leverage	leverage
ret_60_12	long-term reversal	long-term reversal
beta_60m	market beta	market beta
market_equity	market equity	market capitalization
noa_at	net operating assets	net operating assets
oaccruals_at	operating accruals	operating accruals
opex_at	operating leverage	operating leverage
prc_highprc_252d	current price to high price over last year	price relative to its 52-week high
ebit_sale	profit margin	profit margin
niq_at	quarterly return on assets	return on assets
ni_be	return on equity	return on equity
niq_be	quarterly return on equity	return on equity
ebit_bev	return on net operating assets	return on net operating assets
ret_1_0	short-term reversal	short-term reversal
turnover_126d	share turnover	turnover
Panel B. Closely related characteristics		
fcf_me	free cash flow-to-price	cash flow-to-book
sti_gr1a	change in short-term investments	cash-to-short-term-investment
cash_at	cash-to-assets	cash-to-short-term-investment; total assets
gp_at	gross profits-to-assets	gross profitability
ppeinv_gr1a	change PPE and inventory	ratio of change in PPE to the change in total assets
ret_12_1	price momentum t-12 to t-1	return from 12 to 2 months before prediction
ret_6_1	price momentum t-6 to t-1	return from 6 to 2 months before prediction
ret_12_7	price momentum t-12 to t-7	return from 12 to 2 months before prediction
sale_gr1	sales growth	sales-to-assets
prc	price per share	sales-to-price
sale_me	sales-to-market	sales-to-price
dsale_dsga	change sales minus change SG&A	the ratio of SG&A costs to sales
dolvol_var_126d	coefficient of variation for dollar trading volume	unexplained volume
dolvol_126d	dollar trading volume	unexplained volume

The table reports a list of 35 firm characteristics used in our empirical analyses. Building on the large dataset provided by Jensen et al. (2023), we select the 35 characteristics used in Kelly et al. (2019) (and originally in Freyberger et al. (2020)). Panel A presents the firm characteristics that exactly match—namely, those with the same definition in both Jensen et al. (2023) and Kelly et al. (2019). Panel B displays additional characteristics that are close substitutes for those used in Kelly et al. (2019).

Table A2: Key Terms in Q&A and Presentation Topics

Topic name	**Loading**	**Key terms**
		Panel A. Q&A topics
Forecasting and Estimation	0.105	ballpark, delta, time line, baseline, go forward, 3year, sure understand, exact number, missing, contemplate, bit lower, still working, pretty close, somewhere around, noise
Economic Uncertainty	0.066	recession, ball, worried, worry, sudden, throw, make money, shut, whole lot, lot different, bet, massive, amazing, lot money, know exactly, still lot, crystal
Macro Environment	0.061	backdrop, investor day, macro environment, thoughtful, pretty consistent, helpful followup, resilient, normalization, macroeconomic, delta, great followup, framework, data point
Growth Outlook	0.050	midsingle, midsingle digits, automation, high single, investor day, midsingledigit, midteens, high singledigit, margin profile, teens, broadbased, yearonyear, inflationary
Litigation and Negotiation	0.043	time line, burn, language, pilot, patent, timeline, restrictions, royalty, cash burn, letter, dialogue, resolution, resolved, resolve, nearly future, finalize, negotiating
Optimism	0.042	incredible, amazing, massive, thoughtful, great opportunities, solve, tech, team done, come together, runway, lot different, pivot, phenomenal, evolution, early day, tremendous amount
Capital Structure	0.042	covenants, share buyback, return capital, stock price, leverage ratio, capital shareholders, excess cash, terms loan, share price, deleveraging, cost capital, capital market, return shareholders
Pharmaceutics	0.029	dose, therapy, disease, phase iii, efficacy, cohort, enrollment, endpoint, cell, enrolled, cancer, physicians, patients population, clinical trial, tumor
Cloud Service	0.029	cloud, subscription, gotomarket, billings, use case, hardware, logos, sales cycle, smb, new logos, motion, ecosystem, upsell, automation, recurring revenue
Supply Chain	0.025	lead time, shortages, chain issues, cancellations, order book, ordering, automation, chinese, shutdown, absorption, order come, constrained, assembly, supply constraints, meet demand
Inflation	0.023	inflationary, steel, inflationary pressure, price cost, tariffs, input cost, freight, price actions, shortages, volume growth, material cost, cost inflation, resin, metal, increase price
Banking and Lending	0.023	mortgage, originations, loan growth, lending, lenders, borrowers, underwriting, chargeoffs, nim, rate environment, delinquencies, refi, multifamily, beta, deferral, cre, ltv, hikes, deposits growth
Retail and E-commerce	0.016	ecommerce, assortment, merchandise, apparel, women, backtoschool, freight, aur, markdowns, men, shopping, holiday season, new stores, footwear, outdoor
Medical Devices	0.014	reps, surgeons, surgery, physicians, doctors, implant, sales reps, reimbursement, surgical, pump, tissue, knee, consumables, spine, lab, heart, dental, pain, new accounts, wound
Global Markets	0.014	russia, brazil, emea, france, globe, euro, chinese, emerging market, pacific, part world, asia pacific, korea, middle east, italy, rest world, constant currency, ukraine, africa, international business, spain
Renewable Energy	0.013	renewable, solar, electric, hydrogen, carbon, grid, rate case, transmission, megawatts, coal, battery, legislation, emissions, nuclear, reliability
Human Capital	0.013	attrition, tech, health care, wage, new clients, consulting, staffing, clients base, recruiting, merchant, payroll, existing clients, financial services, employers, outsourcing, labor market
Oil	0.012	rigs, drilling, basin, drill, permian, barrels, frac, acreage, lateral, pad, sand, eagle, delaware, midstream, oil price, exploration, commodity price, crew, zone, gulf
Household Goods	0.012	skus, private label, pet, shelf, consumption, elasticity, household, walmart, beauty, ecommerce, mass, grocery, consumer demand, pack, retail partners
Semiconductor	0.010	semiconductor, wafer, chip, memory, foundry, fab, node, semi, laser, silicon, dram, display, logic, inspection, sensors
Transportation: Manufacturing	0.010	dealers, truck, aircraft, boat, electric, dealerships, new car, auction, rental, aftermarket, airlines, defense, ford, aerospace, new vehicle
Health Care	0.009	payers, physicians, medicare, health care, reimbursement, nursing, pharmacy, doctors, medicaid, clinic, acute, health system, primary care, membership, outpatient, clinicians
Insurance	0.009	agents, underwriting, reinsurance, loss ratio, frequency, severity, workers, broker, cat, homeowners, accident, expenses ratio, net retention, institutional, workers comp, book business, annuity
Transportation: Service	0.009	freight, truck, miles, rail, intermodal, container, port, terminal, surcharge, rental, wage, airlines, truckload, hub, transit, trucking
Education	0.008	school, students, education, rigs, enrollment, university, offshore, vessels, campus, college, cohort, district, campuses, graduate, attrition
Electronics	0.007	chip, battery, asp, modules, sensors, cell, silicon, camera, display, laser, computing, design win, hardware, semiconductor, electric
Catering	0.007	restaurants, guests, franchisees, menu, wage, dining, staffing, samestore, samestore sales, chicken, breakfast, beef, kitchen, pizza, loyalty
Defense	0.006	defense, aircraft, aerospace, military, aftermarket, airlines, boeing, flight, aviation, max, satellite, navy, airplanes, army, pilot
Entertainment	0.006	sports, advertisers, audience, streaming, title, entertainment, political, studio, film, theaters, monetization, television, stations, subscription, programming, radio
Biotech	0.006	lab, instruments, pharma, diagnostic, sample, cell, vaccine, consumables, assay, panel, screening, sequencing, placement, academic, life science
Telecommunication	0.005	churn, subscribers, arpu, broadband, fiber, satellite, cable, sub, spectrum, wireless, streaming, tmobile, tower, puerto, rico
Hospitality	0.005	hotel, occupancy, vegas, rental, casino, club, rent, leisure, hospitality, resort, san, revpar, hawaii, vacation, travelers

Panel B. Presentation topics		
Profits	0.049	net revenue, compared quarter, stockholders, period increase, approximately compared, primarily attributable, period prior, increased approximately, expenses third, decrease primarily, adjusted net, decreased compared, revenue third, salaries, offset increase
Organizational Management	0.049	leadership team, stakeholders, pillars, prioritize, bestinclass, transform, strategic priority, streamline, profitable growth, team continue, value creation, team members, unlock, aimed, worldclass
Inflation	0.043	inflationary, price actions, material cost, freight, inflationary pressure, ebit, lower volume, lower sales, input cost, chemical, organic sales, shortages, selling price, cost inflation, volume growth
Pharmaceutics	0.034	disease, cancer, therapy, dose, cell, clinical trial, phase iii, therapeutic, cohort, tumor, efficacy, inhibitor, enrollment, preclinical, endpoint
Events	0.033	amazing, incredible, deck, remarkable, outsized, massive, extraordinary, crisis, friends, war, bright, recession, surge, alltime, letter
Business Operation	0.030	adjusted operating, growth across, driven strong, investor day, macroeconomic, broadbased, profitable growth, point yearoveryear, year driven, macro environment, return shareholders, industryleading, backdrop
Banking and Lending	0.028	mortgage, originations, loan growth, noninterest, lending, interest margin, chargeoffs, loan portfolio, linked quarter, credit loss, tangible, noninterest income, noninterest expenses, total loan, net chargeoffs
Cloud Service	0.027	cloud, subscription, billings, recurring revenue, gotomarket, automation, ecosystem, subscription revenue, intelligence, workflow, hardware, hybrid, professional services
Financial Reporting	0.024	nongaap operating, nongaap net, nongaap basis, quarter nongaap, nongaap gross, gaap net, nongaap earnings, reconciliation gaap, expect nongaap, gaap operating, nongaap results
Retail and E-commerce	0.022	merchandise, ecommerce, assortment, new stores, shopping, stores sales, fulfillment, loyalty, apparel, instore, comparable stores, comparable sales, associates, margin rate, freight
Semiconductor and Electronics	0.021	semiconductor, sensors, chip, battery, display, silicon, modules, optical, wafer, design win, data center, laser, foundry, memory, fab, nextgeneration
International Finance	0.021	constant currency, currency basis, emea, pacific, adjusted operating, asia pacific, brazil, russia, translation, reported basis, impact foreign, local currency, euro, currency translation, currency exchange
Legal and Regulatory Matters	0.016	reliance, company website, warrants, undue, undue reliance, new future, except required, place undue, update revise, per basic, historically fact, events otherwise, including limited, whether results, securities law
Renewable Energy	0.016	renewable, electric, emissions, megawatts, solar, grid, carbon, transmission, rate case, reliability, renewable energy, nuclear, regulated, electricity, storm
Oil	0.014	per day, drilling, barrels, basin, drill, boe, barrels oil, rigs, lateral, permian, eagle, commodity price, acreage, oil equivalents, pad
Fiscal Periods	0.013	fiscal quarter, half fiscal, end fiscal, year fiscal, fiscal fourth, fiscal first, fiscal compared, fiscal second, fiscal third, compared fiscal, full fiscal, fiscal expect, revenue fiscal, months fiscal, guidance fiscal
Defense	0.013	defense, aerospace, military, aircraft, satellite, aviation, booktobill, aftermarket, flight, navy, army, commercial aerospace, rocket, aerospace defense, missile
Medical Devices	0.012	surgery, surgeons, surgical, physicians, therapy, clearance, heart, implant, pain, reps, tissue, medical devices, spine, reimbursement, dental
Insurance	0.011	underwriting, agents, combined ratio, investment income, reinsurance, loss ratio, written premium, expenses ratio, catastrophe, net investment, accident, annuity, workers, investment portfolio, casualty
Household Goods	0.009	pet, consumption, beverage, ingredients, foodservice, household, snacks, coffee, flavor, fresh, wellness, grocery, consumer demand, beauty, skus
Asset Management	0.009	assets management, wealth, aum, institutional, inflows, advisory, advisers, fixed income, capital market, outflows, management fees, index, wealth management, brokerage, assets class
Biotech	0.009	diagnostic, vaccine, lab, instruments, sample, consumables, life science, animal, pharma, assay, sequencing, imaging, panel, laboratory, variant
Health Care	0.008	health care, medicare, physicians, payers, pharmacy, clinic, reimbursement, net revenue, nursing, health system, care provider, prescription, admissions, clinicians, home health
Entertainment	0.008	sports, entertainment, streaming, audience, advertisers, studio, fans, film, music, title, programming, venues, advertising revenue, broadcast, television
Education	0.008	students, school, education, enrollment, university, campus, college, career, district, learners, academic, cohort, campuses, k12, educational
Catering	0.008	restaurants, guests, franchisees, menu, samestore, dining, samestore sales, team members, restaurants sales, wage, systemwide, offpremise, guests experience, companyowned, shack
Energy	0.007	rigs, esg, vessels, offshore, drilling, marine, gulf, lng, pipe, rigs count, activity level, pumping, rental, sea
Steel	0.007	steel, metal, mill, mining, aluminum, per tons, coal, carbon, scrap, metrics tons, valueadded, selling price, imports, export, iron
Hospitality	0.006	hotel, occupancy, resort, revpar, hospitality, casino, vegas, leisure, rent, guests, venues, terms loan, residents, club, vacation
Telecommunication	0.006	subscribers, fiber, broadband, churn, arpu, wireless, cable, spectrum, satellite, tmobile, verizon, subscribers growth, postpaid, voice, rural
Freight	0.005	aircraft, airlines, freight, miles, truck, flight, delta, gallon, passenger, aviation, fuel price, pilot, airport, max, terminal
Metals and Mining	0.004	mining, gold, copper, ounces, metal, coal, silver, pound, creek, grade, carbon, ore, mill, minerals, print

The table presents the topic names, the average loadings of firms' conference call texts on these topics, and the key terms in each topic based on the full-sample LDA estimation. We obtain stable topics following the algorithm described in Appendix B.

Table A3: Robustness Check: Out-of-Sample Sharpe Ratios From 2013

Cross-Sectional Information Sets:	$p = 5$	$p = 10$	$p = 15$	$p = 20$	$p = 25$	$p = 30$
	Panel A. Full-Sample LDA					
32 Q&A topics	0.67	0.89	0.36	0.85	1.31	1.17
32 PRE topics	1.36	1.33	1.29	1.22	1.33	1.35
64 Q&A + PRE topics	1.08	0.91	1.33	1.79	1.71	1.79
35 Firm Characteristics	1.09	1.32	1.74	1.79	1.89	1.91
64 Q&A + PRE topics + 35 Firm Chars	0.42	1.60	2.46	2.46	2.83	2.95
	Panel B. Online LDA					
32 Q&A topics	0.55	0.23	0.84	0.98	1.21	1.20
32 PRE topics	1.00	1.16	1.19	1.32	1.34	1.37
64 Q&A + PRE topics	0.85	0.71	1.37	1.69	1.83	1.85
64 Q&A + PRE topics + 35 Firm Chars	0.37	1.41	2.27	2.23	2.38	2.60

The table reports the annualized out-of-sample Sharpe ratios of the MVE portfolios implied by the IPCA models using as instruments: (i) 32 Q&A topic loadings, (ii) 32 presentation (PRE) topic loadings, (iii) 64 Q&A and presentation topic loadings, (iv) 35 firm characteristics, and (v) both the 64 topic loadings and the 35 firm characteristics. The Q&A and presentation topics are obtained from the full-sample LDA estimation in Panel A and from the online LDA estimation in Panel B. The number of latent factors in the IPCA estimation, denoted by p, is set to 5, 10, 15, 20, 25, and 30. The out-of-sample period is from January 2013 to December 2022.