WORKING PAPER · NO. 2024-155

Learning Fundamentals from Text

Alex G. Kim, Maximilian Muhn, Valeri V. Nikolaev, and Yijing Zhang

DECEMBER 2024

THE UNIVERSITY OF CHICAGO
BECKER FRIEDMAN INSTITUTE FOR ECONOMICS

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

Learning Fundamentals from Text

Alex G. Kim[1] Maximilian Muhn[2]
Valeri V. Nikolaev[3] Yijing Zhang[4]

November 2024

Abstract

We introduce a novel approach to learning the information that investors react to when processing textual information. We use the attention mechanism that learns to identify content that triggers market reactions to disclosed information. The explanatory power of the attention-based model significantly exceeds that of attention-free models. We then develop and analyze a comprehensive set of topics discussed in companies' annual reports. Segment information, goodwill and intangibles, revenues, and operating income are the topics that receive the most attention from investors. Despite their prominence in the public discourse, sustainability and governance are consistently among the least important topics judging by the market reactions. Building on our approach, we show that regulatory interventions can successfully enhance the relevance of textual communication. We also show that firms strategically position information within MD&A to influence investor focus. Our findings underscore the value of attention-based analysis of corporate communications and open new avenues for future work.

Keywords: Attention mechanism, information relevance, fundamentals, LLM, large language model, information processing, corporate disclosure, investor attention.

JEL Codes: C45, C55, G12, G17

[1]The University of Chicago, Booth School of Business, alex.kim@chicagobooth.edu
[2]The University of Chicago, Booth School of Business, maximilian.muhn@chicagobooth.edu
[3]The University of Chicago, Booth School of Business, valeri.nikolaev@chicagobooth.edu
[4]The University of Chicago, Booth School of Business, yijing.zhang@chicagobooth.edu

We appreciate financial support from the University of Chicago Research Support Center and the Stevens Doctoral Program at the University of Chicago Booth School of Business.

I. Introduction

In this study, we introduce and apply a novel approach that identifies the specific language that stock market participants focus on when processing textual information. Since Roll (1988), it is well-understood that firm-specific information explains the bulk of variance in firm stock returns. However, what types of information are, more specifically, moving prices? The key informational events, such as press releases or earnings announcements, present large bundles of textual communications, making it difficult to disentangle the specific information that drives market reactions. While the effect of quantitative fundamentals, such as revenue or profit margins, can be studied directly, qualitative fundamentals, such as changes to the business model, strategic initiatives, competition, regulatory environment, and many others, are much harder to identify. This challenge leaves several key questions unanswered. Are there systematic patterns in the unlimited variety of firm-specific textual data to which markets respond? What are the most important types of information or topics that influence prices? Do companies position this content strategically when communicating information to investors? These and a number of related questions remain to be understood.

The challenge of identifying the most relevant content and learning fundamentals from corporate communications is that of dimensionality. A typical annual report (10-K) disclosed by a company contains, on average, four hundred paragraphs. Large language models efficiently encode a paragraph into a vector of features (i.e., embeddings) whose dimensions vary from hundreds to thousands. These paragraph embeddings easily add up to more than 100,000 characteristics per document and are known to interact in complex non-linear ways (e.g., Kim and Nikolaev, 2023). Estimating a model on a relatively limited sample of financial data, not to mention linking the characteristics to sentences or paragraphs, becomes infeasible.

This paper tackles this problem by introducing the "attention mechanism" from the Transformer architecture applied over paragraphs of a company's annual report instead of words in a sentence (Vaswani et al., 2017). Just like a language model (e.g., GPT), which is trained to learn what words to focus attention on in a lengthy sequence (context) when predicting the next token, we train the model to focus attention on the relevant paragraphs when explaining (or predicting) stock market reactions. This mechanism involves learning the contextual relationships among paragraphs within an annual report and identifying paragraph-level contributions from the markets' perspective.

Several features make the attention-based approach appealing for analyzing complex disclosures. First, the standard approach to processing long textual documents averages features across paragraphs (e.g., logistic regression or LightGBM that cannot accommodate a sequence of vectors as input), making it impossible to distinguish the nuanced importance among different paragraphs, not to mention utilizing the interrelationship among them. Second, investors do not interpret each paragraph in isolation but place it within its broader context. The attention mechanism mirrors this process and contextualizes each paragraph's meaning depending on other paragraphs. Consequently, a paragraph discussing risk factors can receive more attention when the context suggests increased aggregate uncertainty. Third, by analyzing the trained model's outcome, one can trace back how the model prioritizes information, which can directly translate into insights about what content investors find important.

Before discussing our analysis, we highlight two additional distinctions from the prior literature. First, our approach does not involve reducing the information in a lengthy textual document to unidimensional constructs, such as sentiment (e.g., Tetlock et al., 2008; Loughran and McDonald, 2011) or uncertainty (e.g., Campbell et al., 2014; Hassan et al., 2019). For example, a discussion of strategic changes, the competitive landscape, new products, and investment projects, among many other important factors, may be relatively neutral while contributing significantly to stock price movements. Second, our attention-based approach can be applied to different outcome variables, e.g., concurrent or future returns, earnings (or even environmental performance), and effectively allows filtering *relevant* textual information, e.g., about a company's fundamentals, from the narrative. In this sense, our paper complements recent machine learning techniques that filter new information from a large pool of textual (Costello et al., 2024) or quantitative data (Kelly and Pruitt, 2013, 2015).[1]

Our analysis focuses on the entire universe of electronic 10-K filings from the EDGAR database. After deleting non-textual representations such as figures and tables, we split the documents into paragraphs. Using more than 20 million paragraphs, we convert each paragraph into a textual vector using OpenAI's `text-embedding-3-large` model. OpenAI embeddings include 3,072 elements ordered by importance. We condense the dimensionality from 3,072 to 64 for computational efficiency.[2] Thus, the input into our model is an annual

[1] Costello et al. (2024) focus on identifying new information in complex textual filings and study their ability to explain stock returns. The study uses a different methodology and has a different focus. Conceptually, one important distinction is that our approach identifies information relevant to a decision, which need not be new or surprising. Indeed, much of the fundamental information is not "news" to investors (but predicts future returns or earnings), including well-known information size, value, and profitability effects, among others, but also extends to factors such as competition, product market, strategy, regulatory exposure, etc.

[2] OpenAI embeddings are trained using a technique called Matryoshka Representation Learning, which allows the shortening of the dimensionality without losing the embedding's contextual properties (Kusupati et al., 2022).

report represented by a sequence of paragraph-level textual vectors (i.e., it is a $n \times 64$ matrix, where n is the number of paragraphs in a given document).

We include two attention layers that process the input data. The first layer is a traditional Transformer-style attention layer (see Section II.B for technical details) that revises the semantic meaning of each paragraph depending on the surrounding paragraphs (context).[3] The second layer is a standard dot-product attention mechanism trained to aggregate (weight) paragraph vectors into a single vector representing the entire document. This vector is then used in a feed-forward artificial neural network to predict the target variable.[4] Figure I illustrates the overall model architecture and Section II.C contains the training details.

We train our primary model using the directional changes in stock returns realized *around* the 10-K filing dates as the target variable. Specifically, our baseline target variable is an indicator equal to one when the cumulative stock return from one day before the 10-K filing date to 30 days after the 10-K filing date is positive and zero otherwise. This structure enables the model to learn the relevance of the content included in an annual report from the perspective of the stock market.

We evaluate the model's performance by comparing its out-of-sample predictive accuracy with several machine learning benchmarks. Specifically, we separately train a feed-forward multi-layer perceptron (MLP) model without the attention layers. This model treats all paragraphs equally by using their average embedding as input. Additionally, we also estimate a logistic regression that uses the average embedding of the paragraph within a document as the independent variable. We find that the attention-based model achieves the highest predictive performance. In terms of economic significance, the AUC of the attention-based model is 2.71% points higher than that of the MLP model and 4.02% points higher than that of the logit model. The differences are statistically and economically significant.[5]

We corroborate our findings by training the attention model on two additional target variables: future stock returns and directional changes in future earnings-per-share. In both tasks, attention-based models significantly outperform MLP and logit models. Unlike returns around filing dates, these two target variables represent pure prediction tasks. We find that the trading portfolios based on these predictions yield substantially higher Sharpe ratios than those based on the MLP and logit models. Comfortingly, we note that the

[3]Similarly, in the GPT architecture, the attention layer revises the meaning of each word in a sequence depending on the surrounding words. Therefore, when a document-level matrix is processed with the initial attention layer, we obtain an attention-adjusted matrix with a dimension of $n \times 64$.

[4]The ANN model contains two hidden layers and a non-linear activation function, ReLU. The final output layer applies a softmax probability, providing the probability that the stock returns will be positive.

[5]An AUC of 55% is exceptional for return prediction tasks (Chen et al., 2022).

model shows no deterioration in its predictive ability following the period that lies outside OpenAI's embedding model's training window.[6] Overall, our findings suggest that training and applying the attention mechanism to content in corporate communications better helps to uncover fundamental information driving the contemporaneous and future stock market reactions as well as future earnings.

Having established the value of attention, we measure the paragraph-level importance score for each paragraph. To do so, we measure the cosine similarity between the vector representation of each paragraph and the attention-weighted vector representations of the entire annual report. Intuitively, paragraphs that are given a higher amount of attention by the model when explaining stock reactions exhibit a higher importance score. That is, a paragraph is more important if its semantic information aligns more closely with the overall document's attention-based representation.

We begin by examining the importance of different sections of 10-K (i.e., Items) from an investor standpoint. Because 10-K items differ considerably in length, we retain the five paragraphs of highest importance for each item and average them.[7] Our results show that the Management Discussion and Analysis (Item 7) are of the highest importance, followed by Financial Statements and Notes (Item 8), Business Description (Item 1), and Risk Factors (Item 1A). These results are intuitive and line up with the prior research's focus on MD&As (e.g. Kim et al., 2024b), business descriptions (e.g., Hoberg and Phillips, 2016), or risk factor disclosures (e.g., Hail et al., 2021). At the same time, disclosures related to Directors and Governance (Item 10) or Ownership (Item 12) score as least important. Overall, these findings are intuitive and suggest that our attention model is capable of identifying the relevant content.

We then investigate which topics within an annual report drive market reactions. Instead of relying on the previous generation of topic classification models (e.g., LDA), we develop a more comprehensive and adaptable approach that relies on an LLM and representative sampling to generate an extensive and context-sensitive taxonomy of topics and subtopics. Specifically, we use GPT to derive a comprehensive topic list for the most influential 10-K sections, then train a scalable classifier that can readily be applied to the full universe of documents.[8] Using this classification, we observe that the most important topics in the

[6]Predictions starting in the fiscal year 2021 lie outside the `text-embedding-3-large` model training window.

[7]Taking the average of all paragraphs is subject to the number of paragraphs included in each item. As the length of each item varies substantially, the resulting importance measure could be heavily influenced by its length. However, we confirm that the order of the items remains qualitatively similar even when we use the average of all the paragraphs.

[8]For this task, we focus on the four key items in annual reports (Items 1, 1a, 7, and 8). More details on

MD&A section, are segment information, financial performance, and liquidity and capital resources. The top-ranking topics in other sections are financial position (Item 8), financial performance (Item 1), and financial risks (Item 1a). The top-scoring subtopics are segment profitability (Item 7), goodwill and intangibles (Item 8), segment performance (Item 1), and liquidity (item 1a). In contrast, the least important three topics within MD&A are corporate governance, accounting pronouncements, and forward-looking statements. The three lowest-ranking subtopics for MD&A include social responsibility, ethical standards, and assumptions. For the other sections, the lowest-ranking topics are auditor report (Item 8), sustainability and CSR (Item 1), and reputational risks (Item 1a). The corresponding subtopics are audit standards (Item 8), social responsibility (Item 1), and Cyber threats (item 1a).

The evidence above overwhelmingly indicates the primacy of financial performance and revenue-related topics. Interestingly, governance and corporate social responsibility-related topics, which have been the focus of public attention in recent years, are ranked consistently towards the bottom of the topic distribution, judging by stock market responses. Instead, we observe that debt financing and liquidity-related topics score high in terms of investors' attention. Taken together, we find that topic indicators account for almost 60% of the explained variance in the importance of paragraph levels. This suggests that while firm-specific information is idiosyncratic, the factors driving stock price variations exhibit significant commonalities across firms and over time.

We then address two economic questions related to the importance of content in corporate filings. First, we examine whether regulatory interventions can enhance the usefulness of the regulated information, i.e., whether investors start paying more attention to disclosed content. While regulatory burden is often viewed as a source of the increased complexity of companies' communications (e.g. Guay et al., 2016; Dyer et al., 2017), regulation can also discipline and facilitate the revelation of information. To answer this question, we exploit the SEC's Modernization of Regulation S-K as a regulatory shock. The modernization, which is one of the most significant changes regulating textual communication, aimed to enhance the MD&A disclosures to help investors understand the underlying reasons behind the company's financial condition and operational results. Using Financial Statements and Notes (Item 8) as a control group unaffected by the new regulation, we show that the relative importance of MD&As (Item 7) compared to the unaffected textual disclosures (Item 8) increases immediately following the regulatory change. This finding is consistent with an interaction between regulatory requirements and transparent communication of qualitative

the topic classification approach are provided in Section II.E.

fundamentals.

To complement this analysis, we also examine whether the SEC's attention to corporate filings, manifested through comment letters requesting improvements in textual disclosures (Ryans, 2021), enhances the relevance of the disclosed information. We match each comment letter to its corresponding items in the 10-K filings and show that the specific items identified by the SEC exhibit an increase in importance immediately following the comment letter. Overall, our evidence suggests that regulatory interventions shape the flow of relevant information into stock prices.

The second question we explore is whether companies strategically position important information within their disclosures in an attempt to capitalize on investors' limited attention. For this analysis, we focus on MD&A disclosures within which management has considerable discretion (e.g., Cohen et al., 2020). We find that informationally relevant content (paragraphs) is generally placed upfront. However, when companies have to deliver negative news, are subject to more intense competition, report lower profitability, or exhibit higher earnings volatility, they often position important information later in the document. These findings suggest that companies strategically influence the order of communicated information (Bloomfield, 2008).

Our paper contributes to the literature in several ways. First, we address the problem of discovering fundamental information from complex textual disclosures. We introduce a novel approach that allows learning the importance of specific language that matters for investors and their decisions. Our approach learns to dedicate attention to relevant paragraphs and can be applied to a wide range of outcomes (e.g., compensation). This adds to the growing literature that applies machine learning to study information explaining stock prices (e.g., Gabaix et al., 2023; Jiang et al., 2023; Chen et al., 2022; Cao et al., 2024; Costello et al., 2024)

Second, while it is understood that firm-specific information drives most of the variance in stock returns (Roll, 1988), it is less understood what specific content matters and whether it can be systematized. Our study provides direct evidence on this front. We document rich heterogeneity in the importance of individual paragraphs within 10-Ks in explaining market reactions. We document a set of the most important topics, e.g., goodwill and intangibles, segment performance, investment securities, and liquidity, that investors focus attention on when processing corporate disclosures. In contrast, content related to corporate governance and sustainability that dominates public discourse is less important based on investors' reactions. This systematic identification of content that the market reacts to adds new insights into what textual information matters more for investors (e.g., Tetlock et al., 2008;

Cohen et al., 2020).

Third, we show that regulatory requirements directly influence the importance of disclosed content from an investor's standpoint. The SEC's efforts to overhaul Regulation S-K to increase the informativeness of management discussions in annual reports improved investors' attention to this content. A key difference in our study is that we are able to identify which specific content within a report is impacted by the regulatory change. Finally, we provide evidence that while companies tend to front-load the most important news, they also strategically manipulate the positioning of less favorable content within a document. This adds to the literature on strategic disclosures (e.g., Li, 2008; Lo et al., 2017; DeHaan et al., 2021; Kim et al., 2024b).

II. Data and Methodology

In this section, we explain data sources and how we pre-process the data (Section II.A). We then discuss model architecture (Section II.B) and training (II.C).

II.A Data

Our sample of textual disclosures comprises the universe of electronic 10-K filings available on SEC's EDGAR database for the period 1996-2023. For each 10-K filing, we delete non-textual disclosures such as tables and figures and remove HTML syntax and XBRL tags. We then split the documents into paragraphs based on the paragraph separators of the original disclosures.[9] We drop excessively long documents (exceeding 1,000 paragraphs). Finally, we require a valid link between the SEC filings, CRSP, and Compustat. We use CRSP for daily stock returns and Compustat for firm-level financial data, such as earnings. This process results in a sample of 76,929 filings and 20,712,462 paragraphs for our main sample (2004-2024).[10]

II.B The Attention Mechanism

The attention mechanism, which is the key part of the Transformer architecture, is the backbone of state-of-the-art large language models. It enables a model to learn to weigh the importance of different elements in the input sequence and, hence, to focus attention on the most relevant content when generating outputs (Vaswani et al., 2017). This architecture is particularly successful as it allows the computationally efficient processing of long

[9]We concatenate "list items" into a paragraph, preserving both the numbering and line breaks. Our results are qualitatively robust when we treat different list items as separate paragraphs.

[10]This excludes the data used in the initial training performed from 1996 to 2003.

sequences. It enables the training of large-scale models with billions of parameters, leading to high-performing large language models such as Bidirectional Encoder Representations from Transformers (BERT) and Generative Pre-trained Transformers (GPT).

The Transformer computes three vectors for each element in the input sequence: query, key, and value. The value encodes the semantic meaning of a given word (token), which, at each step, is modified by the value of the adjacent context. In particular, for a given token, the model uses its query vector to search for related tokens in the sequence (context) based on the matching tokens' keys. A close match causes the model to revise the value associated with our given token by combining it with the value of the matched token (depending on the strength of the match, i.e., attention). The attention scores measure the (dot product) similarity between the query and key. These scores assign the contextual importance of each token in relation to other tokens in the input sequence.

The transformer architecture is most commonly applied to model relationships among words in a text sequence. However, it can also be applied to other sequences. For example, Gabaix et al. (2023) use transformer architecture to model an ordered sequence of stocks in an investor's portfolio. In this study, we use transformer architecture to model attention over paragraphs.

For example, consider a document with ten paragraphs, and let's assume a researcher focuses on information conveyed by the first paragraph. In this case, the researcher is concerned with the "value" vector that encodes information from the first paragraph. However, this value vector is not informative if the researcher does not consider the context provided by the other nine paragraphs. To tackle this, the researcher will read the entire document, identify the most relevant context (related paragraphs), and revise her knowledge, i.e., the value associated with the first paragraph, accordingly. This is, indeed, what the attention layer aims to do. It first compares the query vector of the first paragraph with the keys of all other paragraphs (including itself) by computing their similarity scores (attention score). Subsequently, the attention layer revises the value vector associated with the first paragraph by incorporating a weighted average value of the most relevant paragraphs. Put differently, the attention mechanism allows us to enrich the meaning of each paragraph in a context-dependent way. Similarly, it allows the model to identify the most relevant paragraphs that help to predict the target variable.

Attention Scores These concepts can be represented mathematically as follows. Assume matrix $X \in \mathbb{R}_{n\times d}$ represents an input text, which is a sequence of n tokens (or paragraphs), each of which is represented by an embedding vector. To capture interdependencies among tokens, we project X into the query, key, and value spaces by computing $Q = XW_Q$, $K = XW_K$, and $V = XW_V$, where W_Q, W_K, and W_V are learnable parameter matrices. The attention score matrix is defined as:

$$\text{Attention Scores} = \frac{QK^T}{\sqrt{d_k}} \in \mathbb{R}_{n\times n}, \tag{1}$$

where d is the dimension of the embedding vector that serves as normalization. The attention score of the i-th token in relation to the j-th, i.e., Attention Scores_{ij}, represents a dot product (analogous to correlation) of ith element query vector with the j-th element key vector. Finally, the output of the attention matrix is the product of Attention Scores and the value matrix.

$$X^{(1)} = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{2}$$

where softmax$(\cdot)$ is a multinomial logit function that converts raw attention scores into weights row-wise, such that they sum to one (highlighting the most relevant positions). Thus, the output from the attention layer is a sequence of now attention-weighted token embeddings.

II.C Model Architecture and Training Details

Our approach applies the above attention layer to paragraph embeddings instead of token embeddings. Put differently, each paragraph of 10-K filings is a "token" within the sequence of paragraphs. The model is then trained to understand the relational importance of the paragraphs within each document and learn to encode their contextualized meaning in vector form. As we train our baseline model on stock returns around the dates information is released to the market (announcement returns), the attention mechanism learns to identify paragraphs (while accounting for the broader context) that are relevant for explaining the stock price movements.[11]

Our model architecture includes three key steps: (i) the first Transformer-style attention layer that contextualizes each paragraph depending on the other paragraphs, (ii) the second customized attention layer, which learns to optimally aggregate over a large number of paragraphs into a document-level embedding, and (iii) the multi-layer perceptron (feedforward neural network) that uses the attention-weighted document embedding to explain the

[11]This further differentiates our approach from traditional large language models, typically trained using unsupervised objectives (e.g., next token prediction) to learn general language patterns. In contrast, we train our model in a supervised manner with a specific financial target (such as earnings or returns), allowing the attention mechanism to focus on content relevant to predicting this outcome.

announcement returns. In this section, we provide more details about each of these steps and on the model's training. Figure I visually displays the overall model architecture.

Embeddings and First Attention Layer The first attention layer follows the structure presented in Section II.B. For each paragraph p_i in a 10-K filing, we retrieve its vector representation (embedding) $e_i = \text{EmbeddingModel}(p_i)$ using OpenAI's `text-embedding-3-large` model. e_i is a dense vector with an original dimension of 3,072. A special feature of this embedding model is that it is trained to order the dimensions by their importance. Hence, to reduce dimensionality, a common approach is just to keep the top k elements of the vector (Kusupati et al., 2022). We use the first 64 elements of each vector. Therefore, a document D with n paragraphs is then represented as a stacked matrix of paragraph embeddings, $X = [e_1; e_2; \cdots; e_n]_{(n\times 64)}$. In addition to the semantic embeddings, we add a positional encoding vector for each paragraph to preserve information about the order of the paragraphs within the document.[12,13] Using equation (2), we obtain the attention matrix $\text{Attention}(X) \in \mathbb{R}_{n\times 64}$.

Second Attention Layer We then design a customized attention layer that takes a sequence of paragraph embedding outputs by the first attention layer as input and learns how to aggregate them optimally to generate a document-level embedding vector. In particular:

$$X^* = \text{softmax}\left(\frac{W^T \text{Attention}(X)^T}{\sqrt{64}}\right)\text{Attention}(X) \tag{3}$$

where $W \in \mathbb{R}^{64}$ is a learnable weight vector. The $\text{softmax}(.)$ operation produces an attention vector whose length equals the number of paragraphs in the document and whose elements sum up to one. The resulting $X^* \in \mathbb{R}^{64}$ is a document-level representation that encapsulates the relational semantic importance in all paragraphs.

[12] The positional encoding vector for each paragraph has the same dimension as e_i, and we combine these two vectors element-wise before constructing the X matrix. pos is the position of the paragraph, and i is the dimension index. We scale the positional encoding by the norm of the input paragraph embeddings.

$$\text{Positional Encoding}(\text{pos}, 2i) = \sin\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right),$$
$$\text{Positional Encoding}(\text{pos}, 2i+1) = \cos\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right)$$

[13] We follow the conventional practice and add a "residual connection" to the first attention layer. The residual connection takes the output from the attention layer, $X^{(1)}$ and adds the original embedding matrix X to this output.

Multi-Layer Perceptron The last step in our model architecture is a feedforward neural network model using X^* (document-level 64-dimensional vector representation X^*) as input. The network has two hidden layers with 128 and 64 neurons each and a one-dimensional output layer. The model's target variable is a binary indicator that equals one when the stock returns surrounding the filing dates are positive and zero otherwise. We use cumulative returns measured from one day before the filing date to 30 days after.[14] We use a rectified linear unit (ReLU) activation in the first two layers (the initial layer and the first hidden layer) to impose non-linearity among the neurons. The model's output represents the probability that filing returns are positive. Note that our model does not incorporate any numeric firm information but relies only on text extracted from annual filings.

Training Details The training, validation, and test samples are constructed on a rolling basis over time. The training sample consists of the six years preceding the validation sample, and the validation sample consists of the two years preceding the test sample (one year). As our initial data is from 1996, the first test sample year is 2004 (i.e., the train set is from 1996 to 2001, and the validation set is from 2002 to 2003). We then roll the training forward each year and, thus, have 20 different test sets up to 2023.

Our training consists of two phases. First, we use a grid search to find an optimal set of hyperparameters based on the validation sample.[15] We use Mean Squared Error (MSE) as our loss function and the Adam optimizer for model training.[16] Second, we select the specification with the optimal set of hyperparameters and train the model over the most recent six years (Gu et al., 2020). This two-stage approach allows us to both optimize our model parameters and incorporate the most current data, enhancing the model's predictive capabilities.[17].

[14]In other model variants, we use future stock returns or future earnings changes as target variables.

[15]We experiment with five different learning rates: 0.00001, 0.0001, 0.001, 0.01, and 0.1, and four different batch sizes: 256, 500, 800, and 1,000. Therefore, we iterate the process 20 times each to find the optimal set of hyperparameters. We adopt an early stopping criteria and stop the training when evaluation metrics do not improve over ten consecutive epochs.

[16]While Binary Cross-Entropy (BCE) Loss is commonly used for classification tasks, we opt for MSE because our experimentations reveal that MSE consistently outperforms BCE in various specifications. We attribute the superior performance of MSE to its more stable gradient behavior. Specifically, MSE yields smoother and more symmetric gradients, thus enabling easier optimization and smoother updates during the training process. In contrast, BCE produces sharper and more volatile gradients when the prediction is incorrect. Although, in theory, it could lead to faster convergences, it introduces instability and volatility, which is the case here. Therefore, MSEs proved more effective for our model's training.

[17]To take advantage of the most recent data, we do not use the validation sample but instead stop training based on the same optimal number of epochs identified during the hyperparameter selection

II.D Paragraph-Level Importance

The key measure used in our analysis is paragraph-level importance. Our measure is based on the attention-adjusted vector representation of the entire document X^* and its semantic similarity with paragraph-level embeddings.

$$\text{Paragraph Importance}_{ikt} = \frac{X^*_{it} \cdot e_{ikt}}{\parallel X^*_{it} \parallel \cdot \parallel e_{ikt} \parallel} \tag{4}$$

where X^*_{it} is the attention-adjusted vector representation of firm i's time t filing, e_{ikt} is the embedding vector of paragraph k in firm i's time t filing. $\parallel \cdot \parallel$ is the L2-norm. Recall that X^* has learned the relative importance of each paragraph in the document and, therefore, has imposed higher (lower) weights on more (less) important content. Following this notion, we compute the cosine similarity between each paragraph embedding e_i and X^*. The more similar the two vectors are, the more important the paragraph is. In some of our analyses, we further normalize this measure of importance within each document.

II.E Topic Classification Methodology

In order to determine what topics investors focus attention on when processing information, we require a comprehensive list of topics encountered in annual reports. We develop a topic classification methodology consisting of three steps: First, we use a large language model (GPT-4o) to obtain exhaustive topic and subtopic lists for each of the four most important 10-K items (Items 1, 1a, 7, and 8). Second, for each item, we create a labeled training set by instructing GPT to classify paragraphs into topics and then subtopics for a sample of documents. Third, we train a scalable multi-label logit classifier on the labeled sample and extend it to the full sample. We choose this approach as, based on our testing, it is more flexible, robust and comprehensive relative to the previous generation of topic classification approaches (e.g., LDA topic modeling).

More specific details of the three steps are as follows. As a first step, we stratify firms based on one-digit SIC codes and four five-year time windows (spanning our sample) into 40 strata (10 industries $\times$ 4 periods). For each of the resulting time-industry strata, we randomly sample five documents and feed these five documents into GPT-4o. The model is then instructed to generate a list of topics that are encountered within the five input documents at the paragraph level. Within each topic, the model then determines subtopics. Thus, the initial sample of 200 documents (40 strata $\times$ five documents) serves as our basis for generating topic lists.

We then consolidate and reconcile topic and subtopic lists across the strata. In particular, GPT first aggregates topic-subtopic lists within each industry. Next, it aggregates the lists across industries to obtain a comprehensive topic-subtopic list for each 10-K item. GPT is instructed to avoid industry-specific topics but rather develop industry independent topic lists. We note that our topic list is robust to the choice of an initial 200 documents and is representative of the entire universe of documents.

As a second step, we create a labeled training dataset using a sample of 2,200 documents (our initial 200 documents plus an additional stratified random sample of 2,000 documents). For each paragraph in these documents, we instruct GPT to classify it into one of the predefined topics and corresponding subtopics (with the option to assign it to "Others" if no existing subtopic fits). Note that a document contains, on average, 400 paragraphs, and hence, we do not process the entire population of paragraphs in this way.

As the third step, we train a multi-label logit classifier on our labeled dataset using 64-dimensional OpenAI paragraph embeddings as input features. Using a random 20% of the data set as a test sample, our classifier achieves an accuracy of 80% in topic classification and 72% in subtopic classification on average.[18] We then deploy this classifier to generate topic-subtopic classifications for the entire sample of paragraphs in each item.

The resulting list of topics for each of the four most important items in an annual report can be found in Appendix E. It encompasses topics such as "financial performance," "corporate governance," "liquidity and capital resources," "forward-looking statements," and many more.

III. Model Performance

In this section, we present the performance of our attention-based models in comparison to other machine learning models. As our primary economic focus is to identify information that investors focus on, our main target variable is contemporaneous stock returns around the filing dates. Section III.A displays the informational gain obtained using the attention layers to predict filing date returns. However, we also test two other target variables (future stock returns and directional changes in future earnings-per-share) later in Section III.B.

To benchmark our main model with attention layers, we train two comparison models: a feedforward neural network (MLP), and a standard logit model. Both of these models use the averaged paragraph-level embeddings as their input. Such a simple average of all paragraph-level embeddings also results in a 64-dimensional vector; however, this approach

[18] We also experiment with an artificial neural network architecture but observe no significant improvements in performance.

is attention-free and thus assumes that all paragraphs have equal importance. For the first benchmark model, the training and modeling details of the neural network are identical to those used in the multi-layer perceptron of our main architecture. Furthermore, for the second benchmark model, we estimate a simple logit model using the same data from the past six years and use the coefficient estimates to test the model.

III.A Using Returns Around Filing Dates

Table I presents the performance of our main model. Our main evaluation metric is the Area Under the ROC Curve or AUC (Chen et al., 2022). We focus on this metric over accuracy, the proportion of correct classifications out of all predictions, because the latter can be sensitive to the threshold values. Furthermore, AUC considers both the true positive rate (sensitivity) and the false positive rate, providing a more accurate measure of the model's ability to distinguish between the two classes. Nonetheless, using 0.5 as a consistent prediction threshold, we report accuracy metrics along with the AUC values.

The table indicates that our attention-based model significantly outperforms the MLP and logit models. The average AUC of the attention-based model is 0.5599, while the AUC of the MLP model is 0.5328, and the one of the logit model is 0.5197. An AUC of 0.5599 is impressive, especially in return prediction tasks (see Chen et al., 2022, for example). Given 0.50 as the baseline, our attention-based model outperforms the baseline by 6% points, whereas the MLP model adds 3.3% points, and the logit adds 2% points. The information gained using the attention-based model is almost three times larger than the logit model and twice as large as the MLP model. Importantly, this outperformance occurs consistently for each prediction year.

Notably, the difference in performance between the attention-based model and the MLP model is the informational gain of adopting the attention mechanism. The difference between the MLP and logit models proxies for the incremental value of introducing non-linearities and interactions among the features. Given these interpretations, the informational value of non-linearity is 0.0131, and the informational value of understanding relational importance is 0.0271. As a formal statistical test, the bottom of the table reports a pairwise t-test (performed across all years, i.e., 19 observations) that indicates that the attention model's AUC is higher than the other two model's AUC at 1% statistical significance level. The results are generally consistent using accuracy as a performance metric.

III.B Other Targets

In addition to explaining announcement returns around filing dates, we demonstrate the value of adopting an attention mechanism further by predicting two additional targets: future returns and changes in earnings. For future returns, we use an indicator that equals one when cumulative returns from one day after the filing date to 30 days after the filing date are positive and zero otherwise. For changes in earnings, we use an indicator that equals one when the earnings-per-share of the following year exceeds the current year's earnings-per-share and zero otherwise.

Table II, Panels A and B, display the model performance for these two target variables. In predicting future returns, the attention-based model achieves an AUC of 0.5600, which adds 0.0267 over the MLP model's AUC (t-value = 3.32) and 0.0406 to the logit's AUC (t-value = 6.42). Similarly, in predicting the direction of future earnings, the attention-based model achieves an AUC of 0.5481, which exceeds the MLP model's AUC by 0.0250 (t-value = 4.08) and exceeds the logit model's AUC by 0.0450 (t-value = 9.34). Overall, our model shows a robust outperformance over other machine learning models in predicting future financial variables. This result highlights the importance of considering the relational importance of information within each document.

Sharpe Ratios As the two additional target variables are based on future financial variables, we construct portfolios using the models' predicted outcomes and evaluate their performance using Sharpe ratios. For future returns, we confine our sample to observations with filing dates in February or March. This restriction reduces our sample size to approximately 70% of the full sample. Since we are predicting one-month-ahead returns, we construct portfolios four times during the two-month period. Specifically, for companies with filing dates between February 1st and 14th, we form a portfolio on February 15th and hold it for two weeks. For filing dates between February 15th and 28th, we form a portfolio on March 1st. We apply the same procedure to the filing dates in March, forming portfolios on March 15th and April 1st, respectively, and holding them for two weeks each. The portfolio is a simple long-short approach based on the model's predictions. We sort the observations on the predicted probability of positive returns and take a long (short) position for the stocks in the top (bottom) quintile. The portfolios are equal-weighted.[19]

For future earnings, we form portfolios on June 30th each year and hold them for one year. The long-short portfolio is constructed on the basis of the predicted values from the model. The higher the predicted probability, the more likely the firm will experience an

[19] In untabulated tests, we also consider value-weighted portfolios and find qualitatively consistent results.

increase in EPS in the next fiscal year. We use the predicted probability values from the three models (attention-based, MLP, and logit) and compare their Sharpe ratios.

We report the results in Table II, Panel C. Consistent with the predictive accuracy results, we find that the portfolios based on the attention-based predictions achieve the highest Sharpe ratios. Specifically, when predicting one-month-ahead stock returns, the attention-based model produces a Sharpe ratio of 1.56, outperforming the MLP model (Sharpe ratio = 1.28) and the logit model (Sharpe ratio = 1.08). When predicting directional changes in future EPS, the attention-based model achieves a Sharpe ratio of 1.40, again surpassing the MLP (Sharpe ratio = 1.13) and the logit (Sharpe ratio = 1.06) models.

Our results consistently underscore the value of incorporating the attention mechanism in prediction tasks that involve aggregating complex textual documents. More specifically, our results suggest that considering interdependencies among paragraphs and learning their attention-based weights can significantly improve the performance of financial prediction tasks.

Look-ahead Bias It is important to consider that relying on pre-trained language models for prediction tasks can be affected by look-ahead bias (Sarkar and Vafa, 2024). The models, which contain knowledge up to the training cutoff, might inadvertently access future information when making predictions. However, our results are unlikely to be influenced by look-ahead bias for various reasons. First, our primary model uses contemporaneous returns as the target variable and thus is geared towards explaining what happened (rather than predicting the future). Second, we do not ask generative AI to provide predictions and do not even directly use OpenAI embeddings. In contrast, our model uses these embeddings as inputs and is trained dynamically and tested purely out of sample.

Nevertheless, as a precautionary check, we exploit the fact that OpenAI's embedding model, trained on a large corpus of textual data, has a knowledge cutoff in September 2021. Thus, we are able to conduct pure out-of-sample tests (e.g., Kim et al., 2024b,a). The last three rows of Table II, in Panels A and B, fall outside of the embedding model's training sample (2021 to 2023). In this table, which focuses on predicting future outcomes, we observe that the dominance of the attention-based model remains consistent after the knowledge cutoff, with little to no deterioration in performance. This result also holds for our main specification in Table I. Thus, look-ahead bias is unlikely to be a major contributor to our findings.

IV. What Topics Matter Most?

In the previous sections, we demonstrate the value of the attention mechanism in identifying information from complex corporate documents that moves stock prices. In this section, we provide evidence on which content investors are focusing on when processing information. We start with item-level (Section IV.B) and, subsequently, analyze topic-level (Section IV.C) importance.

IV.A Descriptive Statistics

Table III, Panel A presents descriptive statistics on paragraph-level importance scores. Paragraph-level importance has a mean of 0.455 with a standard deviation of 0.122. Figure II shows that the distribution of importance scores is slightly left-skewed, but overall resembles the shape of a Normal distribution. The figure indicates substantial heterogeneity across paragraphs.

In Table III, Panel B, we perform a variance decomposition of paragraph-level importance. Firm-fixed effects explain only 8.41% of the variation in our measure, indicating that almost 90% of the variation occurs within firms. Similarly, year-fixed effects explain little of the variation in paragraph-level importance, with an R-squared value of 5.65%. Including firm $\times$ year fixed effects (i.e., document-level fixed effects) increases the explained variation to 18.23%. This result implies that nearly 80% of the variation in our measure exists within a document, indicating that paragraphs have varying levels of importance.

We further add more granular item-level fixed effects to examine whether different 10-K sections (Items) can explain the variation in importance scores. Item-fixed effects alone explain 8.49% of the variation in the importance scores, which is similar to the magnitude of the variation explained by firm-fixed effects. Lastly, we include firm $\times$ year $\times$ item-fixed effects to examine the variation within a specific item of a document. The addition of these fixed effects explains 31.51% of the variation, leaving nearly 70% of the variation within an item. These results jointly imply that paragraph-level importance exhibits substantial heterogeneity both across but especially within different sections of an annual report.

IV.B Ranking of 10-K Items

Our paragraph-level importance scores allow us to examine the relative significance of different sections, topics, and subtopics in the annual report. We begin by exploring the importance of different sections of 10-K, formally referred to as "Items." Because 10-K sections vary substantially in length, we retain the top five paragraphs sorted on the paragraph-level importance within each item and average their importance scores.[20] We drop items with less than five paragraphs.

Table IV presents the importance score computed for different 10-K items. In the first two columns, we report raw paragraph-level importance scores, while in the last two columns, we normalize the paragraph-level importance scores to sum to one within each filing. Regardless of the computation methods, we find that Item 7 (Management's Discussion and Analysis), Item 8 (Financial Statements and Notes), Item 1 (Business), and Item 1a (Risk Factors) are the most important sections. These results line up with prior literature's focus on these parts of annual reports (Hoberg and Phillips, 2016; Cohen et al., 2020; Hail et al., 2021), providing intuitive support for our approach. In contrast, Item 13 (Relations and Transactions), Item 12 (Ownership), and Item 10 (Directors and Governance) are consistently ranked among the least relevant items. Using the average importance scores in column (2), the most relevant item (Item 7) is almost twice as relevant as, for example, Directors and Governance (Item 10).

For each of the top four items (Item 7, Item 8, Item 1, and Item 1a), we plot the time series of their importance in Figure III, Panel B. Notably, we find that Items 7 and 8 exhibit an increasing time trend overall (although they experience a temporary decline during the COVID period), whereas Items 1 and 1a show a decreasing time trend. When regressing annual item-level averages on time trends (i.e., $\text{Item Importance}_{jt} = \gamma_0 + \gamma_1 \text{Year}_t + \varepsilon_t$, where $\text{Item Importance}_{jt}$ is the yearly average of item-level importance and Year denotes fiscal years), the coefficient on the year indicator is 0.0035 (t-value = 2.48) for Item 8 and −0.0021 (t-value = −4.91) for Item 1 (untabulated). The coefficients for the year trends for Item 1a and Item 7 are not statistically significant at conventional levels.

Collectively, our evidence reveals substantial heterogeneity in importance across different 10-K items, with Items 7, 8, 1, and 1a being the most relevant. However, we also observe differences in time trends among these items.

IV.C Analysis of Topics

Next, we use our paragraph-level importance scores to identify "topics" that investors tend to focus attention on when processing information.

Recall that we use an LLM to obtain comprehensive topic and subtopic lists for each of the four most important 10-K items (Items 1, 1a, 7, and 8). Then, for each item, we create a labeled training set by instructing GPT to classify paragraphs into topics and subtopics for

[20] Using the average of all paragraphs within an item yields a qualitatively similar result.

a sample of documents. Finally, we train a multi-label logit classifier on the labeled sample and scale it up to the full sample.

After classifying the paragraphs into topics, we average the importance scores within each topic and subtopic. Figure IV shows the most and least important topics and subtopics for Items 7, 8, 1, and 1a. We report all topics in the order of their average importance. We also show top 10 and bottom 10 subtopics, also sorted by their importance. We discard subtopics that appear less than 30 times in our sample.

Within the MD&A section (Item 7), segment information, financial performance, liquidity, and recent developments (new products, strategic developments, M&A deals) are ranked as the four most important topics. In contrast, corporate governance, accounting pronouncements, and forward-looking statements tend to be the topics that investors find least relevant. Out of the entire set of more granular subtopics, segment-related subtopics are the most important: segment profitability, segment performance, and segment revenue. Subtopics related to capital structure, such as divestitures and capital expenditures, are also among the most important subtopics. On the contrary, subtopics related to social responsibility, executive compensation, and ethical standards are identified as the least important from the market reaction standpoint. The subtopic related to the regulatory forward-looking statements ("assumptions") is also ranked as one of the least important.[21]

For Item 8 (financial statements and notes), topics directly related to operating performance and financial position receive the most attention from investors, whereas topics related to auditing and corporate governance receive the least attention, in line with our results for Item 7. Interestingly, goodwill and intangibles receive the highest attention on average within Item 8, confirming their increasing value relevance (Eisfeldt et al., 2020; De Ridder, 2024; Kepler et al., 2024). Similar to the MD&A, granular segment-level financial data also receives considerably more attention from investors. In contrast, audit-related matters and executive compensation are classified as the least important.

We find consistent results for Item 1 (Business Description) and Item 1a (Risk Factors). Financial performance and strategic initiatives are the most important topics in Item 1, while sustainability and corporate governance are the least important. Similarly, financial and strategic risks receive the highest attention scores for Item 1a, and reputational and cybersecurity risks receive the lowest attention scores.

Overall, our evidence suggests that topics directly related to a firm's segment performance, operating profitability, financial position, and liquidity, as well as recent developments

[21] Among forward-looking statements, the subtopic "business projections" ranks 33rd among all subtopics.

(new product launches and strategic developments), receive the highest amount of attention from equity markets.[22] Topics that are relatively stagnant and less relevant to performance and financial position, such as ESG, tend to receive less attention from investors.

Topic Variance Decomposition We also examine how much of the variation in paragraph-level importance can be explained by the topic indicators corresponding to each paragraph and report the results in Table V. While firm- and year-fixed effects jointly explain 22.10% of the total variation in paragraph-level importance, adding topic-subtopic indicators into the model increases the explained variation to 41.14%. Put differently, topics and subtopics explain approximately 53.72% ($= \frac{22.10\%}{41.14\%}$) of the explained variation in paragraph-level importance. The results are consistent across different sections of an annual report. Topic classifications account for 62.64% of the explained variation in Item 8, 46.85% in Item 1, and 45.30% in Item 1a. Taken together, these results show that topics explain substantial variation in the importance of disclosed content.

Alternative Topic Classifications We also revisit topics using the traditional topic classification method, namely, Latent Dirichlet Allocation (e.g., Dyer et al., 2017; Bybee et al., 2023, 2024). We rely on Dyer et al. (2017)'s topic list (developed specifically for the 10-K documents) and assign each paragraph in an annual report to one of the thirteen topics and one of 150 subtopics based on the cosine similarity of paragraph and topic (subtopic) embeddings.[23] Appendix C displays the most important topics and subtopics following Dyer et al. (2017) topic list. In Panel A, topics such as property and leasing, loans and debt, and business structure and M&A receive the most attention. In contrast, legal matters and executive-related information receive the least attention. In addition to the topic-level importance, we also report subtopic-level importance distribution in Panel B. Subtopics such as the transportation and freight industry, mortgages, and acquisitions receive the highest attention, while options, board of directors meetings, and unions receive the lowest atten-

[22]One potential concern is that our attention mechanism merely learns topics that are generally more or less likely to be relevant without considering the specific circumstances s within a given firm. To probe this, we investigate whether our model adjusts the paragraph-level importance when certain topics should be more important. To do so, we focus on paragraphs that discuss executive backgrounds (which tend to be unimportant) and patents and investigate whether their importance increases when informative changes occur for these topics. Specifically, we identify firm years with CEO changes and new patents and examine whether the paragraph-level importance increases during those events. In Appendix B, we find that paragraph-level importance increases *within* the firm when its CEO changes or when it is granted new patents.

[23]In particular, we use the "representative paragraphs" included in Dyer et al. (2017) topic list as the basis for topic classification. We obtain the contextual embeddings from the representative paragraphs of all 150 topics, which we refer to as topic embeddings. Then, we compute the cosine similarity between our paragraph embedding and all 150 topic embeddings. We assign the topic with the highest similarity score to the paragraph.

tion. Although the topics and subtopics in this analysis cannot be directly matched to the topic list that we develop above, the two approaches generally yield a similar conclusion that financial position and performance are the most relevant topics for investors.

V. Regulations and Disclosure Responses

In this and subsequent sections, we apply our methodology to answer two sets of economic questions. This section focuses on whether regulatory interventions affect the importance of specific information in explaining stock market reactions.[24] In particular, we are interested in whether investors place more attention on corporate communications after regulators intervene to improve their relevance and informativeness. More specifically, Section V.A examines the effect of the SEC's Modernization of Regulation S-K, which governs narrative disclosures, on the importance of the MD&A section. In Section V.B, we study the effect of the SEC's firm-specific actions (comment letters), targeting the informativeness of specific disclosures on the importance of these disclosures in explaining market reactions.

V.A The SEC's Modernization

The SEC's modernization of Regulation S-K significantly reshaped the disclosure of narrative information, with a primary focus on the MD&A section. Its objective was to improve the usefulness of narrative disclosures to investors. To this end, the new requirements aimed to enhance the relevance of disclosed information, emphasizing a principles-based approach and focus on forward-looking information. Companies were encouraged to tailor disclosures to their unique circumstances, highlighting the most critical factors relevant to understanding a company's performance ability to meet financial obligations. The regulatory change went into effect on August 9th, 2021.

Research Design Empirically assessing the effect of the regulatory change, however, on the relevance of disclosed information is not a trivial task. Although one can measure a change in MD&A content (e.g., length, existence of tables), the usefulness of specific disclosures is an elusive concept. Our approach allows us to tackle this issue directly.

Our empirical strategy exploits the feature that the SEC's modernization initiative did not affect narrative disclosures in Item 8 (Financial Statements and Notes) as they are

[24] Prior studies examining the effect of regulations on disclosure quality generally focus on large regulatory changes such as the Security Exchange Acts (Greenstone et al., 2006), IFRS adoption (Leuz and Wysocki, 2008) or European securities regulation (Christensen et al., 2016). Our focus is different.

governed by a different regulation.[25] Therefore, we examine whether the importance of Item 7 increased shortly after the regulatory change went into effect compared to the importance of Item 8 (control group). To do so, we confine our sample to observations from August 9th, 2020 (one year before the change) and August 9th, 2022 (one year after the change) and focus on the importance scores of Items 7 and 8.[26] We then estimate the following difference-in-differences regression:

$$\text{Item Importance}_{ikt} = \beta_1 \text{Treat}_{ikt} + \beta_2 \text{Post}_{ikt} + \beta_3 \text{Treat}_{ikt} \times \text{Post}_{ikt} + \delta_i \times \zeta_t + \varepsilon_{ikt} \tag{5}$$

Item Importance$_{ikt}$ is the item-level importance (average importance of paragraphs within a 10-K Item) from firm i and year t. k is an index for Item 7 or Item 8. Treat$_{ikt}$ is an indicator that equals one when k is Item 7 and zero otherwise. Post$_{ikt}$ is an indicator that equals one when the filing date of firm i's 10-K is after August 9th, 2021, and zero otherwise. If the SEC's Act successfully served its purpose, the coefficient on the interaction term (β_3) should be positive and statistically significant. We include granular *firm* $\times$ *year* fixed effects in our main model specification, which allows us to measure the relative importance between Item 7 and Item 8 within each document. In another specification, we also include firm and year-fixed effects separately to study the within-firm time-series variation in item-level importance. Standard errors are clustered by firm.

Results The results are displayed in Table VI. The first column includes firm $\times$ year-fixed effects, whereas the second column independently includes firm and year-fixed effects. Consistent with our item-level importance results in Table IV, Item 7 is more important than Item 8 ($\beta_1 = 0.0224$, t-value = 28.34).[27] Notably, we find a positive and statistically significant coefficient β_3 ($\beta_3 = 0.0052$, t-value = 5.30). This result indicates that the relative importance of Item 7 compared to Item 8 within a single firm increased immediately after the regulatory change. In terms of economic significance, the new requirements account for an increase of 18.84% ($\frac{0.0052}{0.0052+0.0224}$) in the relative importance of Item 7 or a 0.09 standard deviation ($\frac{0.0052}{0.0573}$) increase in item-level importance. The importance of Item 7 within 10-K filings increased significantly immediately following the mandated compliance date.

Column (2), which allows the measurement of within-firm time-series changes in item-

[25] Several notable changes in accounting standards were introduced in the late 2010s, such as ASC 842 (leases) in 2018, ASC 606 (revenue recognition) in 2017, and ASC 326 (current expected credit losses) in 2019. However, none of these regulations were newly enforced around August 2021.

[26] In the main specification, we use the tight one-year-window to better estimate the effect of the regulation. However, the results are qualitatively similar even when we expand the window to two years surrounding the enforcement date.

[27] The coefficient on Post is omitted as we control for document-level fixed effects.

level importance, shows consistent results. The magnitude of the coefficients remains similar with fixed effect structures. The coefficient on Post, which reflects the time-trend in Item 8's importance, is statistically insignificant (coefficient = -0.0035, t-value = -1.20), which indicates that our control group (i.e., Item 8) did not experience a systematic time-series shift over our two-year sample period.[28]

Collectively, we find evidence that the SEC's Modernization Initiation of 2021, which specifically targeted narrative disclosures in Item 7, increased investors' attention allocated to these disclosures, and hence accomplished its intended purpose.

V.B Comment Letters

In this subsection, we estimate whether regulatory scrutiny of a firm's disclosures affects the importance of these disclosures in explaining stock prices. The SEC's comment letters are a mechanism through which the SEC ensures and enforces the compliance of companies with disclosure regulations (Ryans, 2021). When the SEC identifies potential deficiencies in the disclosed information or seeks clarification, it issues a comment letter to the company, prompting the registrant to respond and, if necessary, amend their filings. The comment letter may ask the company to revise its disclosure, provide additional disclosures, or file a different disclosure in a future SEC filing. Although the SEC's periodic reviews routinely produce comment letters, they can also be triggered by specific concerns regarding a company's filing. Specifically, regarding Regulation S-K, which governs non-financial disclosure requirements, comment letters frequently address issues related to MD&A (Management's Discussion and Analysis), executive compensation, risk factors, and other critical narrative disclosures.

Previous studies show that SEC comment letters can enhance narrative disclosure content and have real effects (Brown et al., 2018; Ryans, 2021). These studies use firm-level outcome variables in measuring the impact of granular SEC comment letters, whereas we are interested in the effect of comment letters on the attention investors place on specific 10-K content.[29] Using our granular setup, we provide the first evidence of whether the item-level importance of the items flagged by the SEC increased after firms received the letters.

Research Design To estimate the effect of the SEC's comment letters on the affected items, we start with the universe of the SEC's comment letters from the Audit Analytics

[28] Note that Post indicator is not subsumed by year fixed effects because the regulatory change took place in the middle of the year.

[29] One exception is (Bernard et al., 2023), who use transaction-level categories based on XBRL tags and show that more complex disclosures lead to more related SEC enforcement letters.

database and limit the sample to letters referencing the sections of Regulation S-K. We then match the sections within Regulation S-K to items in the 10-K filings. Subsequently, we identify item-firm-year observations flagged by the SEC and match them to the item-level importance score. Our final dataset retains the firm-item pairs for the year the item was referenced by the SEC's comment and the subsequent year.[30] To test for the effect of the regulator's attention, we estimate the following regression:

$$\text{Item Importance}_{ikt} = \beta_1 \text{Post}_{ikt} + \delta_i \times \eta_k + \zeta_t + \varepsilon_{ikt} \tag{6}$$

where Item Importance$_{ikt}$ is our item-level importance score for item k in firm i year t 10-K filing. Post$_{ikt}$ is an indicator that equals one when firm i's item k was flagged by the SEC in year $t-1$ and zero otherwise. We include two extensive sets of fixed effects. The first set of fixed effects is the interaction of firm and item-fixed effects. As our dataset includes only two periods, including firm $\times$ item fixed effects allows us to identify the direct effect of the SEC's comment letter on the specific item within the target firm. We also include time-fixed effects to control for time trends in item-level importance. Standard errors are clustered at the firm level.

Results This pre-post analysis provides a powerful way to examine the effects of SEC's attention at the granular level. We report the results in Table VII. In column (1), we include only firm $\times$ item fixed effects, and in column (2), we add time-fixed effects. As we control for firm $\times$ item fixed effects, the coefficient on Post denotes the change in importance of that specific item immediately after the firm receives the SEC's comment letters. We find β_1 to be positive and statistically significant. Even though the inclusion of time-fixed effects in column (2) reduces the magnitude of the coefficient from 0.0142 to 0.0098, its economic significance is sizable. The SEC's comment letter increases the item-level importance by 0.1 standard deviations ($=\frac{0.0098}{0.0984}$).[31]

In sum, we find that the SEC's comment letters increase the importance of the affected items in investors' pricing decisions. Together with our results in Section V.A, our evidence supports the notion that regulatory interventions lead to disclosure responses and that such responses improve the informativeness of narrative disclosures.

[30] For example, if Item 7 for firm A in 2022 is flagged by the SEC, we also include Item 7 for firm A in 2023 in the final dataset, allowing for a longitudinal comparison of flagged items across consecutive years within the same firm.

[31] Sometimes, a firm receives two or more comment letters in consecutive years. In such cases, a firm may receive another comment letter when Post$_{ikt} = 1$. In untabulated analysis, we exclude such instances and re-estimate the regression. The results are qualitatively consistent.

VI. STRATEGIC POSITIONING OF INFORMATION

In this section, we examine whether firms strategically position information communicated to investors. The motive behind the positioning choices stems from the notion that investors with cognitive constraints focus their efforts on the primary message communicated by a firm (Sims, 2003). In particular, we hypothesize that the most relevant information will be discussed relatively early in each section, in line with the informational demands of investors. However, this also creates incentives to front-load more positive information to diffuse negative surprises and obfuscate negative performance. Indeed, prior studies suggest that managers appear to obfuscate poor performance by reducing its readability (Li, 2008; Li et al., 2013; DeHaan et al., 2021) and the quality of vocal delivery (Baik et al., 2023). However, because value-relevant information tends to be more complex and less readable (Bushee et al., 2018; Kim et al., 2024b), directly testing *where* the *important* information is presented provides a valuable opportunity to shed more light on strategic information communication.

To examine whether managers *position* of important information differently, we focus our analysis on the MD&A disclosures as managers have substantial flexibility in terms of how their content is presented (Cohen et al., 2020). We expect that firms place more important information upfront, but at the same time, they may strategically distort the information order to the extent that it communicates negative news or uncertainty. Specifically, when managers need to present negative performance to investors, they will place the information towards the middle or the end of the MD&A section.

VI.A Research Design

Although straightforward, testing the above hypothesis is empirically non-trivial as it requires a paragraph-level (or even sentence-level) importance measure. Taking advantage of our data, we estimate the following OLS regression:

$$\begin{aligned}\text{Paragraph Position}_{ikt} = {} & \beta_1 \text{Paragraph Importance}_{ikt} + \beta_2 \text{Obfuscation Incentives}_{it} \\ & + \beta_3 \text{Paragraph Importance}_{ikt} \times \text{Obfuscation Incentives}_{it} + \varepsilon_{ikt}\end{aligned} \tag{7}$$

Paragraph Position$_{ikt}$ is the position score of paragraph k in firm i's year t MD&A. Positional score is computed as follows:

$$\text{Paragraph Position}_{ikt} = \left(1 - \frac{\text{Position}_{ikt}}{N_{it}}\right) \tag{8}$$

where Position_{ikt} is the ordered location of the paragraph k and N_{it} is the number of paragraphs included in firm i's time t MD&A. Therefore, the first paragraph is assigned the highest score, close to one, and the last paragraph is assigned a score of zero. Obfuscation Incentives_{it} are cross-sectional variables measured at the firm-year level and include sentiment measured using the dictionary by Loughran and McDonald (2011), the degree of competition (Li et al., 2013), operating profitability Ball et al. (2015), and earnings volatility (measured as the standard deviation of the past five years' EPS values). Standard errors are clustered by firms. Continuous variables are winsorized at the 1% and 99% levels.

VI.B Results

The results are presented in Table VIII. In column (1), we begin by testing our prediction that, on average, more important paragraphs are placed earlier in the MD&A compared to less important paragraphs. In this specification, we include firm $\times$ year fixed effects to measure the effect within each MD&A. The coefficient on Paragraph Importance is positive and statistically significant (coefficient = 0.3027, t-value = 88.59). This result indicates that the more important paragraphs are placed up-front in an MD&A. In terms of economic significance, a one standard deviation increase in paragraph-level importance is associated with a 0.12 standard deviation increase in the paragraph position ($= \frac{0.3027 \times 0.1103}{0.2870}$).

In columns (2) - (6), we examine the effect of managerial obfuscation incentives on the positioning of information. Note that we do not include fixed effects in this specification as the obfuscation incentives are measured at the firm-year level.[32] We find that when managers deliver news with negative sentiment, face more intense competition, report low profitability, and have high earnings volatility, the positive association between importance and position becomes weaker. In terms of economic significance, a one standard deviation increase in sentiment leads to an 8% increase in the relation between importance and position (computed as std(Obfuscation Incentives) $\times \frac{\beta_3}{\beta_2}$). Other proxies for obfuscation incentives yield effects of comparable economic magnitudes.[33] In column (6), we include the four proxies of obfuscation incentives simultaneously in one regression. The statistical significance and the magnitude of the coefficients remain quantitatively similar.

Overall, we find evidence consistent with managers' obfuscation incentives weakening the positive association between paragraph importance and its position within the MD&A text.

[32] However, in untabulated tests, we include firm-year fixed effects in these regressions and confirm that the results for our interaction terms are robust.

[33] A one standard deviation increase in competition leads to a 3% decrease in the relation between importance and position; a one standard deviation increase in profitability leads to a 4.3% increase in the relation between importance and position; and a one standard deviation increase in earnings volatility leads to an 8.7% increase in the relation between importance and position.

This implies that managers are less forthcoming when they need to present important, yet potentially negative information to investors.

Document-Level Measurement Based on our finding that more important paragraphs are likely to be placed earlier within an MD&A, we construct a firm-level measure of delayed information positioning. The idea behind the measure is to assign a higher score to an MD&A that is more forthcoming. Specifically, we measure the document-level information positioning as follows:

$$\text{Information Positioning}_{it} = \sum_{k=1}^{N_{it}} \left[\left(1 - \frac{\text{Position}_{ikt}}{N_{it}}\right) \times \text{Paragraph Importance}_{ikt} \right] \tag{9}$$

This measure obtains the highest value if companies present information in descending order of importance. The more distorted this relation is, the lower the measure becomes. In Appendix D, we investigate some descriptive economic determinants of this measure. Large, mature firms (i.e., higher book-to-market ratio and larger size) tend to place more important paragraphs up-front. Consistent with our results in Table VII, we also find that when companies report a loss, have lower profitability, have higher volatility of earnings, report negative sentiment news and have a less readable MD&A, their information positioning score tends to be lower. The results are robust to both firm- and industry-fixed effects.

VII. CONCLUSION

As disclosures become lengthier and costlier to process, theory suggests that investors allocate their restrained cognitive resources to the most relevant information (Simon, 1955; Sims, 2003). However, what information do investors focus on when processing large volumes of unstructured textual data? Answering this question poses a major methodological challenge. In this paper, we introduce an attention mechanism that learns to understand textual information that is more vs. less relevant to investors.

We show that our attention-based model explains contemporaneous market reactions, as well as predicts future returns and EPS changes significantly better than attention-free machine learning methods. This result underscores the informational value of the attention mechanism when learning qualitative fundamentals from text and indicates the importance of incorporating relational importance among paragraphs within corporate disclosures.

We then provide evidence on which sections within an annual report and which disclosure topics are more relevant for investors. We find that Management Discussion and Analysis (Item 7), Financial Statements and Notes (Item 8), Business Overview (Item 1), and Risk

Factors (Item 1a) receive the highest attention. In terms of the topic, investors focus attention on content related to segment performance, profitability, liquidity, and intangibles, among others. At the same time, topics related to compliance, auditing, and ESG receive the least attention judging by their contribution to market reactions.

Using a granular measure of paragraph-level importance, we also answer two economic questions. First, we provide causal evidence that the SEC's regulatory interventions improve the importance of narrative disclosures from an equity investor's standpoint. Second, we find that companies strategically position important information within an annual report presumably to influence cognitively constrained investors' focus.

References

Baik, B., A. Kim, D. S. Kim, and S. Yoon (2023). Managers' vocal delivery and real-time market reactions in earnings calls. *Chicago Booth Research Paper* (23-21).

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. V. Nikolaev (2015). Deflating profitability. *Journal of Financial Economics 117*(2), 225–248.

Bernard, D., E. Blankespoor, T. de Kok, and S. Toynbee (2023). Confused readers: A modular measure of business complexity. University of Washington Working Paper.

Bloomfield, R. (2008). Discussion of "annual report readability, current earnings, and earnings persistence". *Journal of Accounting and Economics 45*(2-3), 248–252.

Brown, S. V., X. Tian, and J. Wu Tucker (2018). The spillover effect of sec comment letters on qualitative corporate disclosure: Evidence from the risk factor disclosure. *Contemporary Accounting Research 35*(2), 622–656.

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

Bybee, L., B. Kelly, A. Manela, and D. Xiu (2024). Business news and business cycles. *The Journal of Finance 79*(5), 3105–3147.

Bybee, L., B. Kelly, and Y. Su (2023). Narrative asset pricing: Interpretable systematic risk factors from news text. *The Review of Financial Studies 36*(12), 4759–4787.

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

Cao, S., W. Jiang, J. Wang, and B. Yang (2024). From man vs. machine to man+ machine: The art and ai of stock analyses. *Journal of Financial Economics 160*, 103910.

Chen, X., Y. H. Cho, Y. Dou, and B. Lev (2022). Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research 60*(2), 467–515.

Chen, Y., B. T. Kelly, and D. Xiu (2022). Expected returns and large language models. *Available at SSRN*.

Christensen, H. B., L. Hail, and C. Leuz (2016, 07). Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement. *The Review of Financial Studies 29*(11), 2885–2924.

Cohen, L., C. Malloy, and Q. Nguyen (2020). Lazy prices. *The Journal of Finance 75*(3), 1371–1415.

Costello, A. M., B. L. Levy, and V. Nikolaev (2024). Uncovering information: Can language models tell us where to look? Working Paper.

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review 114*(1), 199–251.

DeHaan, E., Y. Song, C. Xie, and C. Zhu (2021). Obfuscation in mutual funds. *Journal of*

Accounting and Economics 72(2-3), 101429.

Dyer, T., M. Lang, and L. Stice-Lawrence (2017). The evolution of 10-k textual disclosure: Evidence from latent dirichlet allocation. *Journal of Accounting and Economics 64*(2-3), 221–245.

Eisfeldt, A. L., E. Kim, and D. Papanikolaou (2020). Intangible value. Technical report, National Bureau of Economic Research.

Gabaix, X., R. S. Koijen, R. Richmond, and M. Yogo (2023). Asset embeddings. *Available at SSRN 4507511*.

Greenstone, M., P. Oyer, and A. Vissing-Jorgensen (2006). Mandated disclosure, stock returns, and the 1964 securities acts amendments. *The Quarterly Journal of Economics 121*(2), 399–460.

Gu, S., B. Kelly, and D. Xiu (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies 33*(5), 2223–2273.

Guay, W., D. Samuels, and D. Taylor (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics 62*(2-3), 234–269.

Hail, L., M. Muhn, and D. Oesch (2021). Do risk disclosures matter when it counts? evidence from the swiss franc shock. *Journal of Accounting Research 59*(1), 283–330.

Hassan, T. A., S. Hollander, L. Van Lent, and A. Tahoun (2019). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics 134*(4), 2135–2202.

Hoberg, G. and G. Phillips (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy 124*(5), 1423–1465.

Jiang, J., B. Kelly, and D. Xiu (2023). (re-) imag (in) ing price trends. *The Journal of Finance 78*(6), 3193–3249.

Kelly, B. and S. Pruitt (2013). Market expectations in the cross-section of present values. *The Journal of Finance 68*(5), 1721–1756.

Kelly, B. and S. Pruitt (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics 186*(2), 294–316.

Kepler, J. D., C. McClure, and C. R. Stewart (2024). Competition enforcement and accounting for intangible capital. *Fama-Miller Working Paper, Chicago Booth Research Paper* (24-20).

Kim, A., M. Muhn, and V. Nikolaev (2024a). Financial statement analysis with large language models. *arXiv preprint arXiv:2407.17866*.

Kim, A., M. Muhn, and V. V. Nikolaev (2024b). Bloated disclosures: Can chatgpt help investors process information? *Chicago Booth Research Paper* (23-07), 2023–59.

Kim, A. and V. V. Nikolaev (2023). Context-based interpretation of financial information. *Chicago Booth Research Paper* (23-08).

Kusupati, A., G. Bhatt, A. Rege, M. Wallingford, A. Sinha, V. Ramanujan, W. Howard-Snyder, K. Chen, S. Kakade, P. Jain, et al. (2022). Matryoshka representation learning.

Advances in Neural Information Processing Systems 35, 30233–30249.

Leuz, C. and P. D. Wysocki (2008). Economic consequences of financial reporting and disclosure regulation: A review and suggestions for future research. *Available at SSRN*.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics 45*(2-3), 221–247.

Li, F., R. Lundholm, and M. Minnis (2013). A measure of competition based on 10-k filings. *Journal of Accounting Research 51*(2), 399–436.

Lo, K., F. Ramos, and R. Rogo (2017). Earnings management and annual report readability. *Journal of Accounting and Economics 63*(1), 1–25.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance 66*(1), 35–65.

Roll, R. (1988). Rˆ2. *The Journal of Finance 43*(3), 541–566.

Ryans, J. P. (2021). Textual classification of SEC comment letters. *Review of Accounting Studies 26*(1), 37–80.

Sarkar, S. K. and K. Vafa (2024). Lookahead bias in pretrained language models. *Available at SSRN*.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Tetlock, P. C., M. Saar-Tsechansky, and S. Macskassy (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance 63*(3), 1437–1467.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.

Appendix A. Summaries of the Most and Least Important Paragraphs

This table displays the example summaries of the most and least important paragraphs. We randomly choose five documents and identify their most important paragraphs. We do the same for another five paragraphs and identify the least important paragraphs. We use GPT-4o to summarize the content within these paragraphs.

Panel A. Summaries of the Most Informative Paragraphs
For fiscal year 2007, the company's net revenues increased by 7.1% compared to 2006. After accounting for a 0.4% decrease due to foreign exchange effects, the remaining 7.5% increase was driven primarily by growth in disposable and software support revenues. In particular, disposable revenue growth resulted from higher unit sales in the U.S. for plasma and red cell product lines, along with price improvements in the OrthoPat product line. However, these gains were partially offset by lower unit volumes in Japan for the bloodbank and plasma product lines.
In 2008, the company's revenues grew by 28%, reaching $107.7 million compared to $84.1 million in 2007. This increase was driven by a 7% rise in unit sales, largely to tier one customers, and a 19% increase in average selling prices. The higher prices were primarily due to increased sales of adapter cards, switch systems, and QDR and DDR products, which have higher average selling prices. However, the company cautions that the 2008 revenue results may not reflect future performance.
For the year ended December 31, 2008, the company's revenues increased by 28%, reaching $107.7 million, compared to $84.1 million in 2007. This growth was driven by a 7% increase in unit sales, primarily to tier one customers, and a 19% rise in average selling prices. The higher average selling prices were attributed to increased sales of adapter cards and switch systems, as well as QDR and DDR products, which typically command higher prices. However, the company notes that the 2008 revenue figures may not be indicative of future performance.
The company's future growth through acquisitions and branch expansion could be negatively impacted by several factors. These include strong competition from other financial institutions in the company's current and future markets, which may hinder its ability to acquire new institutions. Additionally, acquisitions and new branch openings require regulatory approval, which depends on factors such as the results of regulatory examinations and the company's Community Reinvestment Act (CRA) ratings. This regulatory scrutiny could also pose challenges to the company's expansion plans.
During fiscal 2022, the company received $3.2 billion in new equity funding related to its majority-owned Flipkart subsidiary, reducing its ownership from 83% to 75%. Short-term borrowings increased to $0.4 billion as of January 31, 2022, compared to $0.2 billion in 2021, with a rise in the weighted-average interest rate from 1.9% to 2.9%. The company also has $15.0 billion in undrawn committed lines of credit in the U.S., providing additional liquidity, alongside access to credit facilities outside the U.S. to support Walmart International operations.
Panel B. Summaries of the Least Informative Paragraphs
The audit was conducted following the standards of the Public Company Accounting Oversight Board (PCAOB). The audit aimed to ensure that internal control over financial reporting was effective in all material respects. Key procedures included understanding internal control, evaluating management's assessment, testing design and operational effectiveness, and other necessary steps. The auditors believe that the audit provides a reasonable basis for their opinion.
The company has implemented an Accounting and Finance Code of Ethics for its top executives and an additional Code of Conduct for its directors and employees. Both codes are available on the company's website. Any waivers or amendments related to these codes, particularly concerning executive officers and directors, will be disclosed on the same site.
Patent litigation could be costly and time-consuming for the company, potentially diverting technical and management personnel from their regular duties. The company may initiate lawsuits to protect or enforce its patents and could also face legal actions aimed at invalidating its patents or preventing future patent approvals. The outcome of such legal disputes could risk patents being invalidated or narrowly interpreted. Additionally, there is a possibility of confidential information being disclosed during litigation, and public announcements related to legal proceedings could negatively impact investor perception and affect the trading price of the company's Class B common stock.
Patent litigation could be costly and time-consuming for the company, potentially diverting technical and management personnel from their regular duties. The company may initiate lawsuits to protect or enforce its patents and could also face legal actions aimed at invalidating its patents or preventing future patent approvals. The outcome of such legal disputes could risk patents being invalidated or narrowly interpreted. Additionally, there is a possibility of confidential information being disclosed during litigation, and public announcements related to legal proceedings could negatively impact investor perception and affect the trading price of the company's Class B common stock.
Challenges exist in developing bispecific T cell-engagers, particularly in designing molecules that activate T cells only in tumor environments to avoid severe toxicities from peripheral T cell activation. Additionally, manufacturing bispecific antibodies at scale is difficult due to the complexity of protein expression in conventional systems. However, advancements in protein engineering are expected to enable the development of new bispecific antibodies with improved therapeutic activity, safety, and manufacturability. In immuno-oncology, cytokines, which regulate immune cell function, represent a potential new approach for cancer treatment by stimulating the immune system to target cancer cells.

Appendix B. Less Important Topics, Exceptional Events, and Importance

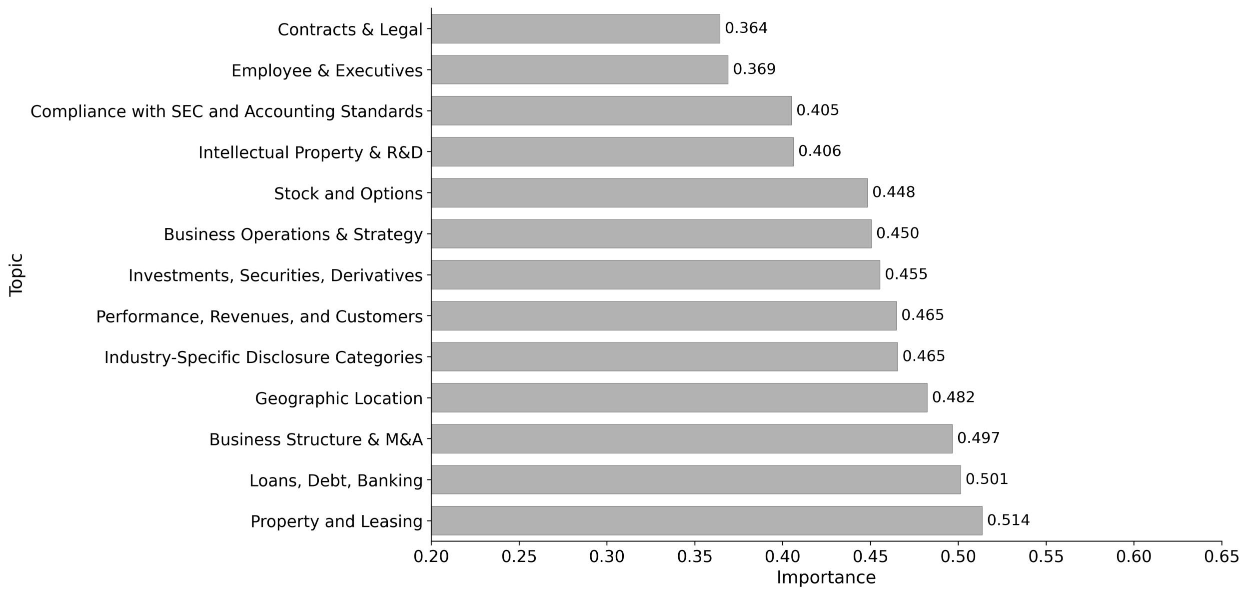
This table reports whether our algorithm can identify unusual yet meaningful events for firms. We choose two topics that are identified to be less important: patents, and executives. We confine our samples to paragraphs with topics related to executive and patents only. Then, we regress paragraph-level importance on the indicators for CEO changes and new patents granted, respectively, along with firm and year-fixed effects. We confine our search to Item 1 and Item 7. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable =	Paragraph-Level Importance			
Topic =	CEO		Patents	
	Full (1)	Item 1&7 (2)	Full (3)	Item 1&7 (4)
CEO Changes	0.0130*	0.0155**		
	(1.98)	(2.10)		
New Patents			0.0055	0.0080*
			(1.33)	(1.70)
Firm Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Adjusted R2	0.2562	0.44	0.3468	0.3627
N	37,038	2,616	887,367	197,820

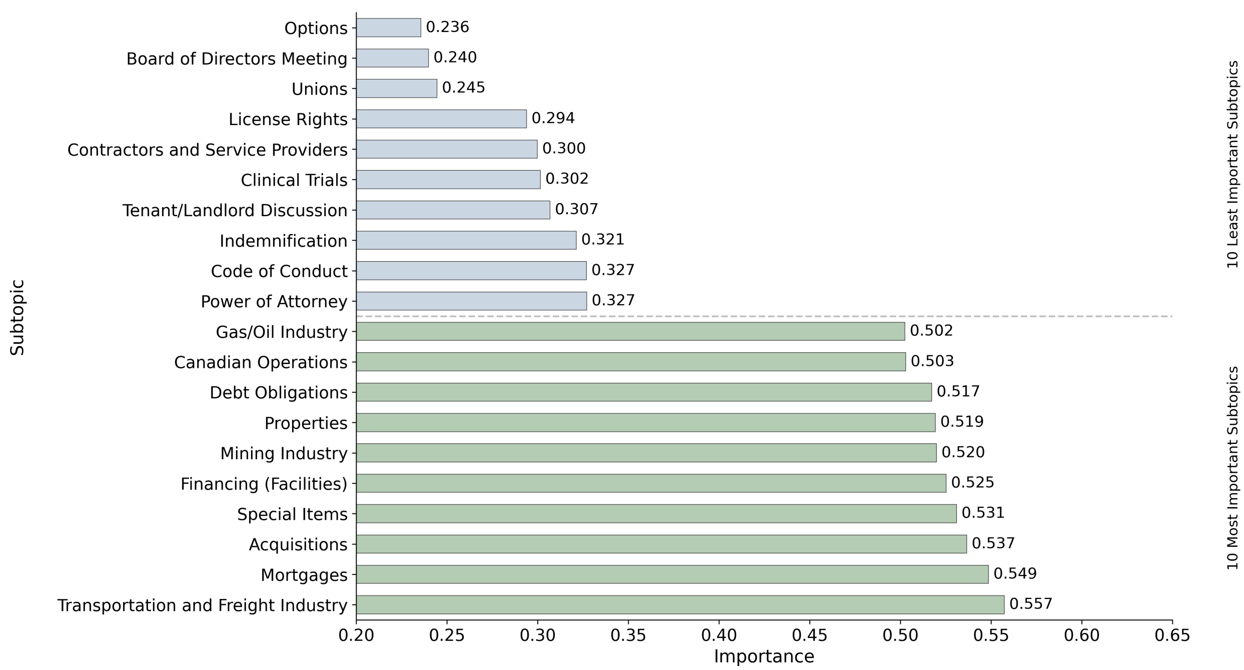
Appendix C. Topics and Subtopics Following (Dyer et al., 2017)

Our methodology involves the following steps: we calculate an importance score for each topic and subtopic by averaging the similarity scores of all paragraphs classified under each category across all 10-K filings. We then rank all topics and subtopics based on their average importance scores. For detailed analysis, we present the 10 most important and 10 least important subtopics, as shown in the figure.

Panel A: Topic Importance



Panel B. Subtopic Importance



Appendix D. Firm-Level Positional Importance

In this table, we report descriptive statistics of our firm-level positional importance measure, examine the determinants, and document its capital market consequences. Firm-level positional importance measures whether relatively more important paragraphs are placed upfront within the MD&A section (Item 7) of a firm's annual filing. Panel A reports the descriptive statistics of firm-level positional importance and other variables used in the regression. Panel B reports the determinants of firm-level positional importance. We include the natural logarithm of the firm's market capitalization (Size), the natural logarithm of book-to-market ratio (Book-to-Market), the standard deviation of the past five years' EPS values (Earnings Volatility), an indicator for operating loss (Loss), the gap between fiscal year and reporting date (Report Lag), return-on-asset (ROA), the textual sentiment of the MD&A section (Sentiment), readability of the MD&A section measured by Fog index (Readability), and the natural logarithm of the number of words contained in the MD&A section (Length). In column (1), we include year and industry (SIC 2-digit level) fixed effects, and in column (2), we include year and firm fixed effects.

Panel A. Descriptive Statistics

Variable	N	Mean	Std	P25	P50	P75
Upfrontedness	66,757	0.5161	0.0243	0.5012	0.5143	0.5283
Size	66,757	6.4086	2.0516	4.9609	6.4021	7.8106
Book to Market	66,757	-0.8062	0.9620	-1.3043	-0.6983	-0.1947
Earnings Volatility	66,757	0.2067	2.3509	0.0141	0.0378	0.1081
Loss	66,757	0.3297	0.4701	0.0000	0.0000	1.0000
ROA	66,757	-0.0273	1.7932	-0.0321	0.0170	0.0610
Report Lag	66,757	51.8779	20.7117	35.0000	51.0000	65.0000
Sentiment	66,757	-0.2920	0.2210	-0.4499	-0.3090	-0.1544
Readability	66,757	14.2080	1.4270	13.2408	14.1140	15.0796
Length	66,757	9.0447	0.5227	8.7497	9.0703	9.3679

Panel B. Determinants of Firm-Level Positional Importance Scores

Dep Var =	Firm-Level Positional Importance Scores	
	(1)	(2)
Size	0.1742***	0.0395*
	(6.99)	(1.99)
Book to Market	0.1523***	0.0565***
	(5.31)	(2.90)
Loss	-0.0114**	-0.0894***
	(-2.26)	(-3.51)
ROA	0.0083***	0.0222**
	(3.44)	(2.52)
Earnings Volatility	-0.0014**	0.0007
	(-2.20)	(0.04)
Report Lag	0.0017	-0.0014*
	(0.32)	(-1.81)
Sentiment	0.0466**	0.0803*
	(2.37)	(1.87)
Readability	-0.0775**	-0.1342***
	(-2.03)	(-4.98)
Length	-1.8366***	-2.0819***
	(-15.58)	(-13.01)
Year FE	Yes	Yes
Industry FE	Yes	No
Firm FE	No	Yes
Adjusted R2	0.2002	0.5476
N	66,757	66,757

Appendix E. Complete Topic-Subtopic List

Panel A: Item 7 (MD&A)	
Topic	Subtopic
Business Overview	Business Strategy Company Description Geographic Presence Industry Trends Market Position Product Offerings Others
Contractual Obligations	Debt Obligations Lease Obligations Purchase Commitments Others
Corporate Governance	Board of Directors Ethical Standards Executive Compensation Shareholder Relations Others
Critical Accounting Policies	Allowance for Doubtful Accounts Goodwill Impairment Income Taxes Inventory Valuation Revenue Recognition Share-Based Compensation Others
Financial Performance	EBITDA Earnings Per Share Expenses Gross Profit Net Income Operating Income Revenues Others
Forward-Looking Statements	Assumptions Future Outlook Growth Strategy Market Opportunities Potential Risks Projections Others
Liquidity and Capital Resources	Capital Expenditures Cash Flow Credit Facilities Debt Management Financing Activities Investing Activities Working Capital Others
Off-Balance Sheet Arrangements	Commitments Contingent Liabilities Guarantees Leases Variable Interest Entities Others

Continued on next page

Panel A: Item 7 (MD&A)	*(continued)*
Topic	Subtopic
Recent Accounting Pronouncements	Adoption Impact Impact Assessment Implementation Plans New Standards Others
Recent Developments	Acquisitions Divestitures New Products Strategic Initiatives Others
Regulatory and Legal Matters	Compliance Environmental Compliance Legal Proceedings Legislative Changes Regulatory Changes Others
Risk Factors	Competitive Risks Economic Conditions Financial Risks Market Risks Operational Risks Regulatory Risks Others
Segment Information	Geographic Segments Intersegment Transactions Product Segments Segment Performance Segment Profitability Segment Revenue Others
Sustainability and CSR	Environmental Impact Social Responsibility Sustainability Initiatives Others

Panel B: Item 8 (Financial Statements and Supplementary Data)	
Topic	Subtopic
Accounting Policies	Amortization Basis of Presentation Consolidation Deferred Revenue Depreciation Derivative Instruments Derivatives and Hedging Equity Method Investments Estimates and Judgments Exit Activities Expense Recognition Fair Value Measurements Foreign Currency Goodwill Impairment Impairment of Assets Income Taxes Intangible Assets Leases Pension and Other Postretirement Benefits Revenue Recognition Segment Information (Accounting Policies) Self-Insurance Share-Based Compensation Stock-Based Compensation Use of Estimates Valuation Methods Vendor Funds Others
Auditor's Report	Audit Opinion Audit Standards Basis for Opinion Critical Audit Matters Internal Control Assessment Opinion Others
Cash Flow	Financing Activities Investing Activities Operating Activities Others
Corporate Governance	Audit Committee Board of Directors Executive Compensation Internal Controls Related Party Transactions Shareholder Rights Strategic Planning Others
Financial Performance	Comprehensive Income Earnings Per Share Expenses Income Tax Expense Interest Income Net Income Noninterest Income Operating Expenses Operating Income Provision for Loan Losses Revenues Segment Information Others

Continued on next page

Panel B: Item 8 (Financial Statements and Supplementary Data) *(continued)*

Topic	Subtopic
Financial Position	Assets Borrowings Cash and Cash Equivalents Debt Deferred Taxes Deposits Equity Goodwill and Intangibles Investment Securities Investments Liabilities Loans Other Assets Other Liabilities Property, Plant and Equipment Regulatory Assets and Liabilities Working Capital Others
Legal and Regulatory	Accounting Standards Capital Requirements Compliance Dividend Restrictions Financial Reporting Legal Proceedings Litigation Tax Compliance Tax Matters Others
Risk Management	Allowance for Loan Losses Capital Adequacy Concentration Risk Credit Risk Derivatives Hedging Activities Interest Rate Risk Legal Contingencies Liquidity Risk Loan Quality Market Risk Nonperforming Assets Operational Risk Troubled Debt Restructurings Others
Supplementary Data	Commitments and Contingencies Discontinued Operations Fair Value Measurements Geographic Information Goodwill and Intangible Assets Investments Major Customers Notes to Financial Statements Quarterly Financial Data Segment Information (Supplementary Data) Subsequent Events Unaudited Quarterly Data Others

Panel C: Item I (Business)	
Topic	Subtopic
Business Overview	Business Model Company Description Company History Core Mission Geographic Presence Industry Overview Market Position Others
Competition	Barriers to Entry Competitive Advantages Competitive Landscape Industry Trends Market Share
Corporate Governance	Board of Directors Corporate Policies Ethical Standards Executive Management Governance Policies Shareholder Relations Others
Financial Performance	Assets Cash Flow Equity Expenses Financial Metrics Liabilities Profitability Revenues Segment Performance Others
Future Outlook	Growth Opportunities Market Trends Strategic Initiatives Others
Human Resources	Compensation and Benefits Diversity and Inclusion Employee Relations Talent Management Training and Development Workforce Others
Legal and Regulatory	Compliance Issues Legal Proceedings Litigation Regulatory Investigations Others
Market Strategy	Competitive Strategy Customer Segments Growth Strategy Market Expansion Marketing Strategy Partnerships Sales Strategy Others
Operations	Distribution Facilities Logistics Manufacturing Operational Efficiency Quality Assurance Supply Chain Technology Infrastructure Others

Continued on next page

Panel C: Item I (Business))	*(continued)*
Topic	Subtopic
Products and Services	Customer Segments Product Development Product Offerings Service Offerings Technology Platform Others
Regulatory Environment	Compliance Data Privacy Environmental Regulations Government Policies Licensing Safety Standards Others
Risk Management	Credit Risk Financial Risks Legal Risks Market Risks Operational Risks Regulatory Risks Others
Strategic Initiatives	Expansion Plans Investment Criteria Mergers and Acquisitions Partnerships Strategic Partnerships Others
Sustainability and CSR	Community Engagement Corporate Citizenship Environmental Impact Social Responsibility Sustainability Initiatives Others
Technology and Innovation	Digital Transformation Innovation Initiatives Intellectual Property Research and Development Technological Advancements Others

Panel D: Item IA (Risk Factor)	
Topic	Subtopic
Environmental and Social Risks	Climate Change Environmental Impact Ethical Practices Natural Disasters Resource Scarcity Social Responsibility Sustainability Others
External Factors	Cross-Border Operations Economic Conditions Geopolitical Risks Geopolitical Tensions Political Instability Public Health Crises Others
Financial Risks	Cash Flow Cost Management Credit Risk Debt Management Indebtedness Interest Rates Investment Strategies Liquidity Liquidity Risk Others
Governance Risks	Board Influence Board Policies Governance Practices Internal Controls Management Retention Ownership Structure Shareholder Rights Others
Human Resources	Employee Retention Key Personnel Labor Costs Labor Relations Leadership Succession Talent Acquisition Talent Retention Workplace Safety Others
Market Risks	Commodity Prices Competition Consumer Demand Currency Exchange Economic Conditions Interest Rates Market Volatility Price Volatility Seasonality Stock Price Volatility Others

Continued on next page

Panel D: Item IA (Risk Factor)	*(continued)*
Topic	Subtopic
Operational Risks	Contract Management Data Security Operational Efficiency Production Costs Production Delays Project Management Quality Control Service Disruptions Service Reliability Supply Chain Technology Integration Weather Impact Workforce Management Others
Regulatory and Legal Risks	Compliance Contractual Obligations Data Privacy Environmental Regulations Intellectual Property Legal Proceedings Litigation Product Liability Regulatory Investigations Tax Compliance Trade Policies Others
Reputational Risks	Brand Image Brand Value Corporate Governance Public Perception Social Media Impact Stakeholder Relations Others
Strategic Risks	Acquisition Strategy Business Integration Business Strategy Capital Allocation Competitive Position Growth Strategy Innovation Market Expansion Mergers and Acquisitions Strategic Alliances Others
Technology and Cybersecurity	Cyber Threats Cybersecurity Data Protection Digital Transformation Innovation Challenges System Failures Technological Advancements Technology Adoption Others

Figure I. Model Architecture

This figure illustrates our model architecture. In the first phase, we obtain embeddings of each paragraph with a document using OpenAI's text-embedding-3-large model. We then use the first 64 elements of the OpenAI's embeddings and construct a document-level matrix ($n \times 64$, n is the number of paragraphs in each document). We then apply the self-attention algorithm from the Transformer architecture to the matrix. The model learns the weight matrices of query, key, and value. This phase generates another n $\times$ 64 matrix, incorporating intra-document relations among paragraphs. We then apply the second custom attention layer, transforming the matrix into a 1×64 vector. In this phase, the model learns a custom weight vector (1×64). Finally, we build a multi-layer perceptron model using the document-level vector that incorporates each paragraph's relative importance as an input. The feedforward neural network contains two hidden layers, each with 128 and 64 neurons. The output layer has one neuron. We use ReLU activation function to impose non-linearity among neurons. For model training details, please refer to Section II.C.

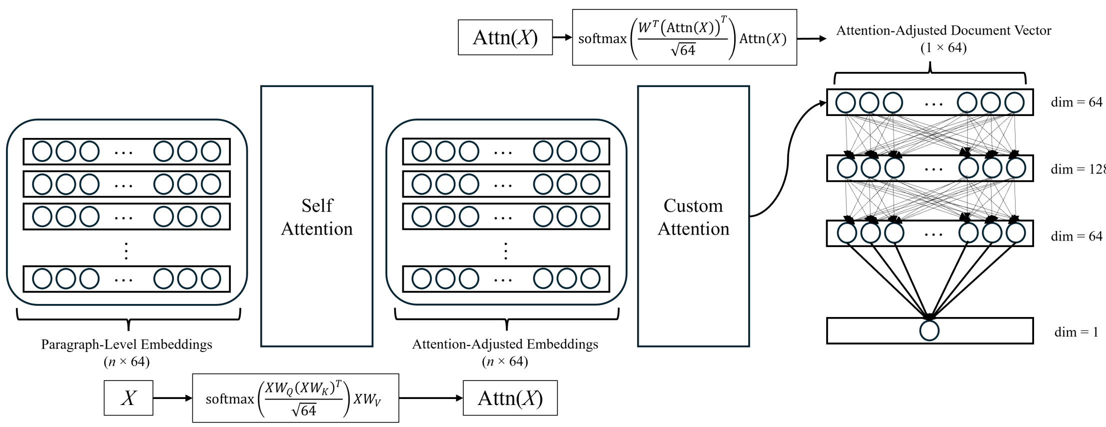


FIGURE II. Distribution in Paragraph-Level Importance

This figure presents the distribution of importance at the paragraph level. The x-axis represents the similarity values between the original OpenAI embedding and the output from the custom attention layer, and the y-axis denotes the frequency of these values.

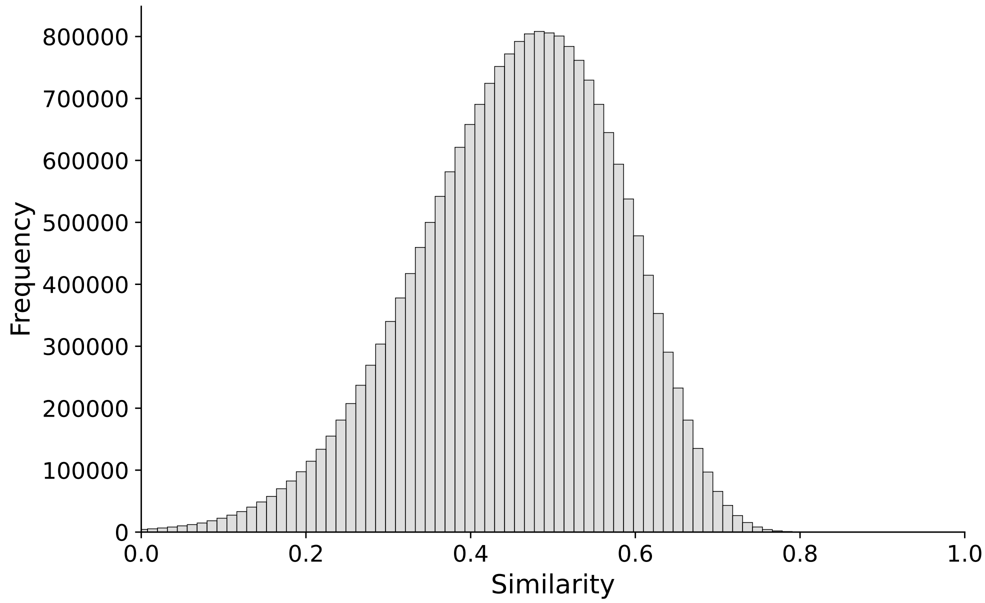


FIGURE III. Time-Series of Item-level Importance

This figure shows the time trend of importance at the item level. We use paragraph-level importance averaged at the item level and plot its annual average across documents. We use five-year moving averages.

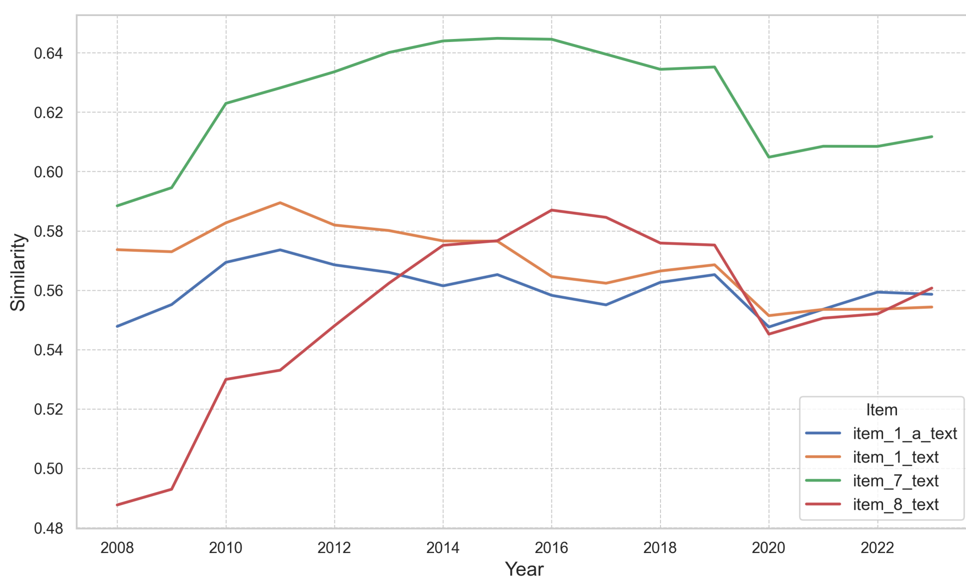
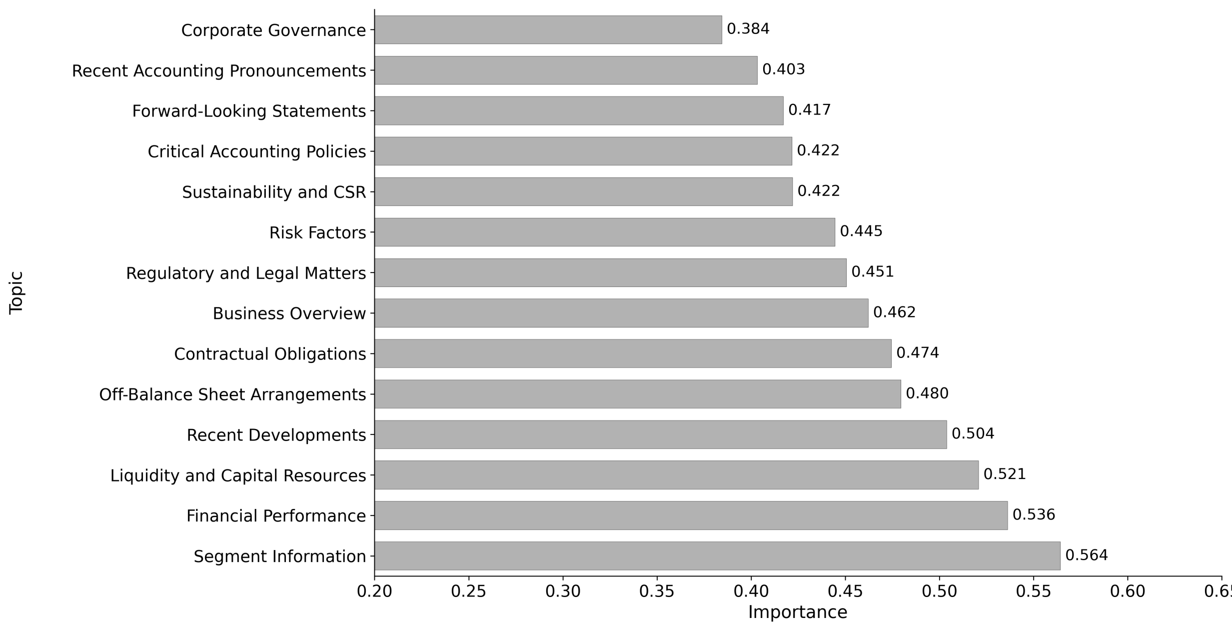


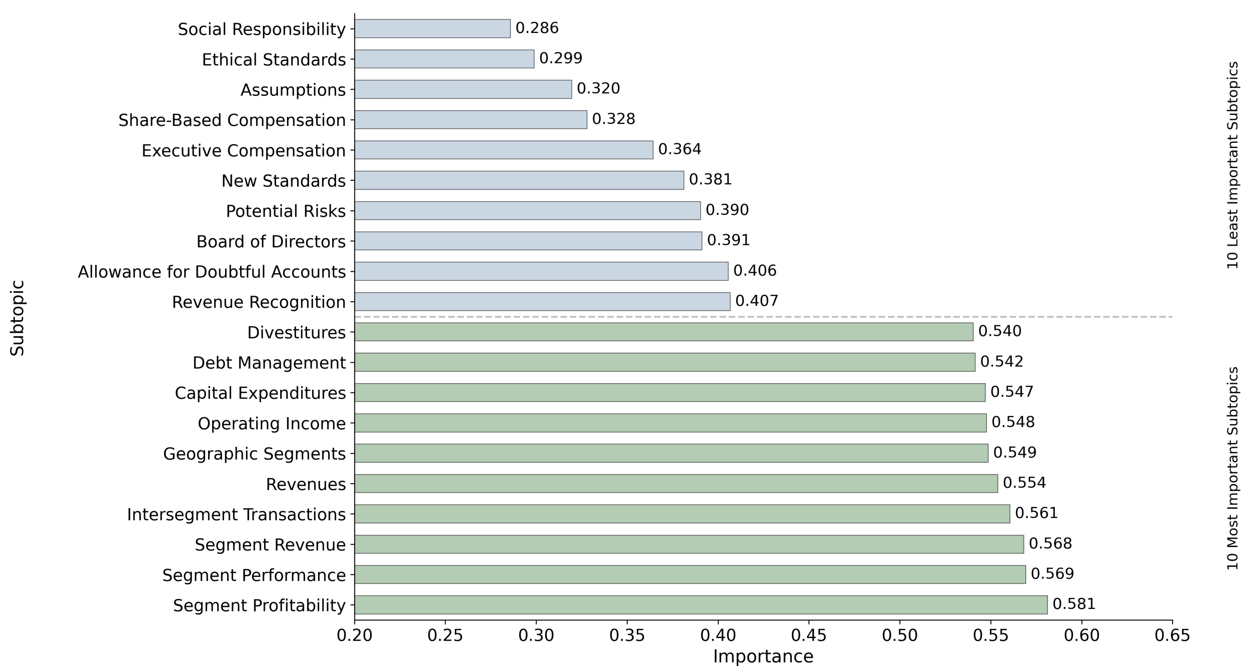
Figure IV. Topic Importance

This figure presents the topic- and subtopic-level importance across different Items of 10-K filings. Topic-level importance is computed as the average similarity score of all paragraphs classified under each topic, while subtopic-level importance follows the same computation within more granular subtopic classifications. For each Item (7, 8, 1, and 1A), we present both topic-level importance (Panels A, C, E, and G) and their corresponding subtopic-level importance (Panels B, D, F, and H).

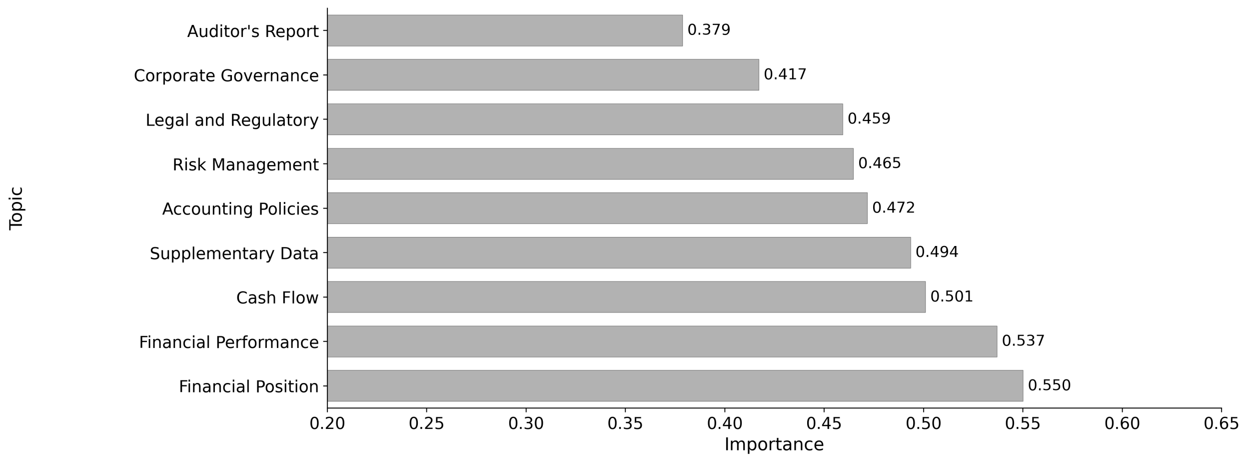
Panel A: Item 7 (MD&A) Topic Importance



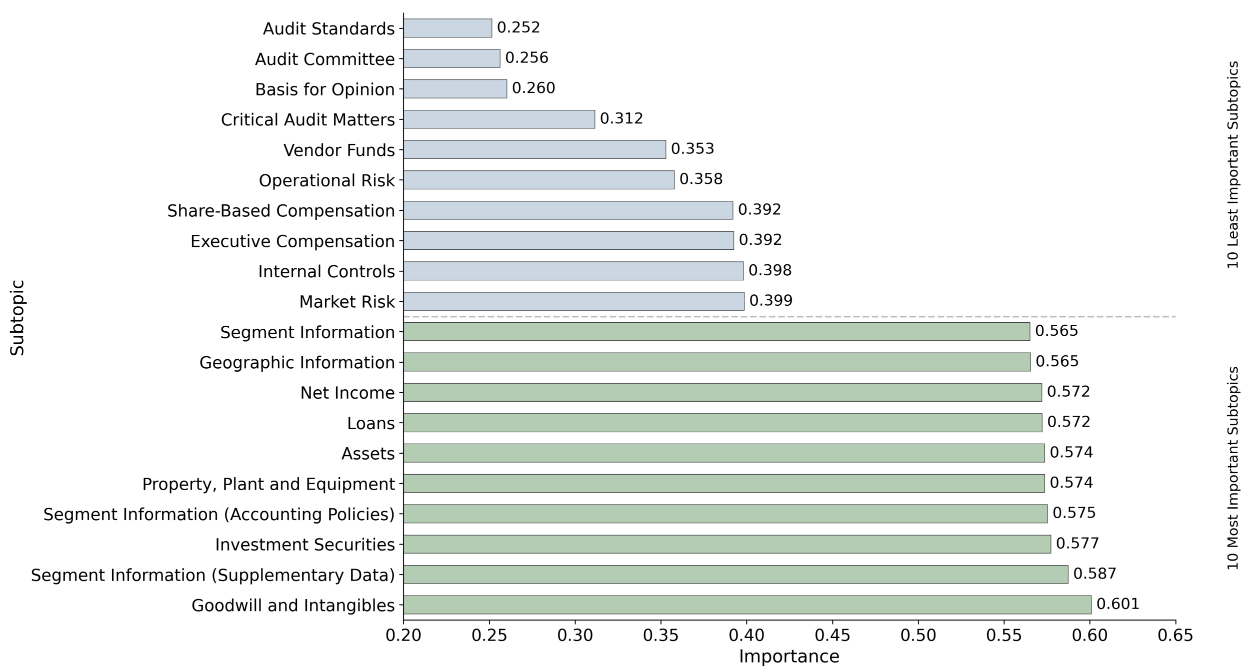
Panel B. Item 7 (MD&A) Subtopic Importance



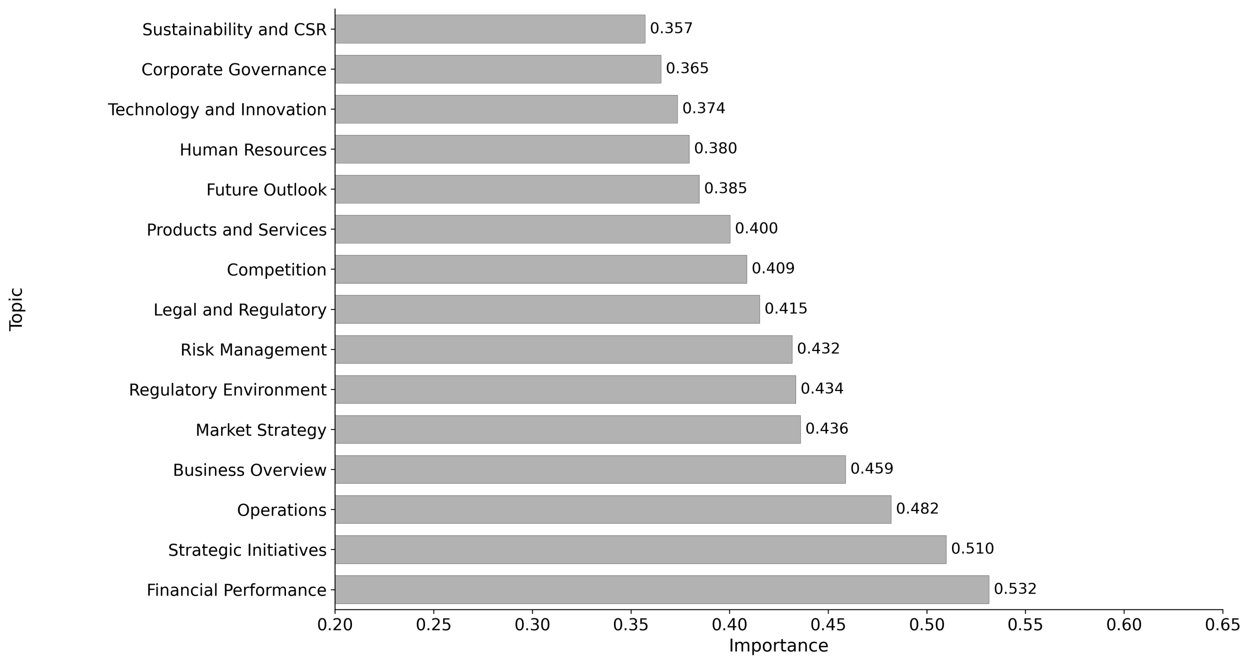
Panel C. Item 8 (Financial Statements and Supplementary Data) Topic Importance



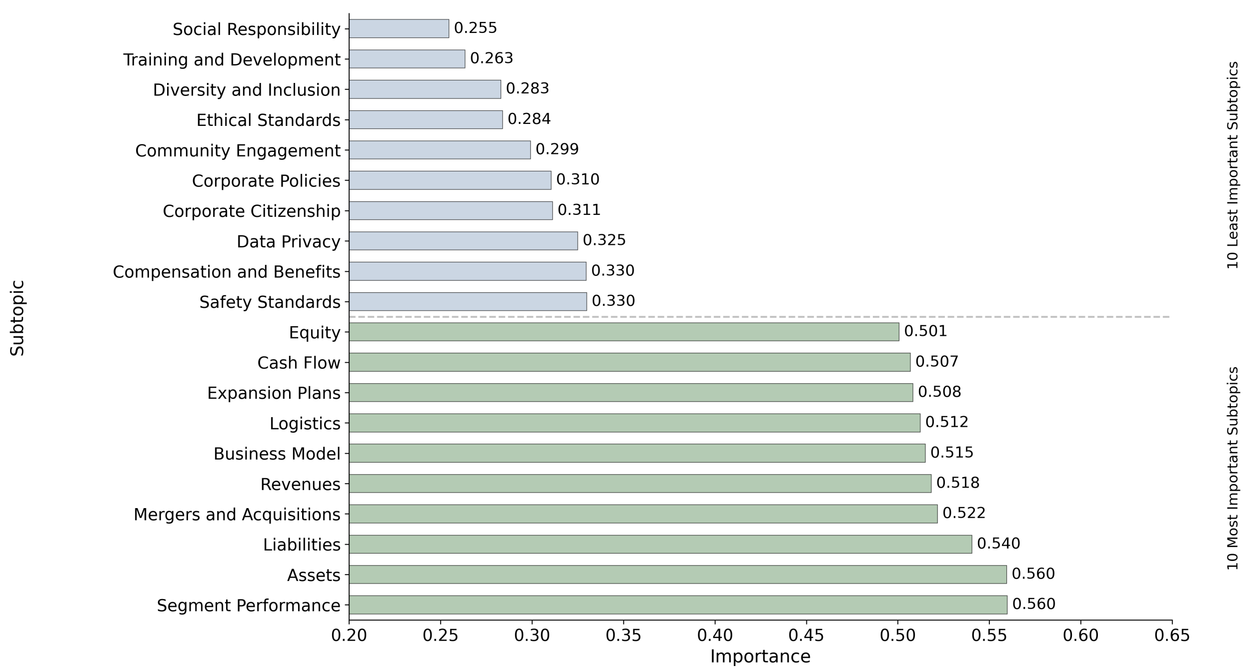
Panel D. Item 8 (Financial Statements and Supplementary Data) Subtopic Importance



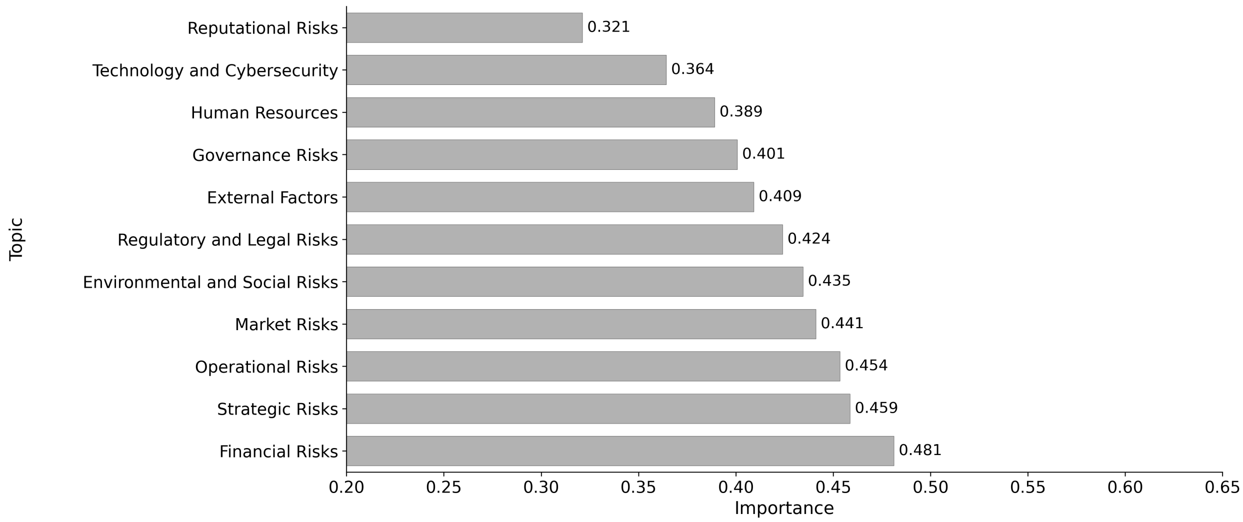
Panel E. Item 1 (Business) Topic Importance



Panel F. Item 1 (Business) Subtopic Importance



Panel G. Item 1A (Risk Factor) Topic Importance



Panel H. Item 1A (Risk Factor) Subtopic Importance

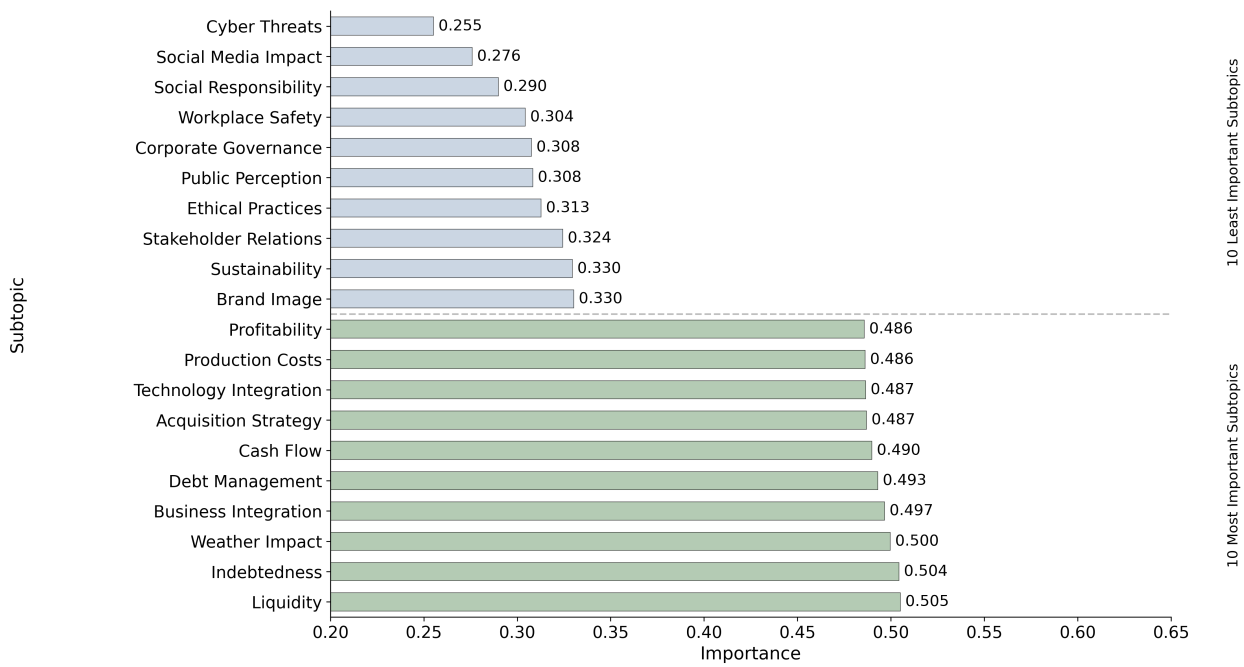


Table I. Baseline Model Performance

This table presents the relative performance of the machine learning models. We first obtain OpenAI embeddings of each paragraph in 10-K filings and reduce their dimensionality from 3,072 to 64. The target variable is a binary indicator that equals one when cumulative returns from one day before the filing date to 30 days after the filing date are positive and zero otherwise. For the Logit and MLP models, we obtain an averaged embedding of all paragraphs in each 10-K filing. We regress the positive return indicator on each embedding element for the Logit model. For the MLP model, we include two hidden layers with 128 neurons and 64 neurons, respectively. The last layer has one neuron. Please refer to Section 2 for detailed model descriptions for the attention-based model. We use rolling windows of the previous six years to train each model. For the MLP model and attention-based model, we use an additional two years as validation samples. We report accuracy and area under the curve as performance metrics. In addition to the AUC and accuracy metrics, we compare the performance of the attention-based and the other two models. Specifically, we perform a paired t-test to compare the average values. *, **, and *** denote statistical significance under the 10%, 5%, and 1% levels, respectively.

Model	Attention-based		MLP		Logit	
Year	AUC (1)	Accuracy (2)	AUC (3)	Accuracy (4)	AUC (5)	Accuracy (6)
2004	0.5133	0.5115	0.5079	0.4941	0.5100	0.4872
2005	0.5242	0.4478	0.5137	0.4778	0.5204	0.4207
2006	0.5336	0.5335	0.5031	0.4938	0.5266	0.5438
2007	0.5407	0.5340	0.5346	0.5329	0.5112	0.5208
2008	0.5493	0.5186	0.5389	0.5320	0.5179	0.5069
2009	0.5428	0.6897	0.5247	0.4942	0.5108	0.4897
2010	0.5806	0.7296	0.5522	0.6867	0.5088	0.7227
2011	0.5434	0.5796	0.5273	0.5687	0.5017	0.5659
2012	0.5709	0.5138	0.5142	0.5265	0.5094	0.5192
2013	0.5792	0.6140	0.5503	0.5894	0.5181	0.6122
2014	0.6163	0.5147	0.5560	0.5012	0.5250	0.4885
2015	0.5662	0.5415	0.5500	0.5587	0.5416	0.5574
2016	0.6012	0.5276	0.5964	0.5127	0.5722	0.6142
2017	0.5388	0.5166	0.5108	0.5038	0.5284	0.5246
2018	0.5492	0.5127	0.5294	0.5181	0.5363	0.5290
2019	0.5782	0.5511	0.5662	0.5482	0.5400	0.5544
2020	0.4616	0.4412	0.3967	0.3738	0.3834	0.3247
2021	0.6693	0.6183	0.5951	0.5375	0.5636	0.5571
2022	0.5981	0.5746	0.5827	0.5898	0.5604	0.5833
2023	0.5417	0.6691	0.5062	0.5836	0.5083	0.6365
Average	0.5599	0.5570	0.5328	0.5312	0.5197	0.5379

Comparisons

	Difference	t-value
Difference in AUC: (1) vs. (3)	0.0271***	5.76
Difference in AUC: (1) vs. (5)	0.0402***	6.13
Difference in Accuracy: (2) vs. (4)	0.0258**	2.24
Difference in Accuracy: (2) vs. (6)	0.0191	1.49

Table II. Predicting Future Returns and Changes in Earnings

This table applies the attention algorithm to other economic target variables. In Panel A, we use the direction of future stock returns as the target. Future stock returns are defined as cumulative returns measured over a period of one day after the filing date until 30 days after the filing date. The target variable takes a value of one when future stock returns are positive and zero otherwise. In Panel B, we use an indicator that equals one when the next period's EPS is higher than that of this period and zero otherwise. In addition to the AUC and accuracy metrics, we compare the performance of the attention-based and the other two models. Specifically, we perform a paired t-test to compare the average values. *, **, and *** denote statistical significance under the 10%, 5%, and 1% levels, respectively. We also report the portfolios' Sharpe ratios in Panel C using the predicted values. For future return predictions, we choose only the observations whose filing dates are in February and March. For the observations whose filing dates are between February 1 and 14, we form a portfolio on February 15 and hold the portfolio for two weeks. For the observations whose filing dates are between February 15 and the last day of February, we form the portfolio on March 1 and hold it for two weeks. For the observations whose filing dates are between March 1 and March 14, we form the portfolio on March 15, and for those between March 15 and March 31, we form the portfolio on April 1. We take the long position of the stocks whose predicted values are in the top 20% and short the stocks whose predicted values are in the bottom 20%. All portfolios are equal-weighted. For EPS predictions, we form portfolios on June 30th of each year and hold the portfolio for a year. We take the long position of the stocks whose predicted values are in the top 20% and short the stocks whose predicted values are in the bottom 20%. All portfolios are equal-weighted.

Panel A: Directions in Future Returns

	Attention-based		MLP		Logit	
Year	AUC (1)	Accuracy (2)	AUC (3)	Accuracy (4)	AUC (5)	Accuracy (6)
2004	0.5140	0.4990	0.5132	0.4897	0.5071	0.4841
2005	0.5118	0.5131	0.4988	0.4082	0.5086	0.4058
2006	0.5409	0.5208	0.5291	0.5284	0.5331	0.5448
2007	0.5351	0.5199	0.5149	0.5187	0.5082	0.5190
2008	0.5589	0.5219	0.5458	0.5287	0.5203	0.5191
2009	0.5554	0.6937	0.5274	0.5277	0.5187	0.5401
2010	0.5845	0.7290	0.5620	0.7378	0.5102	0.7381
2011	0.5387	0.5688	0.5287	0.5744	0.5002	0.5744
2012	0.5617	0.5141	0.4785	0.5165	0.5030	0.5178
2013	0.5774	0.6016	0.5643	0.6104	0.5186	0.6072
2014	0.6126	0.5173	0.6084	0.4729	0.5248	0.4763
2015	0.5464	0.5330	0.5323	0.5441	0.5268	0.5413
2016	0.6209	0.6135	0.6001	0.6043	0.5695	0.6111
2017	0.5401	0.4927	0.5212	0.5108	0.5226	0.5152
2018	0.5412	0.5311	0.5423	0.5311	0.5367	0.5344
2019	0.5971	0.5585	0.5806	0.5604	0.5500	0.5625
2020	0.4687	0.4092	0.3278	0.3322	0.3807	0.3336
2021	0.6582	0.5651	0.6198	0.5830	0.5723	0.5633
2022	0.5829	0.5686	0.5983	0.5706	0.5671	0.5818
2023	0.5535	0.6510	0.4717	0.6043	0.5092	0.6377
Average	0.5600	0.5561	0.5333	0.5377	0.5194	0.5404

Comparisons

	Difference	t-value
Difference in AUC: (1) vs. (3)	0.0267***	3.32
Difference in AUC: (1) vs. (5)	0.0406***	6.42
Difference in Accuracy: (2) vs. (4)	0.0184*	1.73
Difference in Accuracy: (2) vs. (6)	0.0157	1.54

Panel B: Directions in EPS Change

	Attention-based		MLP		Logit	
Year	AUC (1)	Accuracy (2)	AUC (3)	Accuracy (4)	AUC (5)	Accuracy (6)
2004	0.5072	0.5508	0.4711	0.5756	0.4980	0.5931
2005	0.5339	0.5588	0.5207	0.5822	0.5029	0.5825
2006	0.5150	0.5552	0.4937	0.5562	0.4980	0.5577
2007	0.5235	0.5131	0.4879	0.5093	0.4968	0.5095
2008	0.5473	0.5045	0.5121	0.4000	0.5086	0.4038
2009	0.5271	0.5338	0.5005	0.4965	0.4937	0.4875
2010	0.5458	0.6427	0.5254	0.5538	0.5192	0.5641
2011	0.5464	0.5911	0.4942	0.5411	0.4912	0.5587
2012	0.5531	0.5599	0.5575	0.5420	0.4863	0.5312
2013	0.5445	0.5745	0.5373	0.5510	0.4942	0.5491
2014	0.5616	0.5809	0.5603	0.5592	0.5066	0.5613
2015	0.5914	0.5332	0.5867	0.5373	0.5177	0.5326
2016	0.5460	0.5787	0.5421	0.5666	0.5217	0.5577
2017	0.5376	0.5546	0.5077	0.5153	0.4964	0.5331
2018	0.5641	0.5879	0.5559	0.5771	0.5194	0.5793
2019	0.5682	0.5301	0.5738	0.5277	0.5284	0.5353
2020	0.5376	0.5350	0.4521	0.4657	0.4891	0.4809
2021	0.6218	0.6878	0.5968	0.6407	0.5194	0.5903
2022	0.5563	0.4987	0.5492	0.5061	0.5049	0.4861
2023	0.5336	0.5442	0.4364	0.4664	0.4702	0.4614
Average	0.5481	0.5608	0.5231	0.5334	0.5031	0.5328

Comparisons

	Difference	t-value
Difference in AUC: (1) vs. (3)	0.0250***	4.08
Difference in AUC: (1) vs. (5)	0.0450***	9.34
Difference in Accuracy: (2) vs. (4)	0.0274***	3.34
Difference in Accuracy: (2) vs. (6)	0.0280***	3.23

Panel C. Sharpe Ratios

Method	Using Future Return Predictions (1)	Using EPS Change Predictions (2)
Attention-Based	1.56	1.40
MLP	1.28	1.13
Logit	1.08	1.06

Table III. Descriptive Statistics of Importance

This table presents descriptive statistics of paragraph-level importance. Paragraph-level importance is computed as the similarity between the original paragraph-level OpenAI embedding and the output vector of the custom attention layer. The output vector of the custom attention layer is a document-level vector that incorporates the semantic and contextual relations among the paragraphs. Panel A demonstrates several descriptive statistics. In Panel B, we include the variance decomposition results by reporting the R-squared values from regressing paragraph-level importance on firm-fixed effects, time-fixed effects, firm-time fixed effects, item-fixed effects, and firm-time-item fixed effects.

Panel A. Descriptive Statistics

level	N	Mean	Std	P25	P50	P75
paragraph	20,712,462	0.455	0.122	0.377	0.465	0.544

Panel B. Variance Decomposition

Fixed Effects (1)	R-Squared (2)
Firm FEs	8.41%
Year FEs	5.65%
Firm-Year FEs	18.23%
Item FEs	8.49%
Firm-Year-Item FEs	31.51%

Table IV. Item-level Importance

This table reports averaged item-level importance within each 10-K. We first choose the top five paragraphs within each item based on the paragraph-level importance. We then measure the item-level importance using the average paragraph-level importance of these top five paragraphs within each item. In column (2), we report the average importance values. In column (4), we report the ranking using the normalized importance values within each document.

Not Normalized		Normalized within Document	
Item (1)	Importance (2)	Item (3)	Importance (4)
Item 7 - Management's Discussion & Analysis	0.6281	Item 7 - Management's Discussion & Analysis	0.0060
Item 8 - Financial Statements	0.6142	Item 1 - Business	0.0058
Item 1 - Business	0.5704	Item 8 - Financial Statements	0.0052
Item 1A - Risk Factors	0.5679	Item 1A - Risk Factors	0.0051
Item 2 - Properties	0.4741	Item 2 - Properties	0.0047
Item 7A - Market Risk	0.4737	Item 4 - Mine Safety Disclosures	0.0044
Item 5 - Market for Securities	0.4474	Item 6 - Selected Financial Data	0.0042
Item 6 - Selected Financial Data	0.4452	Item 9 - Accounting Disagreements	0.0042
Item 1B - Unresolved Comments	0.4053	Item 5 - Market for Securities	0.0041
Item 14 - Accountant Fees & Services	0.3981	Item 7A - Market Risk	0.0041
Item 4 - Mine Safety Disclosures	0.3977	Item 14 - Accountant Fees & Services	0.0040
Item 11 - Executive Compensation	0.3945	Item 1B - Unresolved Comments	0.0038
Item 9A - Controls & Procedures	0.3907	Item 11 - Executive Compensation	0.0038
Item 9B - Other Information	0.3803	Item 9A - Controls & Procedures	0.0037
Item 3 - Legal Proceedings	0.3684	Item 3 - Legal Proceedings	0.0037
Item 13 - Relations & Transactions	0.3642	Item 12 - Ownership	0.0036
Item 12 - Ownership	0.3641	Item 10 - Directors & Governance	0.0036
Item 10 - Directors & Governance	0.3542	Item 13 - Relations & Transactions	0.0035
Item 9 - Accounting Disagreements	0.3524	Item 9B - Other Information	0.0035

Table V. Variance Decomposition Using Topic Classifications

In this table, we perform variance decomposition of paragraph-level importance scores at the item level. We regress paragraph-level importance scores on various sets of fixed effects. Within each item, we include firm and year fixed effects separately, and both fixed effects together in one regression. We then include topic-subtopic classifications as fixed effects. Finally we include firm, year, and classification fixed effects simultaneously in one regression. We reporte R-squared values for each specification.

Fixed Effects	Item 7 (1)	Item 8 (2)	Item 1 (3)	Item 1a (4)
Firm FEs	10.80%	7.61%	17.81%	14.89%
Year FEs	12.85%	10.11%	7.26%	9.29%
Firm and Year FEs	22.10%	16.86%	23.94%	23.39%
Topic-Subtopic FEs	21.82%	24.95%	16.06%	17.01%
Firm, Year, and Topic-Subtopic FEs	41.14%	39.83%	34.28%	37.55%

Table VI. Disclosure Regulations and Item-Level Importance

In this table, we examine whether disclosure regulations by the SEC impact the importance of the targeted items. We use the SEC's Modernization Act, enforced on August 09, 2021, for MD&As. In the Act, the SEC refocused the MD&A on providing a clear narrative that allows investors to understand the company's financial condition and results through management's eyes. We construct our sample using observations whose filing dates are between August 09, 2021, and August 09, 2023. The sample is a firm-item-year-level dataset containing item-level importance of Item 7 (MD&A) and Item 8 (Financial Statements and Notes). We use Item 8 as a control group as there was no material change in disclosure regulations related to the notes of financial statements around August 2021. The dependent variable is the item-level importance, the averaged paragraph-level importance within each item. Treat is an indicator that equals one when the item is MD&A and zero otherwise. Post is an indicator that equals one when the filing date is after August 9, 2021, and zero otherwise. We include firm-year fixed effects in column (1) and firm and year fixed effects in column (2). Standard errors are clustered at the firm level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable	Importance	
	(1)	(2)
Treat (=Item 7)	0.0224***	0.0225***
	(28.34)	(28.34)
Post (=After Aug 09, 2021)		-0.0035
		(-1.20)
Treat × Post	0.0052***	0.0058***
	(5.30)	(5.67)
Average Dependent Variable	0.5346	0.5346
Standard Deviation of Dependent Variable	0.0573	0.0573
Firm-Year Fixed Effects	Yes	No
Firm Fixed Effects	No	Yes
Year Fixed Effects	No	Yes
Adjusted R^2	0.6758	0.4083
N	14,830	14,830

Table VII. SEC Comment Letters and Disclosure Response

In this table, we examine whether the SEC's comment letters affect the quality of narrative disclosures. We collect the entire universe of the SEC's comment letters on Regulation S-K from 2005 and match them with the corresponding items in the 10-K filings. We then construct an item-firm-year-level dataset and merge it with the comment letter database. We retain the items pointed out by the SEC comment letters and the same item of the same firm one year afterward. The observations one year afterward are flagged as Post = 1. We then regress item-level importance on Post with firm-item-fixed effects and year-fixed effects. Standard errors are clustered at the firm level. t-statistics are reported within parentheses. *, **, and ***, denote statistical significance at the 10%, 5%, and 1% level, respectively.

Dependent Variable =	Item-Level Importance	
	(1)	(2)
Post (= One Year After Receiving the Comment Letter)	0.0142**	0.0098***
	(8.33)	(2.98)
Average Dependent Variable	0.3915	0.3915
Standard Deviation of Dependent Variable	0.0984	0.0984
Firm-Item Fixed Effects	Yes	Yes
Year Fixed Effects	No	Yes
Adjusted R^2	0.7734	0.7988
N	31,533	31,533

Table VIII. Positional Importance and Information Environment

This table examines the relation between paragraph-level importance and its position within an MD&A. The dependent variable is the position of each paragraph within an MD&A, with the first paragraph having a score of 1 and the last paragraph having a score of 0. Sentiment is measured using the keyword dictionary of Loughran and McDonald (2011), competition is measured using the approach of Li et al. (2013), profitability follows the definition of Ball et al. (2016), and earnings volatility is the standard deviation of earnings-per-share values over the past five years. We include firm-year fixed effects in column (1) and do not include fixed effects in the remaining columns, as interacting variables are measured at the MD&A level. Standard errors are clustered at the firm level. t-statistics are reported within parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable =	Position within the MD&A Section					
Cross-Section	Full (1)	Sentiment (2)	Competition (3)	Profitability (4)	EarnVol (5)	Pooled (6)
Importance	0.3027***	0.2468***	0.2195***	0.2235***	0.2033***	0.2494
	(88.59)	(57.10)	(70.40)	(53.35)	(57.35)	(54.69)
Sentiment		-0.0425***				-0.0398***
		(-7.44)				(-6.94)
Importance × Sentiment		0.1029***				0.0983***
		(8.86)				(8.42)
Competition			0.0852***			0.0769***
			(4.08)			(3.64)
Importance × Competition			-0.0841*			-0.0662*
			(-1.87)			(-1.70)
Profitability				-0.0843***		-0.0755***
				(-10.33)		(-8.45)
Importance × Profitability				0.2831***		0.2535***
				(6.76)		(5.97)
EarnVol					0.0037***	0.0032***
					(9.11)	(7.85)
Importance × Earnvol					-0.0959***	-0.0842***
					(-3.35)	(-3.03)
Firm-Year Fixed Effect	Yes	No	No	No	No	No
Cluster	Firm	Firm	Firm	Firm	Firm	Firm
Adjusted R^2	0.0698	0.0875	0.0743	0.0755	0.0902	0.1003
N	5,170,592	5,170,592	5,170,592	5,170,592	5,170,592	5,170,592