From Transcripts to Insights: Uncovering Corporate Risks Using Generative AI

Alex G. Kim*   Maximilian Muhn†   Valeri V. Nikolaev‡

First draft: October 5, 2023
This draft: October 26, 2024

Abstract

We explore the value of generative AI in uncovering corporate risks. We use a Large Language Model (LLM) to construct firm-level measures of political, climate, and AI-related risks. The proposed measures possess significant information content and outperform dictionary-based measures in predicting firm-level volatility and explaining investment decisions. Information in risk *assessments* dominates that in risk *summaries*, establishing the value of general AI knowledge. Our measures continue to perform well outside the LLM's training window and at the aggregate level. They are also priced in equity markets. Collectively, generative AI offers an effective solution to the measurement of hard-to-quantify corporate risks.

Keywords: GPT, ChatGPT, large language models, generative AI, risk information, firm-level risk exposure, conference call, political risk, AI risk, climate change risk

JEL Codes: C45, D81, G12, G30, G32, M41

CAVEAT NOTICE:

The findings presented here are preliminary and subject to revision. We are currently reevaluating our data, methods, and conclusions.

1 Introduction

In the era of global political instability, climate uncertainty, and rapid technological change, corporations face various hard-to-assess risks that extend beyond traditional financial metrics. New challenges like shifting regulations, environmental issues, and AI-related risks are emerging quickly, each impacting long-term growth and stakeholder interests significantly. This study explores how large language models (LLMs) can detect and analyze these critical aspects of corporate risk. Specifically, we aim to understand the capabilities of AI in assessing the complex corporate risk landscape and, ultimately, helping stakeholders to make more informed decisions in the face of growing uncertainties. Such evaluation is particularly timely as we are witnessing the emergence of financial services that use generative AI to query, analyze, and summarize corporate disclosures.

The evaluation of firm risks through textual analysis of corporate disclosures received substantial attention in recent literature (see in particular Hassan et al., 2019, 2024; Chava et al., 2022; Sautner et al., 2023a,b; Li et al., 2024). A distinctive feature of these studies is the utilization of dictionary-based bigram (n-gram) frequencies to quantify various risk types.[1] This literature has laid important groundwork that improves our understanding of corporate risks. It also shows that corporate disclosures are a valuable information source to learn about firms' risk exposure (e.g., Hassan et al., 2019). However, the recent developments in AI technology offer a valuable opportunity to explore textual data more deeply and gain a richer, more nuanced insight into corporate risks. The new generation of language models can understand complex relationships within a text, incorporate a wide context within which relevant topics (words) are discussed, and make inferences. These aspects are critical for an accurate and comprehensive analysis of complex corporate risks. Furthermore, rapidly evolving political, environmental, and technological changes in recent years may quickly render existing dictionaries outdated or incomplete.

---

[1] This approach counts the presence of risk-related words mentioned in the vicinity of risk-topic-specific bigrams from a pre-constructed dictionary (e.g., the algorithm searches for instances where the bigram "economic policy" is used along with the word "risk").

Two additional features make generative language models particularly attractive in the analysis of corporate risks. First, their general nature allows them to go beyond a given text. Unlike traditional methods that analyze risks based on a single document, such as a conference call transcript, generative language models are trained on vast corpora that enable them to leverage *general knowledge* acquired from similar documents or documents featuring related topics. This "general AI" feature holds promise in improving the measurement of corporate risks because companies need not explicitly discuss risks in their disclosures.[2] Second, an important advantage of the new generation of language models is that they synthesize the extracted information into coherent, understandable narratives, thus providing not only a quantitative assessment but also an explanation to support it.

To illustrate the usefulness of the AI-based approach, consider the example of SK Telecom Inc., a Korean cell phone service provider cross-listed and operating in the US. Figure 1 illustrates a significant wedge between the bigram-based measure of political risk, which implies no risk in 2018, and our GPT-based measure, which places the company towards the top of risk distribution. What is causing this difference? Around 2018, an active discussion took place worldwide regarding regulating the bundling of phone handset contracts and telecom services.[3] During SK Telecom's 2018 earnings call, analysts actively asked questions concerning the bundled sales. The transcript, however, does not explicitly discuss political or regulatory risks; hence, bigrams do not capture these risks. In contrast, GPT identifies "*the separation of handsets and telecom services*" as a potential source of regulatory uncertainties and provides a rich explanation of this issue (see Appendix A). The example illustrates that LLMs connect diverse pieces of information in a given text and make judgments by interacting context with the model's general knowledge.

Despite this potential, LLMs' ability to evaluate firm-level risks is yet to be understood. This study addresses this question by developing firm-level risk exposure measures using

---

[2]For example, political risk is implicitly implied if an executive states that a firm is subject to a new regulation.

[3]Indeed, Japan banned the bundling in 2019. See these two articles: (link 1) and (link 2). Such bundling is considered detrimental for consumers since it makes pricing opaque (Hazlett et al., 2018).

OpenAI's GPT3.5-Turbo LLM. In particular, we extract the relevant risk-related information from earnings conference call transcripts spanning January 2018 to March 2023. We then evaluate how GPT-based risk exposure measures compare to the existing measures in predicting stock market volatility and related economic outcomes.

We focus on the sources of corporate risks of the highest significance to firms' stakeholders: political risk (Hassan et al., 2019), climate-related risk (Sautner et al., 2023a), and AI-related risk.[4] Because AI risk is a new phenomenon, it enables us to probe the ability of language models to assess emerging risks.[5] To further explore the value of LLMs' knowledge, for each of the three measures, we generate two types of output: (1) risk *summaries* and (2) risk *assessments*, as discussed next.

Risk summaries aim to identify risks solely based on the earnings call contents. Accordingly, we provide the model with instructions that minimize GPT's reliance on its general knowledge. The risk summaries are thus human-readable reorganizations of risk-related discussions in conference call transcripts. In contrast, the goal of risk assessments is to utilize the ability of the LLM to integrate the discussion in the call with its general knowledge and make judgments. We develop a prompt to generate an assessment of a given risk that is not limited to the contents of a given transcript. Therefore, each assessment integrates the document's context with the model's broader understanding.

There are a number of possible approaches to converting readable summaries (assessments) into quantitative risk exposures. One could, for example, rank the GPT output by using it (its embeddings) in a machine learning model trained to construct risk scores

---

[4]To substantiate our focus on the three key categories of risks, we randomly sample 100 conference calls after November 2021. For each sentence of the transcripts, we instruct GPT-4-Turbo to identify whether the sentence discusses financial or non-financial risks. Among sentences that discuss non-financial risks, we further instruct the model to determine the topic of the sentences into one of the four categories: political risk, climate change risk, AI-related risk, and other non-financial risks. Among all sentences, 6.55% contain content related to financial risk, and 7.56% contain content related to non-financial risk. Among sentences that contain information about non-financial corporate risk, 31.08% of sentences are about political risks, 21.56% are about climate change risks, 24.60% are about AI-related risks, and 22.76% are about other non-financial risks.

[5]A complementary approach to assessing firms' exposure to AI risk, rather than analyzing firm communications, involves examining labor inputs likely to be disrupted by this new technology (Eisfeldt et al., 2024).

predictive of future volatility. While such approaches are likely to add further value, we follow perhaps the most straightforward approach that divides the length of a risk summary (assessment) by the length of the underlying transcript. An important advantage of this approach is that our measures are more directly comparable to the bigram-based proxies, which also rely on risk-related word counts scaled by the transcript length.[6]

Our measures show intuitive variation. The tobacco industry exhibits the highest political risk, coal mining exhibits the highest climate risk, and business services exhibit the highest AI risk exposures. The aggregate time series of our political and climate risks covary positively with the corresponding bigram-based exposure measures, culminating in the first quarter of 2020.[7,8] AI-related risks exhibit a different pattern. Instead of an uptick around 2020, we observe a sharp increase in the most recent years, consistent with the emergence of AI technologies.

Variance decomposition analysis shows that time- and industry-time-invariant characteristics can only explain 10-15% of the total variation across the three risk measures. In line with Hassan et al. (2019), about 90% of variance in political risk is firm-specific. This number goes down somewhat (is similar) for climate risk (AI risk) proxies. We further decompose the residual variation and observe that firm-invariant factors explain only 20-30%. Thus, time-varying firm-specific components account for the bulk of the variation in firms' exposures to political, climate, and AI risks.

Our main analysis uses a market-based approach to evaluate whether GPT-based proxies are effective at measuring firm-level risks. Specifically, we examine whether risk exposure measures explain future stock price volatility (Engle, 2004). We use two forward-looking firm-level volatility metrics: implied volatility derived from option prices (Hassan et al.,

---

[6]Higher ratios are interpreted as higher risk exposure.

[7]This spike is likely to be attributable to the presidential election of the US and the outbreak of COVID-19. As the Biden administration is known to put more emphasis on global environmental issues than the previous administration, we see an elevated level of climate change risk after the first quarter of 2020. Furthermore, in line with the Russia-Ukraine war, we also observe an increase in political risk (and also climate change risk) in the first quarter of 2022.

[8]For the AI-related risk exposure measures, there is no corresponding bigram-based measure.

2019; Sautner et al., 2023a), and abnormal realized volatility, which builds on Loughran and McDonald (2014).

We begin by examining political and climate risk exposures as they received the most attention in prior literature. Focusing on the main sample period from 2018 to 2021, we show a strong and robust positive relation between the GPT-based risk exposure measures and stock price volatility. Across different fixed effects structures, GPT-based political and climate risk measures exhibit positive associations with both of our volatility proxies. We also consistently find evidence that GPT-based proxies are more informative compared to the bigram-based proxies in explaining stock price volatility for these two types of risks, indicating the significant value derived from the new technology. In particular, GPT-based measures subsume information in bigram-based measures. More importantly, when comparing our GPT-based measures against each other, risk assessments perform better than risk summaries both in the case of political and climate risks. This result implies that AI general-knowledge-based insights are useful in uncovering corporate risks.

One concern with pre-trained LLMs is their potential for look-ahead bias (e.g. Sarkar and Vafa, 2024), where they may incorporate future event information learned during their training phase. Although we do not instruct the model to *predict* firm-level risks, we examine our risk proxies using a time period unseen by the language model during its training phase (pure out-of-sample period). Since GPT3.5 is trained on texts that precede September 2021, we focus on a sample of transcripts from January 2022 until March 2023 for these additional tests. Even within this limited sample, we find strong and robust evidence that GPT-based risk exposure measures are positively associated with volatility variables, suggesting that our results are unaffected by GPT's possible ex-post knowledge.

Turning to the analysis of AI-related risks, based on our main sample, we find limited evidence that AI risk is an important determinant of stock price volatility. It shows up only in one of the two models featuring risk assessment-based proxies. This result is not unexpected due to the recency of AI disruptions (Eisfeldt et al., 2024). Indeed, we show that

5

the AI-related risk proxy becomes significant in predicting stock price volatility after 2021.

Having established the validity of the GPT-based risk proxies, we turn our attention to firms' actions predicted to change after being exposed to a given risk. First, we investigate whether our risk measures can explain capital investments. In theory, riskier companies experience higher financing costs and value the option of waiting (Dixit and Pindyck, 1994). Thus, ceteris paribus, they are less likely to make investments. This force is expected to be pronounced for political and environmental risks. However, for technology-related risks, the effect is less clear because addressing AI challenges requires significant investments in new technology. Indeed, the political and (less so) climate-risk exposures exhibit negative associations with investment. On the other hand, AI risk exhibits a positive albeit insignificant relation, which becomes significant during the 2022-2023 period.

Furthermore, we find that firms further adjust their behavior in response to specific risks they are facing. They increase (1) lobbying activity in response to political risk, (2) green patent filings in response to climate risk, and (3) AI-related patent filings in response to AI risk. These findings continue to hold in the period outside of GPT's training window.

We also expand our analysis by constructing the economy-wide measure of political risk. We show that an LLM-based economy-wide measure of political risk has more information about future stock market volatility, investment, and unemployment as compared to a widely used keyword-based measure of the economic policy uncertainty (EPU) (Baker et al., 2016). The former is also a more forward-looking measure. Finally, we show that environmental and AI risks command significant equity risk premia when assessed based on the traditional asset pricing methodology.

We make the following contributions to the literature. First, we probe the economic usefulness of AI-powered large language models in risk assessment. Although the financial industry is currently at the forefront of the adoption of LLM-based technology, the economic usefulness of LLMs in risk assessment and risk management is yet to be understood. We add to a growing body of research on the potential value of LLMs (see Bernard et al., 2023;

Cao et al., 2023; Lopez-Lira and Tang, 2023; Jha et al., 2023; Chen et al., 2023; Eisfeldt et al., 2024; Kim et al., 2023, 2024), by showing that AI tools are effective at distilling disclosures to extract information about diverse risk categories. Recent work has shown that LLMs are capable of generating informative summaries of corporate disclosures (Kim et al., 2023), evaluating corporate investment strategies (Jha et al., 2023), innovation success (Yang, 2023), job substitutability and exposure to AI (Eisfeldt et al., 2024), as well as detecting private information communicated by fund managers (Cao et al., 2023). Additionally, our work contributes to the broader literature on machine learning applications in finance and accounting (e.g., Frankel et al., 2022; Chen et al., 2022; Cao et al., 2023, 2024).

Second, we contribute to the recent literature that uses corporate disclosures to construct firm-level measures of risk exposure: political risk (Hassan et al., 2019), country risk (Hassan et al., 2024), climate risk (Sautner et al., 2023a,b; Li et al., 2024), inflation risk (Chava et al., 2022), offshore activity exposure (Hoberg and Moon, 2019), and pandemic risk (Hassan et al., 2023). We complement and build on this influential work by adopting AI-based technology to analyze risks. In contrast to existing studies that rely on topic-based bigram dictionaries, LLMs are trained to understand the deeper context in which bigrams are encountered and thus have the potential to improve accuracy.[9] Indeed, we document that GPT-based measures are more informative than bigram-based measures and generally subsume their information content.

Last but not least, we contribute by establishing the value of AI knowledge acquired during pre-training to understand complex topics, such as corporate risk. We show that LLMs successfully leverage their general knowledge to derive *insights* about corporate risks from a given context. These insights go beyond the information discussed in the processed document. The AI-knowledge-based evaluations of risks, unrestricted to the document context, are incrementally informative and generally outperform summaries that are based on localized knowledge only.

---

[9]Chava et al. (2022) employ a deep-learning approach to classify the topic of each sentence and measure firm-level inflation exposure.

2 LLMs' Usefulness in Measuring Risk Exposure

The recent emergence of Large Language Models (LLMs) has fundamentally transformed our ability to analyze and generate text. Generative models, such as GPT (by OpenAI), LLaMa (by Meta), and Gemini (by Google), are deep neural networks trained on a large corpus of text data with the purpose of predicting the next word in a sentence conditional on previously seen or generated text. These tools have proven to be effective across various domains, from natural language processing to content creation. In this section, we discuss the foundations that make LLMs particularly well-suited for the analysis of hard-to-quantify corporate risks. In particular, by relying on its vast knowledge embedded within these models, they can be used to uncover valuable insights into corporate risks. Indeed, risk exposures are often subtly implied in conference call discussions rather than explicitly stated. Evaluating these exposures requires connecting the information found in conference call transcripts with a broader context, i.e., users' pre-existing knowledge. As we discuss next, LLMs are particularly well suited for this task.

2.1 Why Are LLMs Useful in Detecting Risks?

Bigram-based techniques provide a straightforward way to identify risks, but they are known to generate a significant number of false positives and false negatives when assessing risks (Li et al., 2024).[10] For example, when measuring climate risk, the word 'climate' preceded by the words 'political' or 'economic' can be easily misclassified as being indicative of climate risk exposure (false positive). LLMs are well suited to addressing these problems. They use word embeddings (high dimensional word vectors that encode word meaning) and rely on the revolutionary Transformer architecture (Vaswani et al., 2017), which allows them to encode the meaning of words and sentences depending on their context. Transformer employs a so-called attention mechanism, which effectively directs the model to focus attention on the most relevant words. For example, sentences that appear early in the call may be rel-

---

[10]For an overview of the advancements in different textual analysis methods, see Cong et al. (2020).

evant when encoding the meaning of words later on. This ability to recognize related words is instrumental in minimizing false negatives in sparsely distributed risk-related information in a text.

Specifically, when a user prompts a query requesting to summarize or assess risk exposure, LLMs choose which words and, consequently, sentences are most relevant to the query without confining the analysis to the vicinity of specific risk-related words. In contrast, dictionary-based algorithms have to focus on specific bigrams and typically employ a relatively narrow "attention" window to determine their relevance. For example, bigrams that encode political or regulatory topics (e.g., environmental regulation) are counted as long as they are within a ten-word window relative to the word "risk." Furthermore, many companies' executives may avoid the explicit mention of "risk" or related words while, at the same time, a mention of, e.g., environmental regulation or similar topics implies an important form of uncertainty for the company.

2.2 The Value of General Knowledge

Besides allowing for deep context when analyzing risk-related content, a key feature that distinguishes LLMs is their extensive general knowledge. Because the models are pre-trained on vast amounts of textual data and feature billions of parameters, in addition to learning language structure (Klafka and Ettinger, 2020), they acquire a massive knowledge base (Bubeck et al., 2023) and inherit abilities for human-like reasoning (Liu et al., 2023; Binz and Schulz, 2023). This general knowledge forms a "prior" that the model uses when presented with a new text, affecting how the model interprets the new data.

Fundamentally, the presence of extensive general knowledge allows the model not only to summarize risk-related content effectively but also to make its own assessments and judgments of risks based on the context provided by the conference call transcripts. Such assessment is thus a synthesis of the model's general knowledge and the context provided by the document. The Transformer's architecture enables this integration of knowledge at every

step (when predicting each word), allowing it to generate coherent and contextually relevant insights.

3 Methodology and Implementation

Earnings calls are an attractive source of new information about corporate risks because they contain an informational exchange between the demand (analysts representing stakeholders) and the supply (executives) sides. This choice follows a number of prior studies measuring firm-level risk exposures based on earnings call transcripts (Hassan et al., 2019, 2023, 2024; Sautner et al., 2023a; Chava et al., 2022). In contrast to these studies, we employ AI techniques to uncover risk exposures. Specifically, we use OpenAI's GPT3.5-Turbo LLM to analyze these call transcripts.

A visual representation of our GPT processing pipeline is provided in Figure 2. GPT3.5-Turbo limits combined (input and output) tokens to 4,000, whereas call transcripts are 7,000 tokens on average. Thus, we chunk transcripts into several parts, which is a common practice in the literature (see Goyal et al., 2022; Zhang et al., 2021; Kim et al., 2023, for example). Chunking can improve the quality of GPT's output because the model struggles to generate detailed summaries when processing long documents, whereas it produces summaries on par with humans when processing shorter essays (Choi et al., 2023). Chunking also significantly reduces the computational burden of calculating self-attention scores, which increases quadratically in document length.[11]

To ensure the model's flexibility, we allocate 2,000 tokens for the input text and the remainder for the output text. We divide the transcript into a presentation and Q&A sections to maximize chunk quality. In the presentation part, we avoid splitting the speech by the same executive into different chunks. Similarly, we do not split the Q&A in the middle

---

[11]For example, consider two documents A and B, with similar complexity. Assuming all other things equal, A has 10 tokens and B has 100 tokens. When calculating positional encodings and self-attention, the model has to consider 100 ($=10^2$) pairwise relations. However, when it comes to document B, the model needs to consider 10,000 ($=100^2$) pairwise relations.

of the answer to an analyst's question.[12] We concatenate the model's output at the call level (pooling across chunks).[13] In the absence of risk-related information, we instruct GPT to print "NA" and subsequently purge NAs from the output. Doing so ensures zero values for the exposure measures when calls have no risk-related information.

We design separate prompts for risk assessments and risk summaries. For summaries, we instruct the model to ignore external information sources. For assessments, in contrast, the model is instructed to make judgments accompanied by narrative reasoning. In both cases, we provide the model with a context specifying that the input text is excerpted from an earnings call transcript. The prompt also provides an explanation of each risk and a list of sample questions relevant to understanding risk exposures.

The risk exposures corresponding to summaries, $RiskSum$, and assessments, $RiskAssess$, are constructed as follows:

$$RiskSum_{it} = \frac{\sum_{l=1}^{K_{it}} \text{len}(\mathbf{S}(c_{it}^l))}{\text{len}(c_{it})} \tag{1}$$

$$RiskAssess_{it} = \frac{\sum_{l=1}^{K_{it}} \text{len}(\mathbf{A}(c_{it}^l))}{\text{len}(c_{it})} \tag{2}$$

where c_{it} is earnings call transcript for a company i in quarter t divided into K_{it} chunks $c_{it}^1, c_{it}^2, \cdots, c_{it}^{K_{it}}$. $\mathbf{S}(\cdot)$ is a GPT-based function that generates risk summaries and $\mathbf{A}(\cdot)$ is the corresponding risk assessment function. $\text{len}(\cdot)$ measures the number of words in a given text.[14] We set the temperature parameter for the text generator to zero and do not restrict the maximum output length.[15]

---

[12]When a question from an analyst prompts responses from multiple executives, we consolidate these answers into a single chunk to maintain coherence. Very occasionally, chunks may exceed the assigned input token limit. In such cases, we further divide the chunk into smaller chunks. Further chunking accounts for less than 3% of the total GPT processing.

[13]We do not apply a summarization layer when concatenating the output texts, although this is an option since our purpose is to measure the degree of risk exposure. Even though some content is repeated in different chunks, we interpret this as the topic (risk) being important for the firm. Similarly, bigram-based measures count the total number of occurrences rather than isolating the count of unique mentions.

[14]Note that $\text{len}(c_{it})$ might not equal $\sum_{l=1}^{K_{it}} \text{len}(c_{it}^l)$ since we drop several chunks such as operator instructions.

[15]High temperature might generate creative yet less replicable answers. We set the temperature to zero (its minimum value) to keep the summaries as close as possible to the actual content of the transcript. Language

Our prompts are designed to capture three different types of *RiskSum* and *RiskAssess*. Specifically, we measure (i) political risk summary, *PRiskSum*, and assessment, *PRiskAssess*, (ii) climate risk summary, *CRiskSum*, and assessment, *CRiskAssess*, and (iii) AI risk summary, *AIRiskSum*, and assessment, *AIRiskAssess*. For political risk, GPT's output generally includes explanations of political or regulatory uncertainties, e.g., whether the company is likely to be affected by a new regulation. For climate risks, GPT's answers typically encompass how extreme weather or environmental policy changes might impact the company's operations.[16] Finally, for AI risks imposed by AI itself, GPT often discusses how a firm's primary operations might be replaced or assisted by AI and whether the company's business depends on AI technologies.

4   Data

4.1   Earnings Call Transcripts

We source quarterly earnings call transcripts from the S&P Capital IQ Global Transcript database. Our analysis focuses on US firms' transcripts available between January 2018 and March 2023. We selected this period because (1) generating risk summaries and assessments for each risk metric is costly and time-consuming, (2) a considerable part of the sample is outside of GPT's training window (allowing for pure out-of-sample tests), and (3) this time period is characterized by significant changes in political, climate and AI uncertainty.

A typical earnings call lasts approximately an hour and has two parts: a presentation session during which executives describe the company's performance and a Q&A session during which analysts ask questions. Our data identifies portions of the transcript attributable to

---

models are probabilistic in nature, and zero temperature does not necessarily mean that the output texts will be completely identical across trials. However, setting the temperature parameter to zero, we randomly sample 100 observations and measure $PRiskSum$ and $PRiskAssess$ 30 times each. For all observations, the standard deviation of $PRiskSum$ and $PRiskAssess$ is less than 1% of the mean values. Thus, averaging across multiple trials will only minimally improve the performance.

[16]We focus on the overall climate risk exposure. Our methodology can be readily applied to quantify more specific components, such as 'acute climate risk,' 'chronic climate risk,' and 'transition climate risk' (Li et al., 2024).

a given speaker, allowing us to distinguish between presentation and discussion sessions and between questions by different analysts.

As our analysis focuses on capital market outcomes, we restrict our sample to publicly traded companies. We exclude very short calls and calls without a discussion session (Brochet et al., 2016; Baik et al., 2023).[17] We exclude calls that are conducted in languages other than English, including the ones that are machine-translated into English. After applying these filters, our final sample consists of 69,969 transcripts from 4,983 distinct firms. We further split the sample into two periods: our baseline tests use calls from January 2018 to December 2021, and our post-GPT-training sample includes calls from January 2022 to March 2023.[18]

4.2 Capital Market Variables

Our primary validity test for firm-level risk measures is based on the association between these risk measures and stock price volatility (Engle, 2004). Specifically, we use two forward-looking volatility proxies: implied volatility and abnormal volatility, discussed next. Stock return data is sourced from the Center for Research in Security Prices (CRSP), and market returns and risk-free rates are from Ken French's website.

Implied Volatility We obtain firm-quarter implied volatility from OptionMetrics, following Hassan et al. (2019) and Sautner et al. (2023a). OptionMetrics calculates implied volatility based on the Black-Scholes model for European options and the Cox-Ross-Rubinstein model for American options. We use the implied volatility derived from the 90-day at-the-money options measured as of the end of each fiscal quarter.

---

[17]Such calls account for 1.4% of the total transcript sample. We also drop operator instructions and chunks shorter than 50 tokens as they are more likely to be greetings or irrelevant conversations.

[18]Our benchmarks are bigram-based risk exposure measures from Hassan et al. (2019) and Sautner et al. (2023a). Their sample of released risk exposure measures ends in early 2022, and we match our sample period to theirs. Relatedly, we do not incorporate the climate measures of Li et al. (2024) as their sample ends in 2018.

Abnormal Volatility Since implied volatility hinges on option models' assumptions, which may not reflect the most recent risk-related information (it is also limited to stocks with actively traded options), we also use firms' realized *abnormal* volatility. Following Loughran and McDonald (2014), we measure volatility as the root mean squared errors (RMSE) from the market model residuals. Abnormal volatility is the ratio of post-conference call RMSE to pre-call RMSE calculated as follows. Post-call RMSE is from a market model estimated over the period starting six and ending 28 trading days after the conference call. Specifically, the market model is: $r_{it} = \beta_0 + \beta_1 r_{mt} + \varepsilon_{it}$, where r_{it} is the daily stock return for company i on day t and r_{mt} is the market return on day t. Similarly, using returns data from -257 to -6 days relative to the conference call date, we estimate pre-conference call RMSE.[19] We then take the ratio of the two RMSE values to calculate the abnormal volatility:

$$\text{Abnormal Volatility} = \frac{\text{Post-Conference RMSE}}{\text{Pre-Conference RMSE}} - 1.$$

4.3 Economic Variables

Besides focusing on market-based outcomes, we also study firms' responses to risks by examining their real activities, discussed next.

Investments We follow Stein and Stone (2013) to construct a capital investment measure using the ratio of current quarter capital expenditures ($CapEx_t$) to the recursively updated cumulative capital taken as of the previous quarter end (K_{t-1}). For the initial time period in our sample, K_1 is set to the quarter-end value of property, plant, and equipment from the Compustat Quarterly database. Subsequently, capital is recursively updated as follows: $K_t = K_{t-1} \times (1-\delta) \times (1+\rho_t) + CapEx_t$, where δ is the depreciation rate assumed to be 10%, and ρ_t is the inflation rate measured by the change in the monthly Producer Price Index

---

[19]We require at least ten valid observations to estimate the post-conference RMSE and at least 60 valid observations for pre-conference RMSE. Additionally, in untabulated analyses, we also use alternative estimation windows for post-conference call RMSE estimation and obtained quantitatively and qualitatively similar results.

published by the Federal Reserve Economic Data (FRED). Accordingly, capital investment in period t is measured as $\frac{CapEx_t}{K_{t-1}}$.

Lobbying Activity The Lobbying Disclosure Act of 1995 mandates that each lobbying firm reports its lobbying expenditures and their recipients. We obtain lobbying data from the Center for Responsive Politics, which maintains an archive of lobbying records (as long as lobbying amounts exceed the mandated disclosure threshold). We fuzzy match the names disclosed in lobbying reports to firm names on Compustat. Because lobbying amounts are highly skewed, we use a lobbying indicator $\mathbf{1}\,(\$\text{ Lobby Amount} > 0)$, which takes the value of one when a firm engages in any lobbying and zero otherwise.

Green Patents and AI-Related Patents We obtain patent data from the United States Patent and Trademark Office (USPTO). We then match the assignee's firm name to Compustat following the approach in Kogan et al. (2017). We use patent filings to measure a firm's patenting activities during a given quarter. Following Skinner and Valentine (2023), we use International Patent Classification (IPC) codes to classify patents related to green technology. Similarly, by following Toole et al. (2020), we identify IPC codes related to AI technology.[20] As with the lobbying amounts, the distribution of patent filings is heavily skewed (the 90th percentile constitutes one filing per quarter). Therefore, we use an indicator variable $\mathbf{1}\,(\#\text{ Green Patent} > 0)$ that takes the value of one when a firm files at least one green patent in the following quarter and zero otherwise.[21] The variable $\mathbf{1}\,(\#\text{ AI Patent} > 0)$ is defined similarly.

5 Descriptive Statistics

To provide preliminary evidence that GPT-based risk exposure measures capture risks in the intended ways, we start by examining descriptive statistics. We explore cross-industry

---

[20]The IPC codes that we use to classify green and AI-related patents are available upon request.

[21]For patent and lobby variables, unmatched firm-quarter observations are assumed to be zero. Nonetheless, we find that our results are robust to excluding unmatched observations or zero values (untabulated).

variation in our risk scores and compare the time trends in our measures with those of bigram-based measures. We also directly investigate the face validity of our measures.

5.1 Descriptive Statistics and Industry-Level Averages

In Table 1, Panel A, we report descriptive statistics for risk measures as well as for the variables used in our regression analyses. We observe that risk exposure *assessments* are more than double in length compared to risk exposure *summaries*. Consistent with Hassan et al. (2019) and Sautner et al. (2023a), for most firms, the levels of risk exposure are relatively close to zero. In particular, the median values of $PRiskAssess$, $CRiskAssess$, and $AIRiskAssess$ are 0.011, 0.001, and 0.000, respectively. However, there are also observations with large values. This result implies that in the majority of earnings calls, some risk information is disclosed even though it does not account for a large portion of the call. In addition, political risk is the most commonly mentioned risk, followed by climate risk.

In Panel B, we present Pearson correlations among the risk measures, including the bigram-based political and climate risk exposures. As expected, bigram-based measures and GPT-based measures are positively correlated, although the correlation coefficients are relatively low (ranging from 0.116 to 0.147). Our corresponding assessment and summary measures are highly correlated with the correlation coefficients of 0.742 (political risk), 0.66 (climate risk), and 0.512 (AI risk). The political and climate risks also exhibit a correlation, with correlation coefficients of 0.349 for summaries and 0.483 for assessments. These moderate correlations are expected because environmental and regulatory risks are inherently intertwined. The positive correlation is potentially further exacerbated by our sample period, during which several politically and environmentally significant events happen to coincide.

We verify that the political and environmental risk measures are not capturing the same content.[22] First, we estimate the correlation between the two assessment measures after

---

[22]In Section 7.4, we conduct additional placebo tests and show that both measures capture a distinct risk dimension.

demeaning them by the quarter average. The correlation decreases considerably from 0.483 (overall) to 0.351 (within a quarter). Second, we present word clouds in Figure 5, which display very different keyword patterns for each risk type. Third, for non-zero risk exposure documents, we estimate the pairwise cosine similarity between the two risk assessments. Their average similarity is only 0.421.[23] Lastly, we manually check contents for a random sample of generated summaries and assessments and verify that the two risk categories deal with substantially different topics.

We further validate our measures by examining the two-digit SIC industries with the highest risk assessment scores (Figure 3). For political risk, the high-risk industries include, for example, tobacco products and heavy construction. For climate risk, the high-risk sectors include coal mining, electricity and gas, textile, and paper products. In line with observed positive correlations between the two risk topics, there is some (expected) overlap in the high-risk industries (e.g., textile mill products, electricity & gas). Lastly, for AI risk, the highest risk industry composition is quite different. Business services, engineering services, and electronic equipment are among the highest-ranked industries. We provide a heat map for one-digit SIC industry risk exposure in the Online Appendix, Figure 1. The results are consistent with those in Figure 2 and Figure 3.

5.2   Time Series Variation

We next explore the time series properties of our risk measures, which are depicted in Figure 4. Panel A presents political risk exposure based on GPT summaries and assessments, as well as a bigram-based measure based on Hassan et al. (2019). GPT-based and bigram-based measures exhibit similar time trends. Irrespective of the metric, political risk spikes in 2020, which is the year of the COVID-19 pandemic and the US Presidential Election. We also see another increase in political risk in 2022, during which the outbreak of the Russia-

---

[23]As a benchmark, the cosine similarity between two continuous MD&A reports is 0.8 to 0.9 (Brown and Tucker, 2011). Also, for firms in the same industry, the pairwise similarity is, on average, 0.55 (Peterson et al., 2015).

Ukraine war took place. Thus, GPT-based measures appear to reflect notable political events in a timely and intuitive manner and are aligned with bigram-based measures.

Panel B of Figure 4 plots GPT-based and bigram-based climate risk exposure measures. All three climate risk measures co-move over time, although they are not as closely aligned as the political risk measures. We observe spikes in 2020, the first quarter of 2021, and the first quarter of 2022. The 2020 spike follows the UN Climate Action Summit,[24] and coincides with the outbreak of COVID-19 that introduced uncertainty regarding climate action. The first quarter of 2021 was impacted by the Texas power crisis, where a severe winter storm brought to light the vulnerability of the electricity supply to extreme weather events. Again, observable time trends in our GPT measures line up with notable environmental events.

Lastly, in Panel C, we plot AI-related risk summaries and assessments. Because no public bigram-based AI risk measure is available from prior studies, we do not have a natural benchmark. However, AI risk exposure exhibits an entirely different time trend than the two preceding risks. We observe a steady increase over time, reflecting the heightened significance of AI technologies. A notable increase took place between the third quarter of 2019 and the second quarter of 2020. This period coincides with the release of Transformer-based BERT (Google) and GPT-2 (OpenAI) models, which have since been widely adopted by companies. The highest increase in AI risk occurred in the first quarter of 2023, immediately following the release of GPT3.5, which became viral around the world.[25]

5.3 Face Validity

To examine the face validity of our measures, we read a number of generated summaries and evaluations for the three types of risk and confirm that they address the relevant topic. Appendix B presents several snippets from GPT-processed risk exposure documents. Each summary and assessment pair is from the same respective earnings call transcript. Assess-

---

[24]During which climate activist Greta Thunberg delivered her widely resonating speech accusing world leaders of failing to tackle climate change.

[25]See this article published in Fortune for anecdotal evidence.

ments, as expected, contain richer contents that often include GPT's judgments based on the given context.

In example B1 (Appendix B), the political risk summary expresses a cautionary note that the recent government auction of the HS1 project (high-speed railway line) in the UK creates exposure to political and regulatory uncertainties in Europe. In turn, the corresponding political risk assessment (see example B2) highlights that "*deficit reduction in these regions may lead to an increased flow of government disposals and potentially PFI (Private Finance Initiative) opportunities.*" In the earnings call transcript, there is no mention of government disposals or PFI in conjunction with deficit reductions. Rather, PFI is mentioned in the later part of the executive's presentation. This example highlights GPT's ability to connect information pieces into a logical assessment.

The examples of climate risk summaries (B3) and assessments (B4) display a starker difference. The text of the climate risk summary only contains "NA," implying no direct mention of climate-related uncertainties. However, despite the lack of explicit discussion of climate risks, the corresponding climate risk assessment states that the company is subject to risks related to significant energy consumption and electronic waste. It also mentions a potentially high carbon footprint and the need for compliance with environmental regulations.

Turning to the examples of AI Risk summaries (B5) and assessments (B6), we observe that the summary identifies the firm's active efforts to incorporate AI into its operations and provides more specific details. In contrast, risk assessment (B6) takes a more holistic view of whether the company's business is dependent on AI technologies.

In sum, the above-mentioned examples illustrate that the generated risk summaries and assessments are readable and generally contain relevant information when it comes to understanding firm-level risks. Further, risk summaries and assessments are distinct from each other in logically expected ways.

5.4 Variance Decomposition

In this subsection, we perform a variance decomposition of our proposed risk exposure measures into time, industry, and firm-level variation.[26] The results are presented in Table 2, Panel A1, B1, and C1. The table shows the incremental R-squared values attributable to each set of fixed effects. While the descriptive evidence above reveals clear time trends and industry variation in our risk exposure measures, they jointly explain only a small portion of the variation in risk exposures. These results are closely in line with the findings in Hassan et al. (2019), suggesting that macroeconomic and sector-level risk variance is only the tip of the iceberg. There are some differences in the explained portion of the variance depending on the specific risk type and the measurement approach, but the general consistency is apparent and noteworthy. The results also imply that the variation attributable to firm-level factors is in the range of 80% to 90%, highlighting the importance of firm-level risk measurement.

We further decompose firm-level variance into time-varying and invariant portions. Specifically, we zoom in on residual variance by including firm-level fixed effects and report their corresponding R-squared values. Firm fixed effects explain between 21% and 35% of firm-level (residual) variance, thus implying that almost 65-79% of variation is time-varying.[27] Pairwise comparisons of different types of risk measures (summary-, assessment-, and bigram-based) reveal that the assessment-based measures tend to be the more stable within a firm and across time.

In sum, although the risk measures exhibit notable aggregate time trends (as seen in Figure 4), rich firm-level heterogeneity exists in exposures to these aggregate trends across different types of risks.

---

[26]We use two-digit SIC industry classification. In untabulated analyses, we also use three-digit and one-digit level industry classifications and find consistent results.

[27]To further validate our measures, we estimate the measurement errors in our variables following Sautner et al. (2023a) (Online Appendix Table 1). Measurement errors for *PRiskAssess* and *CRiskAssess* are 2.71% and 8.50%, respectively. *AIRiskAssess* has a relatively higher measurement error of 27.63%, which is due to the high proportion of zero values in earlier years. Once we exclude zeros from analysis, the measurement error decreases to 6.59%.

20

6 Capital Market Consequences

A valid measure of risk must exhibit association with volatility (Engle, 2004). In this section, we examine associations between our risk proxies and two stock price volatility variables introduced previously, namely, implied volatility and abnormal volatility (described in Section 4.2). Specifically, we estimate the following OLS regression:

$$Volatility_{it+1} = \beta Risk_{it} + \gamma \mathbf{X}_{it} + \delta_x + \varepsilon_{it}, \tag{3}$$

where $Volatility_{it+1}$ is a proxy for the volatility of a firm i's stock price in quarter $t + 1$.[28] $Risk_{it}$ is one of the firm-level risk exposure measures for firm i in quarter t. Depending on the specification, we include a GPT summary-based risk exposure, $RiskSum$, and a GPT assessment-based risk exposure, $RiskAssess$, as well as a bigram-based measure, $RiskBigram$. Prefix P, C, or AI in front of $Risk$ denotes political, climate, and AI-related risks, respectively. $\mathbf{X}_{it}$ is a vector of firm-quarter controls. Following Hassan et al. (2019), we control for the log of total assets in our main analysis. We also include an extensive set of additional controls in the Online Appendix and verify the robustness of our results. δ_x denotes various fixed effects. Specifically, we estimate two different sets of fixed effects in our main tests: (1) time (δ_q) and industry (δ_s) fixed effects, and (2) the interactions of time and industry fixed effects ($\delta_q \times \delta_s$). We do not control for firm-fixed effects, as firm-level variation in our measures is of interest. However, in the Online Appendix, we confirm that our results are robust to controlling for firm fixed effects. Robust standard errors are clustered at the firm level. All continuous variables are winsorized at 1% and 99% levels.

6.1 Political Risks

We start discussing our main results with respect to political and climate risk because they received considerable attention in prior studies (and because bigram-based benchmarks

---

[28]In the Online Appendix, we also use realized volatility, which is measured as the standard deviation of stock returns over a 90-day window, as an additional volatility variable.

are available for comparison purposes). Thus, we confine our main sample to the 2018-2021 period as the bigram-based measures are updated until early 2022 (Hassan et al., 2019; Sautner et al., 2023a).

The estimates for political risk are reported in Table 3. As shown in columns (1) and (2), both GPT-based measures of political risk, *PRiskSum* and *PRiskAssess*, exhibit positive and highly statistically significant associations with both implied volatility and abnormal volatility. This result holds for industry-fixed effects (Panel A) and industry-time fixed effects (Panel B). In columns (3) and (4), we add the bigram-based risk measure as an additional control variable to the regressions.[29] We find that the effect of our GPT-based measures does not attenuate and that GPT-based measures subsume information in *PRiskBigram*. Overall, this evidence suggests that LLM-based risk measures effectively capture political risk and dominate dictionary-based measures.

We next turn to the comparison of GPT-based summaries vs. assessments in explaining volatility. When adding the risk measures simultaneously to the same regression, as we do in column (5), we find that the assessment-based measure of political risk dominates the summary-based measure for both implied and abnormal volatility. This result is important as it implies that GPT's ability to synthesize general knowledge acquired from its training is valuable when assessing political risk exposure.[30] Overall, these findings reinforce the notion that LLMs can produce insights by synthesizing the input text and general knowledge.

In terms of the economic significance of the results, a one standard deviation increase in *PRiskAssess* translates to a 0.04 standard deviation increase in implied volatility (using column (5) of Panel A), which is a non-trivial amount. Similarly, one standard deviation increase in *PRiskAssess* is associated with a 0.06 standard deviation increase in abnormal volatility (using column (5) of Panel A).

---

[29]This measure is statistically significant when used on its own.

[30]For abnormal volatility, we observe negative and significant coefficient for *PRiskSum* when including industry-time fixed effects. This phenomenon arises due to the high positive correlation between *PRiskAssess* and *PRiskSum*. The effect is expected when both proxies contain correlated measurement errors and one of the proxies is the dominant one.

6.2 Climate Risks

Next, we estimate the same model using the climate risk proxies as our independent variables. The results are reported in Table 4. GPT-based measures bear a positive and statistically significant association with stock price volatility both on a standalone basis or when used jointly with *CRiskBigram*. The statistical power to detect climate risk is highest in the case of GPT-based assessments, *CRiskAssess*, which exhibit several times higher t-values compared to the summary-based measure *CRiskSum*. Furthermore, when the three proxies are included in the model simultaneously, *CRiskAssess* is the only one that consistently remains statistically significant and exhibits the expected sign. These results are consistent regardless of the fixed effect structures and hold for implied and abnormal volatility.

Overall, LLMs perform well at identifying climate-related uncertainty and subsume bigram-based measures. Furthermore, the analysis also confirms the value of LLM's general knowledge – GPT assessments – in evaluating climate risks, going beyond the information discussed during the conference calls.

6.3 AI-Related Risks

Lastly, Table 5 examines the association between AI-related risk exposure and stock price volatility. As there is no established bigram-based AI risk measure in the literature, we only analyze the GPT-based measures. In our primary sample, we find limited evidence that AI-related risk is associated with stock price volatility. For example, we do not find a statistically significant relation between AI risk proxies and implied volatility. In the case of abnormal volatility, we only find significant associations in the case of *AIRiskSum*. This result is not surprising as AI risk has emerged as an economically important phenomenon only very recently. Indeed, from 2018 to 2021, nearly 60% of the observations have zero AI-related risk exposure. In the following section, we explore whether AI risk becomes significant in explaining stock price volatility in the most recent period.

6.4 Analysis Outside GPT's Training Period

Note that the analyses in our study are out-of-sample in the sense that the language model is not trained to predict any of the outcome variables that we examine. Nevertheless, one potential concern is that our main sample overlaps with the GPT model's training period, which extends to September 2021. As noted by Sarkar and Vafa (2024), large language models may suffer from a look-ahead bias. In particular, the model trained on a large corpus of textual data might incorporate information about future events when assessing risks based on the current earnings call.[31]

To further address this concern, we conduct "true" out-of-sample tests by using conference calls from January 2022 to March 2023. This period is outside of the GPT's training sample. We present the results of this analysis in Table 6. Panels A and B use GPT-based political risk and climate risk proxies, respectively, as their main independent variables. Even in this more limited sample, we continue to find positive and highly significant associations between GPT-based risk measures and firms' stock price volatility. The only exception is that *PRiskSum* and *CRiskSum* lose statistical significance when examining abnormal volatility (column (3) of both panels). Nevertheless, they remain positive and retain their economic magnitudes. Most importantly, the assessment-based exposure measures are significantly associated with both volatility across all model specifications.

For AI Risk, reported in Panel C, in fact, we now find stronger results during this more recent time window. Specifically, we observe that both *AIRiskSum* and *AIRiskAssess* become positive and significantly significant determinants of volatility. This finding directly maps into our evidence in Figure 4C, which indicates that AI risk has soared in the most recent year in our sample. In terms of economic magnitudes, a one standard deviation increase in

---

[31]In our setting, we do not instruct the model to predict future risk. Instead, our prompts explicitly instruct the model to use the information included in the transcript and synthesize it with its general knowledge, making it less subject to the look-ahead bias. As a simple validity check, we examine whether our output texts contain the word "COVID-19" before January 2020. We do not find a single instance out of 115,860 (19,310 multiplied by six different outputs) output texts that mention COVID-19. This result confirms that the model is not using future knowledge to compose its answers. Rather, it seems that the model uses the provided information to construct and synthesize information.

AIRiskAssess is associated with a 0.03 standard deviation increase in implied volatility and a 0.03 standard deviation increase in abnormal volatility, which is once again an economically sizable effect.

These findings lead to two important insights. First, GPT's ability to produce valid risk exposure measures cannot be attributed to its look-ahead bias. Rather, it suggests that large language models are useful for investors' decision-making when used outside of the training window. Second, the fact that AI risk is correlated with volatility only in the period after it soars constitutes further evidence of the construct validity of our measures.

6.5   Robustness Checks

We also conduct a battery of robustness tests in the Online Appendix. First, motivated by our findings in Section 5.4, we include *firm* fixed effects into the model (Online Appendix Table 2). Second, we expand the set of our control variables to include leverage, cash holdings, tangible assets, profitability, and capital expenditures (Online Appendix Table 3). In both sets of robustness tests, we continue to find strong results for the GPT-based assessment measures, which outperform the dictionary-based proxies. Third, we use realized volatility (instead of implied volatility) as another dependent variable. We show that *PRiskAssess* and *CRiskAssess* are positively associated with this volatility measure (Online Appendix Table 4). Fourth, to mitigate the concern that extreme risk exposure values are driving our results, we plot the association between the risk assessment variables and implied volatility in Online Appendix Figure 2. We randomly sample observations whose risk assessment values are within 10% and 90% percentile and show that the positive association remains.

Overall, our analysis lends support to the notion that GPT-based firm-level political risk exposure and climate change risk exposure are useful in capturing corporate risks. These measures are superior to traditional bigram-based measures, indicating that LLMs can produce substantially informative risk assessments.

7 Risk Exposures and Firm Decisions

An alternate approach to evaluate the economic usefulness of AI-based risk measures involves examining their ability to explain firms' economic decisions. We focus on capital investment decisions, followed by the examination of corporate actions aimed at diminishing risk exposures, such as lobbying or developing intellectual property.

7.1 Investment Decisions

Theoretical considerations suggest that heightened uncertainty regarding a firm's future commands a risk premium that investors require in exchange for capital (e.g., Pástor and Veronesi, 2013). Consequently, higher risk exposure should make investments in capital projects harder to finance. Supporting this notion, Baker et al. (2016) and Hassan et al. (2019) show a decline in capital expenditures in response to escalating political uncertainty.[32]

To examine the association between risk exposures and investment, we estimate the following OLS regression:

$$Investment_{it} = \beta Risk_{it} + \gamma \mathbf{X}_{it} + \delta + \varepsilon \tag{4}$$

where $Investment_{it}$ is capital expenditure intensity (see Section 4.3 for more details). We include quarter and industry-fixed effects. The remaining variables are the same as in Equation (3). We winsorize continuous variables at 1% and 99%. Standard errors are clustered at the firm level.

Table 7 presents the results. In Panel A, columns (1) and (2) show that *PRiskSum* and *PRiskAssess* have a negative and statistically significant association with capital expendi-

---

[32]In theory, investments are also expected to decline in response to increased uncertainty because the option to wait becomes more valuable (Dixit and Pindyck, 1994; Bloom et al., 2007). Although the cost of capital effect on investment applies to all risk types, there is a countervailing effect when capital expenditures become an important part of a firm's strategy to counteract a specific risk. For example, companies might be less inclined to rely on tangible investments to alleviate political risk, favoring methods like lobbying or political donations instead. However, to address risks posed by AI, significant infrastructure investments are likely to be needed to bolster resilience to AI risk.

tures. In line with prior results, the coefficient of *PRiskAssess* is not only larger than that of *PRiskSum* but also more significant. When we jointly include LLM-based and bigram-based measures in columns (3) and (4), the LLM-based measures remain significantly associated with investments and bigram-based measures are not. In line with prior evidence, we also find that *PRiskAssess* is also the dominant proxy when considering several proxies simultaneously (column 5). In terms of economic significance, a one standard deviation increase in *PRiskAssess* translates to a 1.3% decrease in investment, on average (based on estimates in column (5)).

In Panel B, we present the results for climate risk and observe a similar pattern as in Panel A. *CRiskAssess* bears a negative relation with investments, which remains significant with and without the bigram-based measure or *CRiskSum* in the same regression. In terms of economic magnitude, a one standard deviation increase in *CRiskAssess* is associated with a 0.8% decrease in average investment.

In Panel C, we do not find a significant relation between AI risk and investments. In both columns, we find positive yet not statistically significant coefficients on both risk exposure measures.

To shed further light on the relationship between AI-related risk and capital expenditure, we focus on the most recent period starting in January 2022 and ending in March 2023 (during which AI risks skyrocketed) and perform several cross-sectional tests. In particular, we examine whether companies that are financially constrained, i.e., lack internal funds to finance investments, have a decrease in their capital expenditures in response to increasing AI risks (cost of capital channels). In contrast, we expect that financially unconstrained firms will increase their investments to tackle AI-related challenges and opportunities.

For this analysis, we partition the sample based on the levels of cash holdings and leverage measured as of 2021 and report the analysis in the Online Appendix, Table 5. We observe that the least financially constrained firms exhibit the largest increase in capital expenditures in response to *AIRiskAssess*. In contrast, we find that the most constrained firms reduce

their investments. These findings help to understand the seemingly counterintuitive impact of AI risk on investments.

Overall, our results indicate that GPT-based risk exposure measures explain corporate investment decisions in intuitive ways, in line with theoretical priors. However, the effect on investments is nuanced and depends on the type of risk.

7.2 Firm Responses to Mitigate Risk Exposures

In this subsection, we turn to the examination of firms' actions to mitigate their risk exposures. To do so, we focus on three outcome variables that correspond to three different risk types. First, we gauge our political risk measures by probing their ability to explain companies' lobbying activities (Peltzman, 1976). Second, we investigate firms' reactions to climate threats by examining the acquisition of green patents (Blyth et al., 2007; Stern and Valero, 2021). Finally, we measure firms' responses to AI risk by focusing on AI-related innovation (the filing of AI-related patents).

We estimate the following OLS regression:

$$Response_{it+1} = \beta Risk_{it} + \gamma \mathbf{X}_{it} + \delta + \varepsilon \tag{5}$$

where $Response$ is one of the three indicator variables discussed in Section 4.3 in more detail: $\mathbf{1}(\$ \text{ Lobby Amount} > 0)$ is an indicator variable that equals one if firm i lobbying expenditures are greater than zero in the quarter $t + 1$ and zero otherwise, $\mathbf{1}(\# \text{ Green Patent} > 0)$ is an indicator variable that equals one when firm i files at least one green patent in the quarter $t + 1$ and zero otherwise, and $\mathbf{1}(\# \text{ AI Patent} > 0)$ is an indicator that equals one when firm i files at least one AI-related patent in the quarter $t + 1$ and zero otherwise. We expect a positive association between these variables and their corresponding risk exposure measures. As before, we use quarter and industry fixed effects and the same set of control variables as in Equation (3).

Table 8 presents how firms respond to a change in firm-level risk exposure. Panel A

28

focuses on lobbying activity in response to political risk. We find a positive and significant association between *PRiskAssess* and lobbying, whereas *PRiskSum* only shows a limited association. When *PRiskAssess* is included jointly with *PRiskBigram*, the two measures both remain incrementally informative. A one standard deviation increase in *PRiskAssess* is associated with a 1.40% point increase in the likelihood of lobbying in the following quarter (based on estimates in column (5)).[33]

Panel B studies the association between green patents and climate risk exposure. In column (2), we find a positive and significant association between *CRiskAssess* and green patent filings. When we include all three risk proxies in one regression, *CRiskAssess* comes out as a dominant proxy. In terms of economic magnitude, based on our estimates in column (5), a one standard deviation increase in *CRiskAssess* is associated with a 0.73% point increase in the likelihood of filing a green patent in the following quarter.

Finally, in Panel C, we repeat the analysis for AI risk and AI patent filings. We find that one standard deviation increase in *AIRiskAssess* is associated with a 2.39% point increase in the likelihood of filing at least one AI patent in the subsequent quarter.

Overall, we find that companies respond to risks captured by our measures by taking actions to mitigate them. We view this finding as further evidence of the validity and economic usefulness of GPT-based risk exposures.

7.3 Further Out-of-Sample Analysis

Similar to Section 6.4, we perform a subsample analysis from January 2022 to March 2023 to ensure that our results are not attributable to GPT seeing the underlying data during its training phase. Table 9 presents the results of this analysis. Overall, our results are consistent with those in Tables 7 and 8, addressing concerns about possible in-sample bias. In fact, for lobbying activity, *PRiskSum* becomes even more significant out-of-sample.

---

[33]Note that *PRiskSum* is negative and statistically significant in column (5). To the extent the same information is repeated in the assessment summaries, the model performs better by filtering it out from the assessment-based measure.

One notable difference in this analysis is that AI risk, on average, shows a positive and significant association with investments in the most recent year. This finding reconciles with the additional analysis in section 7.1 and is noteworthy because standard theory would predict that investment should decrease in risk. As discussed previously, the likely explanation for this finding is that companies respond to soaring AI risk exposures by making investments in AI-related technology and infrastructure.

7.4 Robustness Checks

We perform several robustness checks. We gauge the robustness of our results by including firm fixed effects (Online Appendix Table 6), as well as by including an extensive set of controls (Online Appendix Table 7).[34] Our results are similar to these alternative specifications. To mitigate the concern that extreme values are driving our results, we visualize the linear association between risk exposure measures and investments after excluding extreme values (Online Appendix Figure 3).

To bolster our findings that each risk is exclusively associated with its corresponding corporate action (and not just proxies for general risk), we perform several placebo tests. Specifically, we regress a lobbying activity indicator on *CRiskAssess* and *AIRiskAssess*, a green patent indicator on *PRiskAssess* and *AIRiskAssess*, and an AI patent indicator on *PRiskAssess* and *CRiskAssess*. We report the results in Online Appendix Table 8. In general, each economic outcome is most significantly associated with its corresponding risk measure. One exception is *AIRiskAssess*, which is also positively associated with lobbying and green patent activities.

Finally, because patenting and lobbying activity variables are heavily skewed, we use their natural log as dependent variables. Additionally, following Cohn et al. (2022), we also use Poisson regressions for patent counts (Online Appendix Table 9). Our results remain qualitatively similar. Overall, our findings highlight that firms respond to risk by subsequently

---

[34]In Tables 7 and 8, we only report the results using industry and quarter-fixed effects. However, the results are almost identical when we allow for the interactions of quarter and industry-fixed effects.

mitigating their risk exposure.

8 Aggregate Risk Exposures

It is also interesting to examine whether our firm-level measures can be extended to measure uncertainty at the macroeconomic level. To do so, we focus on aggregate political uncertainty that received attention in the prior literature. Specifically, we construct a quarterly aggregate political risk measure (PRA^{Agg}) by averaging $PRiskAssess$ across firms. As almost all US public firms host quarterly earnings calls, our measure is expected to represent the entire economy. We contrast our measure of aggregate political uncertainty against a widely-used aggregate measure of economic policy uncertainty (EPU) developed by Baker et al. (2016) – an index based on a static keyword list encountered in the business press.

8.1 Volatility, Investment, and Uncertainty

Because PRA^{Agg} is a quarterly measure and EPU is monthly, we obtain quarterly EPU scores by taking its three-month average. Following Baker et al. (2016), we validate our measure by examining its association with stock market volatility, capital investments, and unemployment. We use one-quarter-ahead realized volatility and implied volatility as proxies for stock market volatility. We construct aggregate volatility and investment variables by averaging the corresponding firm-level values each quarter. Unemployment is a one-quarter-ahead unemployment index obtained from FRED.

Figure 6 illustrates the univariate associations between the volatility variables and the two uncertainty indexes (PRA^{Agg} and EPU). We find pronounced (largely linear) positive associations between PRA^{Agg} and aggregate volatility proxies. EPU also exhibits positive associations, however, they appear to be less stable.

To analyze these relations more formally, we regress quarterly volatility (and other outcomes) variables on EPU and PRA^{Agg}, while clustering standard errors at the quarter level. Results are reported in Table 10, Panel A. In column (2), we find that PRA^{Agg} strongly

31

predicts future stock market volatility. The adjusted R-squared is nearly 0.316, which is an order of magnitude higher than that obtained using only EPU (0.025). In column (3), we include both variables in the same regression and find that PRA^{Agg} dominates EPU, which loses its economic magnitude and statistical significance.

In columns (4), (5), and (6), we use implied volatility as our dependent variable. Both EPU (column (4)) and PRA^{Agg} (column (5)) show positive associations with implied volatility. However, the adjusted R-squared is 0.749 when using PRA^{Agg} vs. 0.434 in the case of EPU. When the two variables are simultaneously included in column (6), both EPU and PRA^{Agg} remain statistically significant, though EPU exhibits a substantially lower t-value. The economic significance of PRA^{Agg} is also considerably higher compared to that of EPU.

We find generally similar results for aggregate investment and unemployment rates (Panel B). PRA^{Agg} is negatively associated with capital investment and positively associated with future unemployment. PRA^{Agg} does not lose its explanatory power when we control for EPU, which is consistently dominated, both statistically and economically. Overall, our findings suggest that an LLM-based aggregate political uncertainty measure provides useful insights not captured by EPU.

8.2 Future Volatility vs. Past Volatility

In the previous section, we use PRA^{Agg} to predict realized volatility in the next quarter. We extend this analysis by examining whether PRA^{Agg} and EPU can predict future volatility or explain past volatility for various horizons. Specifically, we regress future volatility measured one to four quarters ahead on PRA^{Agg} and EPU. Also, we regress past volatility starting four to one quarters before the quarter of PRA^{Agg} and EPU measurement.

We summarize the results for this analysis graphically in Figure 7. The figure plots the coefficients on PRA^{Agg} (Panel A) and EPU, and their corresponding t-statistics (Panel B). We find that PRA^{Agg} can predict future volatility of up to four quarters ahead. However, when explaining past volatility, PRA^{Agg} loses statistical significance. Notably, EPU becomes rela-

32

tively more significant in explaining past volatility and loses its significance when predicting future volatility.

These results are noteworthy as they suggest that *EPU*, which is based on newspaper articles, is more backward-looking. This is intuitive as the press often covers developments concurrently or even with a significant lag. In contrast, PRA^{Agg}, which gauges the corporate pulse based on the current conference call discussion, is more forward-looking.

8.3 Other Results and Robustness

Controlling for VIX Index Another commonly used measure of stock market uncertainty is the VIX index developed by S&P. In Online Appendix Table 10, Panel A, we repeat our aggregate analysis controlling for VIX. Given that VIX and EPU are defined on a monthly basis, we also construct monthly *PRA* scores based on cross-sectional averages of all conference calls in a given month.[35] We then regress one-month-ahead volatility or implied volatility on *EPU*, PRA^{Agg}, and *VIX*. Notably, in columns (1) and (4), we show that *VIX* dominates *EPU* index, whereas columns (2) and (5) indicate that PRA^{Agg} dominates *VIX*. Lastly, columns (3) and (6) show that, when all three variables are included simultaneously, PRA^{Agg} retains statistical significance.

Market Weights To bolster the robustness of our results, we also perform several alternative measurements. First, in Online Appendix Table 13, Panel B, we include *PRAW*, which is the weighted average of *PRiskAssess* scores by their respective market capitalization values. Overall, the results are qualitatively similar.

Other Market Volatility Measures We also test whether our results hold using alternative market volatility measures. We use the market-wide volatility index from Baker et al. (2019) and the standard deviation of market returns from the CRSP database. We present the results in the Online Appendix Figure 4. The findings are very similar.

---

[35]In untabulated tests, we also confirm that we obtain very similar results in the analyses of Table 10 when we use monthly cross-sectional or rolling averages of *PRA* scores (instead of quarterly scores).

9 Equity Market Pricing

In our last set of tests, we probe the asset pricing implications of our GPT-based risk exposure measures. In theory, higher risk exposures should be associated with higher expected equity returns. Establishing this link is challenging in our setting as the asset pricing methodology requires a relatively long time series. We use our entire sample period, starting in January 2018 and ending in March 2023, for this test. As this period is still relatively short, our results (t-statistics) are likely to understate the significance of risks picked up by our proxies.

We test the pricing of GPT-based proxies by running Fama and MacBeth (1973) regressions and performing portfolio analysis for our risk variables. In this analysis, we focus on risk assessment-based measures as our prior tests imply their dominance. As Fama and Mac-Beth (1973) regressions typically use annual characteristics, we annualize our risk proxies by taking their average across four quarters each year ($PRiskAssess^{ann}$, $CRiskAssess^{ann}$, and $AIRiskAssess^{ann}$).[36] We construct our portfolios on March 31 of the subsequent year to allow for three or more months for stock prices to incorporate information disclosed during earnings calls (e.g., we compute annualized risk exposure over 2021 and form the portfolios on March 31, 2022).

We regress monthly stock returns on the natural log of each risk exposure, log of the market value of equity ($\log(ME)$), log of book-to-market ratio ($\log(BE/ME)$), operating profitability (*Profitability*), investment (*Investment*), lagged one-month return ($r_{0,1}$), and lagged annual return after excluding the most recent month ($r_{2,12}$) (Fama and French, 2015).[37] We report Newey-West t-values with a lag of 3. All continuous independent variables are trimmed at 1% and 99% levels.

---

[36]We exclude observations with zero annualized risk exposures since they are likely not to feature any discussion of political, climate change, or AI-related risks during earnings calls (or, alternatively, they might be instances where the model fails to generate meaningful output).

[37]*Investment* is given by growth in total assets; *Profitability* is (total revenue – cost of goods sold – (sales, administrative expense – R&D expense)) scaled by total assets; $\log(ME)$ is the natural logarithm of the market value, and $\log(BE/ME)$ is the natural logarithm of the book-to-market ratio.

We report Fama and MacBeth (1973) analysis in Table 11, Panel A. After controlling for asset characteristics, all three risk exposure proxies exhibit positive coefficients. The coefficients on $CRiskAssess^{ann}$ (0.211) and $AIRiskAssess^{ann}$ (0.317) are positive and statistically significant at conventional levels. The coefficient on $PRiskAssess^{ann}$ is positive (0.077), yet not statistically significant (t-value = 1.36). Overall, the positive associations support the pricing of the corresponding risks in equity markets.

For the portfolio analysis, we construct quintile portfolios based on the values of each annualized firm-level risk exposure measure. As earlier, we form portfolios on March 31 of the subsequent year relative to risk exposure and hold them for one year.

Table 11, Panel B presents equal-weighted high-minus-low portfolio alphas based on Fama and French (2015)'s five-factor model. Columns (1), (2), and (3) report monthly alphas from $PRiskAssess^{ann}$, $CRiskAssess^{ann}$, and $AIRiskAssess^{ann}$, respectively. We observe that portfolio alphas almost monotonically increase across risk exposure quintiles. Accordingly, all three columns report positive high-minus-low alphas. For political risk exposure, the annualized alpha is 5.28%, which is economically sizable yet does not attain statistical insignificance (t-value = 1.51). For climate change risk exposure, the annualized alpha is 6.72% and is statistically significant (t-value = 1.90). For AI-related risk exposure, the annualized alpha is 6.36% and is also statistically significant (t-value = 2.31).

Overall, our results suggest that GPT-based risk exposure measures successfully extract firm-level risks that are priced in equity markets. Further, we show that such risk exposure measures are associated with sizable alphas unexplained by the five-factor asset pricing model.

10 Conclusion

In this paper, we evaluate whether recent advances in AI technology can help investors assess critical aspects of corporate risks. We evaluate these risks based on the information disclosed during companies' earnings calls. More specifically, we use a Large Language Model,

GPT 3.5 Turbo, to develop and validate three proxies for firm-level exposure to political, climate, and AI-related risks, all of which have been of primary concern to firms' stakeholders in recent years. We also investigate LLM's ability to leverage general knowledge to enhance summaries of risk-related content from conference calls by adding their own insights.

Our measures of firm-level risk exposures exhibit large within-firm variation and move across industries and over time in intuitive ways. More importantly, each of the three measures is a powerful predictor of future stock price volatility and helps in explaining firms' policies, such as investment and innovation. Furthermore, GPT-based assessments are consistently more informative about firms' risks than GPT summaries, highlighting the value of LLM's general knowledge for the analysis of corporate communication. We also find that GPT-based risk measures consistently subsume the existing bigram-based risk proxies when subjected to the task of predicting stock market volatility and explaining economic decisions. Our findings are robust to a number of research design choices and hold outside of the training window of the GPT model.

We also show that the LLM-based measure of aggregate political risk contains valuable information about the macro-economy, generally dominating the existing word count-based measure of economic policy uncertainty (EPU) in predicting aggregate stock market volatility, investment, and unemployment.

We conclude that generative AI technology enables users to obtain valuable insights about firm-level risks at a relatively low cost. The generated risk summaries and evaluations are not only human-readable but also easily convertible to risk exposure scores. Overall, our study highlights the economic usefulness of generative AI tools, such as ChatGPT, in providing helpful input into investors' decisions based on unstructured textual information (Kim et al., 2023).

We also acknowledge several limitations that come with the reliance on generative LLMs. GPT outcomes can be sensitive to prompt quality, and prompt engineering becomes an important starting point. Unless provided with relevant context, GPT may generate incorrect

evaluations or provide random answers (known as hallucinations).[38] We recommended that users examine and reconcile GPT's answers with the corresponding source documents when making decisions. Nevertheless, in a large sample, the usefulness of generative AI in assessing risks is apparent.

---

[38]Even with perfect prompts, GPT might still generate incorrect information. This is because GPT is a language model whose primary purpose is completing a sentence in the most natural manner. This process is inherently *not* deterministic, and GPT might produce incorrect answers.

References

Baik, B., A. G. Kim, D. S. Kim, and S. Yoon (2023). Managers' vocal delivery and real-time market reactions in earnings calls. *Available at SSRN 4398495*.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Baker, S. R., N. Bloom, S. J. Davis, and K. J. Kost (2019). Policy news and stock market volatility. Technical report, National Bureau of Economic Research.

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

Binz, M. and E. Schulz (2023). Using cognitive psychology to understand gpt-3. *Proceedings of the National Academy of Sciences 120*(6), e2218523120.

Bloom, N., S. Bond, and J. Van Reenen (2007). Uncertainty and investment dynamics. *The Review of Economic Studies 74*(2), 391–415.

Blyth, W., R. Bradley, D. Bunn, C. Clarke, T. Wilson, and M. Yang (2007). Investment risks under uncertain climate change policy. *Energy Policy 35*(11), 5766–5773.

Brochet, F., P. Naranjo, and G. Yu (2016). The capital market consequences of language barriers in the conference calls of non-us firms. *The Accounting Review 91*(4), 1023–1049.

Brown, S. V. and J. W. Tucker (2011). Large-sample evidence on firms' year-over-year md&a modifications. *Journal of Accounting Research 49*(2), 309–346.

Bubeck, S., V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg, et al. (2023). Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*.

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.

Cao, S., W. Jiang, B. Yang, and A. L. Zhang (2023). How to talk when a machine is listening: Corporate disclosure in the age of ai. *The Review of Financial Studies 36*(9), 3603–3642.

Cao, S., B. Yang, and A. L. Zhang (2023). Beyond the lines: Deciphering private information from fund managers' narratives. *Available at SSRN 3713966*.

Chava, S., W. Du, A. Shah, and L. Zeng (2022). Measuring firm-level inflation exposure: A deep learning approach. *Available at SSRN 4228332*.

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 (2023). Expected returns and large language models. *Chicago Booth Research Paper*.

Choi, J. H., K. E. Hickman, A. Monahan, and D. Schwarcz (2023). Chatgpt goes to law school. *Available at SSRN*.

Cohn, J. B., Z. Liu, and M. I. Wardlaw (2022). Count (and count-like) data in finance. *Journal of Financial Economics 146*(2), 529–551.

Cong, L. W., T. Liang, B. Yang, and X. Zhang (2020). Analyzing textual information at scale. *Information for Efficient Decision Making 10*, 239–271.

Dixit, A. K. and R. S. Pindyck (1994). *Investment under uncertainty*. Princeton University Press.

Eisfeldt, A. L., G. Schubert, M. B. Zhang, and B. Taska (2024). The labor impact of generative ai

on firm values. *UCLA Working Paper*.

Engle, R. (2004). Risk and volatility: Econometric models and financial practice. *American Economic Review 94*(3), 405–420.

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

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

Frankel, R., J. Jennings, and J. Lee (2022). Disclosure sentiment: Machine learning vs. dictionary methods. *Management Science 68*(7), 5514–5532.

Goyal, T., J. J. Li, and G. Durrett (2022). News summarization and evaluation in the era of gpt-3. *arXiv preprint arXiv:2209.12356*.

Hassan, T. A., S. Hollander, L. Van Lent, M. Schwedeler, and A. Tahoun (2023). Firm-level exposure to epidemic diseases: Covid-19, sars, and h1n1. *The Review of Financial Studies 36*(12).

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.

Hassan, T. A., J. Schreger, M. Schwedeler, and A. Tahoun (2024). Sources and transmission of country risk. *The Review of Economic Studies 91*(4).

Hazlett, T. W., S. Oh, and B. Skorup (2018). Mobile phone regulation: the effects of prohibiting handset bundling in finland. *Journal of Competition Law & Economics 14*(1), 65–90.

Hoberg, G. and S. K. Moon (2019). The offshoring return premium. *Management Science 65*(6), 2876–2899.

Jha, M., J. Qian, M. Weber, and B. Yang (2023). Chatgpt and corporate policies. *Chicago Booth Research Paper* (23-15).

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

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

Klafka, J. and A. Ettinger (2020). Spying on your neighbors: Fine-grained probing of contextual embeddings for information about surrounding words. *arXiv preprint arXiv:2005.01810*.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Li, Q., H. Shan, Y. Tang, and V. Yao (2024). Corporate climate risk: Measurements and responses. *The Review of Financial Studies 37*(6), 1778–1830.

Liu, H., R. Ning, Z. Teng, J. Liu, Q. Zhou, and Y. Zhang (2023). Evaluating the logical reasoning ability of chatgpt and gpt-4. *arXiv preprint arXiv:2304.03439*.

Lopez-Lira, A. and Y. Tang (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *arXiv preprint arXiv:2304.07619*.

Loughran, T. and B. McDonald (2014). Measuring readability in financial disclosures. *The Journal of Finance 69*(4), 1643–1671.

Pástor, L. and P. Veronesi (2013). Political uncertainty and risk premia. *Journal of Financial Economics 110*(3), 520–545.

Peltzman, S. (1976). Toward a more general theory of regulation. *The Journal of Law and Economics 19*(2), 211–240.

Peterson, K., R. Schmardebeck, and T. J. Wilks (2015). The earnings quality and information processing effects of accounting consistency. *The Accounting Review 90*(6), 2483–2514.

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

Sautner, Z., L. Van Lent, G. Vilkov, and R. Zhang (2023a). Firm-level climate change exposure. *The Journal of Finance 78*(3), 1449–1498.

Sautner, Z., L. Van Lent, G. Vilkov, and R. Zhang (2023b). Pricing climate change exposure. *Management Science 69*(12), 7540–7561.

Skinner, A. N. and K. Valentine (2023). Green patenting and voluntary innovation disclosure. *Available at SSRN 4321932*.

Stein, L. C. and E. Stone (2013). The effect of uncertainty on investment, hiring, and r&d: Causal evidence from equity options. *Hiring, and R&D: Causal Evidence from Equity Options (October 4, 2013)*.

Stern, N. and A. Valero (2021). Innovation, growth and the transition to net-zero emissions. *Research Policy 50*(9), 104293.

Toole, A., N. Pairolero, A. Giczy, J. Forman, C. Pulliam, M. Such, and B. Rifkin (2020). Inventing ai: Tracing the diffusion of artificial intelligence with us patents. *US Patent and Trademark Office, Alexandria* (5), 2020.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin (2017). Attention is all you need. *Advances in Neural Information Processing Systems 30*.

Yang, S. (2023). Predictive patentomics: Forecasting innovation success and valuation with chatgpt. *Available at SSRN 4482536*.

Zhang, Y., A. Ni, T. Yu, R. Zhang, C. Zhu, B. Deb, A. Celikyilmaz, A. H. Awadallah, and D. Radev (2021). An exploratory study on long dialogue summarization: What works and what's next. *arXiv preprint arXiv:2109.04609*.

Appendix A. Example of SK Telecom

The following is an excerpt from the GPT-processed political risk exposure assessment from SK Telecom's 2018 Q1 earnings call transcript.

The firm is subject to political and regulatory risks and uncertainties related to its 5G CapEx plans and the discussions within the National Assembly regarding the separation of handset contracts and telecom service contracts.

[...] Regarding the discussions on the separation of handset contracts and telecom service contracts, the company states that it is difficult to predict whether it will be introduced and the impact it may have. However, the company believes that if introduced, it could resolve issues related to the blending of service fees and handset pricing. The company also expects a shift in the market competition paradigm towards prices and service quality, leading to greater user benefits. The company emphasizes the need for measures that minimize negative effects on the overall ecosystem, including the distribution structure. The company looks forward to a reasonable conclusion being generated through the discussion and reviewing process, with a focus on market competition and enhancing customer convenience.

[...] The analyst mentioned the government's talk about the separation of device contracts and service contracts, which could potentially affect the company's business model. However, the executives mentioned that there are still ongoing discussions and it is uncertain what will happen. The company has its own stance on the issue but is open to various options before making a decision. This indicates that the company may be affected by new regulations in the future, but the extent of the impact is uncertain at this point.

Appendix B. Sample Snippets

B1. Political Risk Summary

The company is subject to political and regulatory risks and uncertainties in Europe and North America. The recent government auction of HS1 in the U.K. is mentioned as an example.

B2. Political Risk Assessment

The firm is subject to political and regulatory risks and uncertainties in Europe and North America. The focus on deficit reduction in these regions may lead to an increased flow of government disposals and potentially PFI (Private Finance Initiative) opportunities. The recent government auction of HS1 in the U.K. is mentioned as an example. Additionally, the flow of non-core disposals by corporate and financial institutions is continuing, as evidenced by the firm's recent investment in Eversholt, which was purchased from HSBC. These political and regulatory factors could impact the firm's operations and investment opportunities in these regions.

B3. Climate Change Risk Summary

NA

B4. Climate Change Risk Assessment

It is worth noting that the company's use of leading-edge technologies in wafers, silicon wafers, substrates, and packaging may have implications for their environmental footprint. These technologies often require significant energy and resource consumption during production and may generate electronic waste at the end of their lifecycle. Additionally, as the company's networking products are being used in the creation of Ethernet fabric for AI clusters, there may be indirect environmental risks associated with the energy consumption and carbon footprint of these clusters. It is important for the company to consider the sustainability of its networking products and ensure they are aligned with environmental regulations and standards.

B5. AI-Related Risk Summary

Based on the given information, the firm is heavily dependent on AI technologies and is actively incorporating AI into every layer of its stack, including productivity and consumer services. The executives mention that they believe the next big platform wave is AI and that they are working on building training supercomputers and inference infrastructure. They also mention specific AI capabilities in their products, such as robotic process automation and workflow automation, as well as the incorporation of AI in their consumer services. The firm has a partnership with OpenAI and is excited about their innovation and commercialization of products.

B6. AI-Related Risk Assessment

Based on the given information, it is clear that Microsoft is heavily invested in emerging technologies, particularly artificial intelligence (AI). The company highlights its leadership in the AI era and its commitment to developing AI-powered products and services. Microsoft's Azure platform is being used by customers and partners to train state-of-the-art AI models and services, and the company is positioning itself as a leader in AI with its powerful AI supercomputing infrastructure. Additionally, Microsoft's AI services, such as Azure ML, have seen significant revenue growth, indicating a strong demand for AI capabilities. The company is also leveraging AI in its developer tools, such as GitHub Copilot, which is an AI-powered product that transforms developer productivity. Furthermore, Microsoft is integrating AI into its business applications, such as Dynamics 365, to help businesses digitize their operations and improve efficiency. Overall, Microsoft's business is heavily dependent on AI technologies, and the company is at the forefront of AI innovation.

Figure 1. Time Trend in SK Telecom's Political Risk Exposure

This figure shows the time trend in SK Telecom's political risk exposure from 2018 to 2022. The solid line represents GPT-based risk exposure assessment (*PRiskAssess*) and the dotted line represents the bigram-based risk exposure score by Hassan et al. (2019) (*PRiskBigram*).
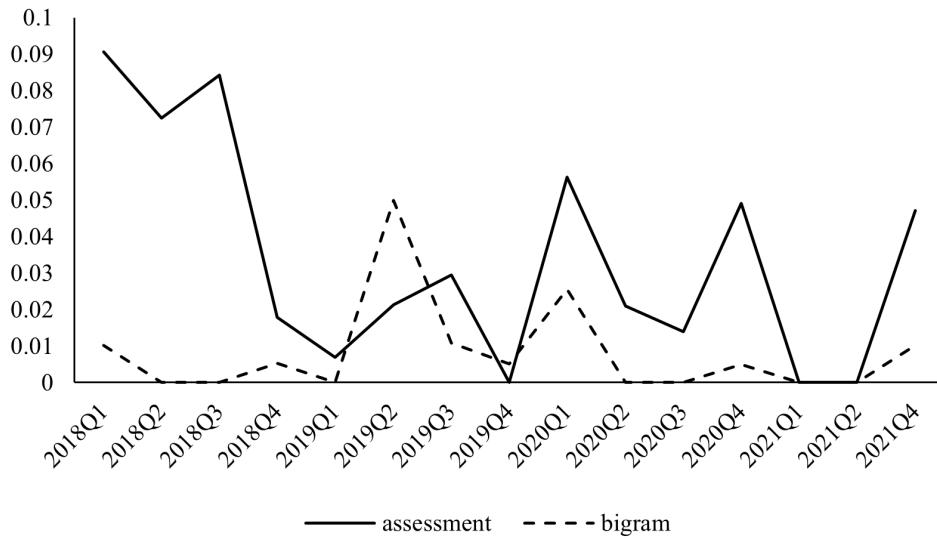
Figure 1. Time Trend in SK Telecom's Political Risk Exposure

Figure 2. Measuring Risks with Generative AI

This figure summarizes how we process earnings call transcripts with GPT to generate firm-level exposure measures. Refer to Section 3 for detailed explanations.

Figure 2. Measuring Risks with Generative AI

44
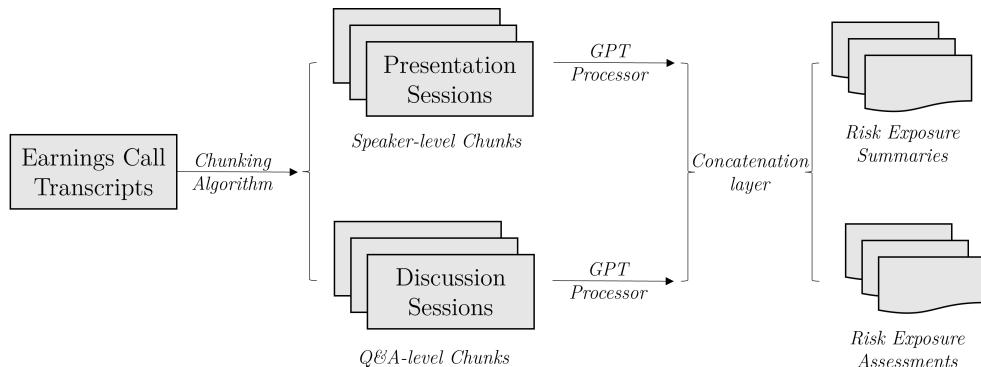
Figure 3. Industry Averages of Risk Exposure Assessments

This figure shows the SIC two-digit level industry averages of GPT-based risk exposure assessments (RiskAssess). We regress *RiskAssess* on dummy variables that represent each industry and report the eight industries with the largest coefficients. We also plot 95% confidence intervals. Standard errors are clustered at the firm-level. Figure 3A shows industry-level averages of political risk exposure, Figure 3B shows averages for climate change risk exposure, and Figures 3C shows averages for AI-related risk exposure. Variables are defined in Section 3.
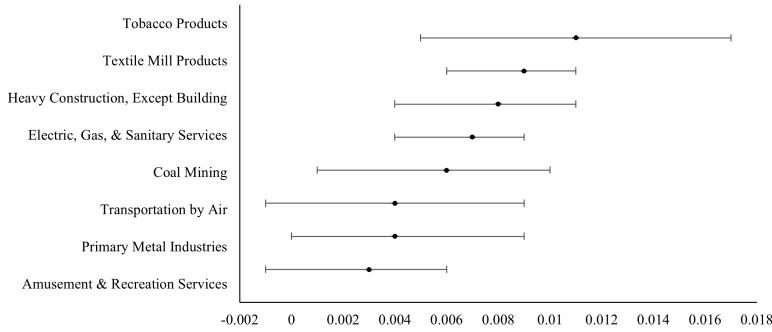
Figure 3A. Industry Averages of *PRiskAssess*
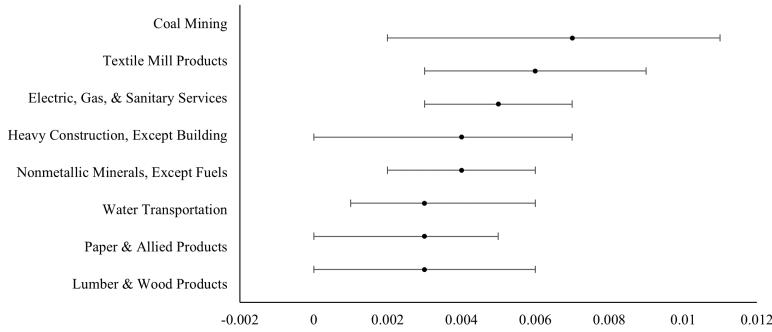
Figure 3B. Industry Averages of *CRiskAssess*
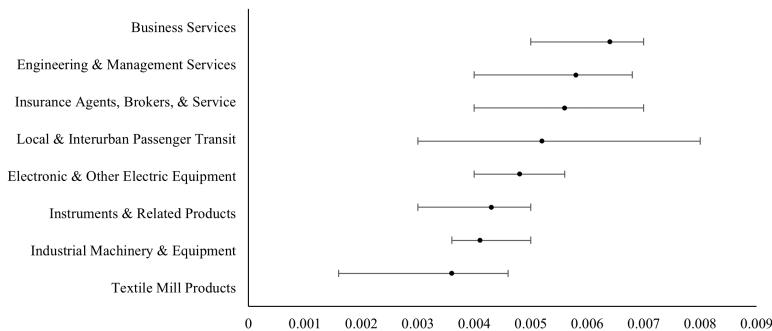
Figure 3C. Industry Averages of *AIRiskAssess*

45

Figure 4. Time Trend of Risk Measures

This figure shows the time series variation in firm-level risk exposure measures. Figure 4A shows political risk exposure measures and Figure 4B shows climate change risk exposure measures. Besides the GPT-based summary and GPT-based assessment measure, we also provide the bigram measure. Figure 4C shows AI-related risk exposure measures. For 4C only, we include the GPT-based summary measure and GPT-based assessment measure. Shaded areas denote notable economy-wide events related to each risk. Variables are defined in Section 3.
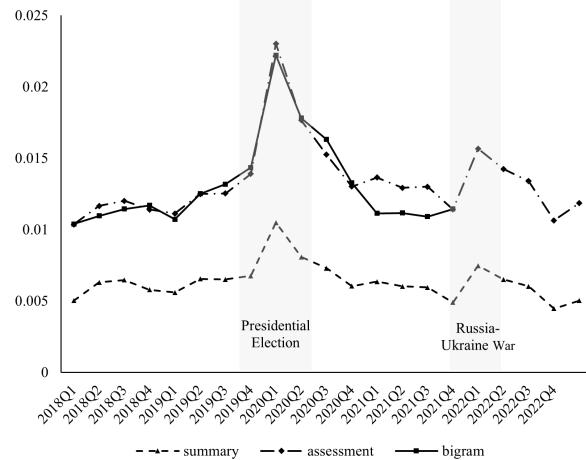
Figure 4A. Time Trend of *PRiskAssess*
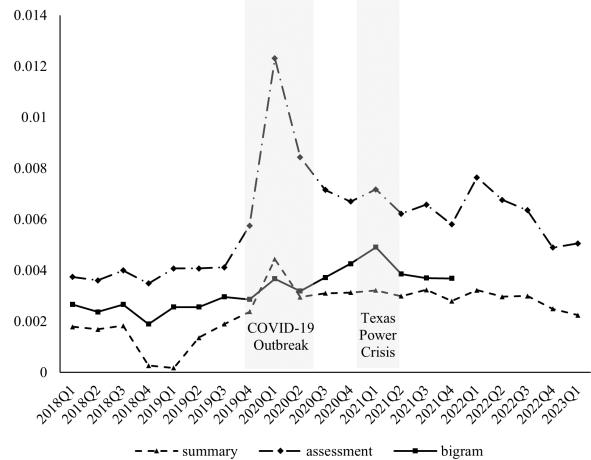
Figure 4B. Time Trend of *CRiskAssess*
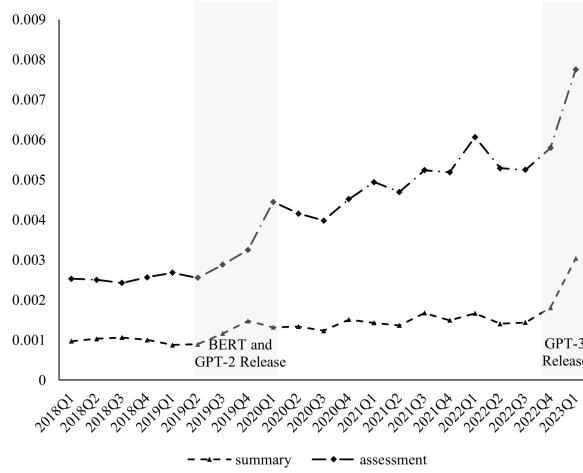
Figure 4C. Time Trend of *AIRiskAssess*

46

Figure 5. Word Clouds

This figure shows the word clouds extracted from the underlying documents of *PRiskAssess*, *CRiskAssess*, and *AIRiskAssess* (defined in Section 3).

Figure 5A. Word Cloud of *PRiskAssess*

Figure 5B. Word Cloud of *CRiskAssess*
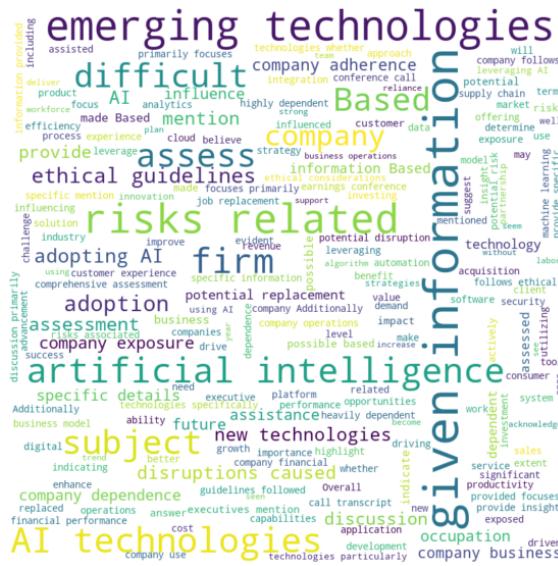
Figure 5C. Word Cloud of *AIRiskAssess*

Figure 6. Aggregated Risk Measures and Volatility

This figure shows the univariate associations between quarterly political risk exposure measures and volatility variables (Panels A and B), and EPU scores and volatility variables (Panels C and D). Variables are defined in Section 8.
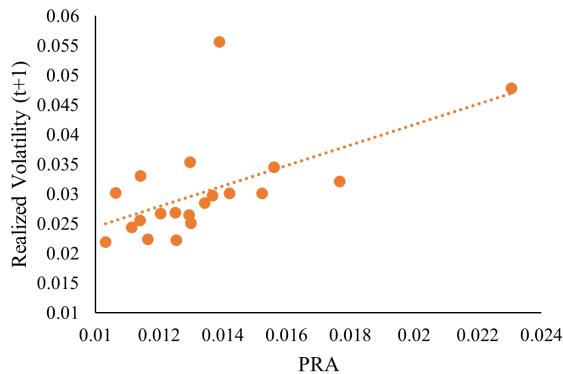
Figure 7A. PRA^{Agg} and Realized Volatility
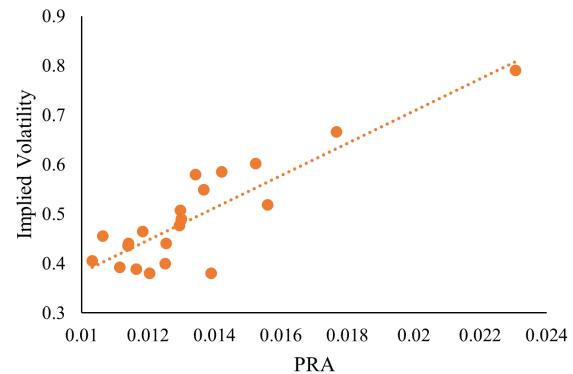
Figure 7B. PRA^{Agg} and Implied Volatility
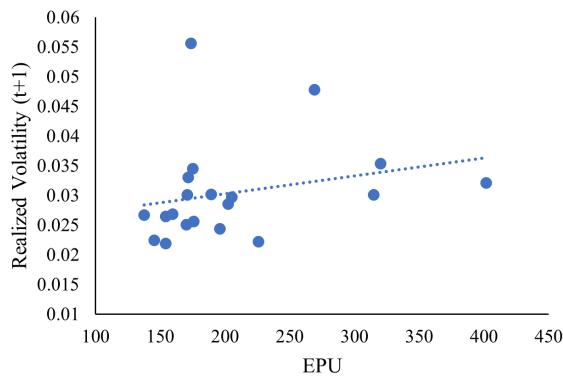
Figure 7C. EPU and Realized Volatility
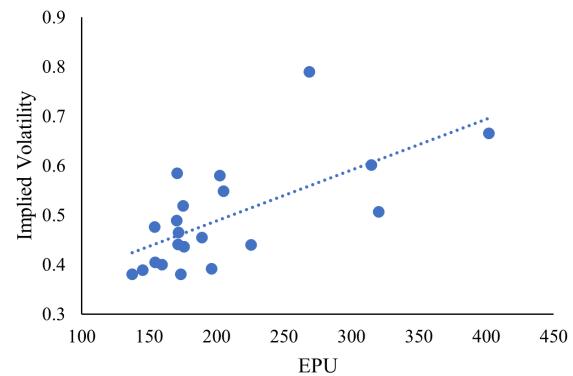
Figure 7D. EPU and Implied Volatility

48

Figure 7. Future vs. Past Volatility

We regress realized volatility variables on PRA^{Agg} and EPU. We start from past $(t-4)$ realized volatility and end with future $(t+4)$ realized volatility. Panel A plots coefficient values on PRA^{Agg} and EPU. Panel B plots t-statistics clustered by quarter of each coefficient. Variables are defined in Section 8.
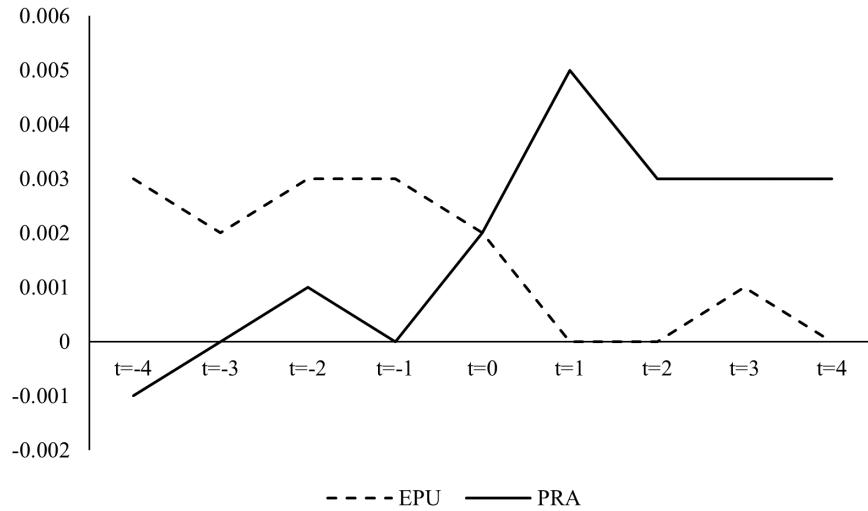
Figure 8A. Coefficient Values
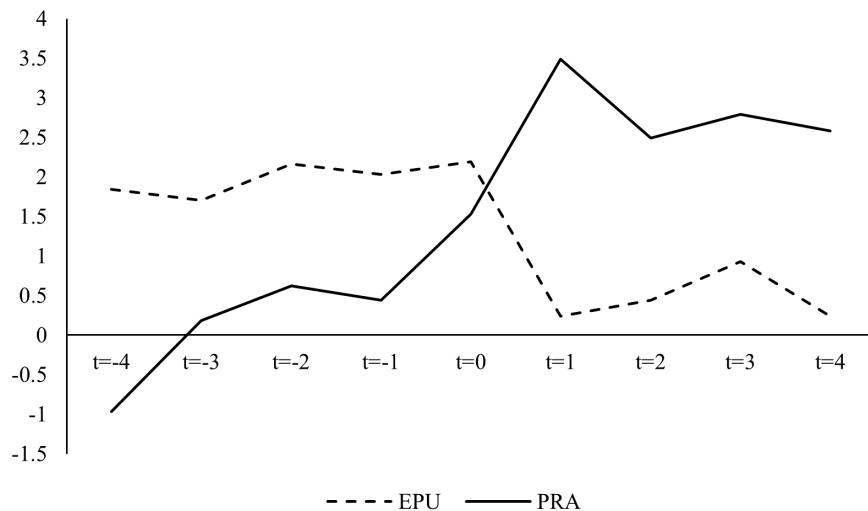
Figure 8B. t-Statistics

49

Table 1. Descriptive Statistics

This table reports the descriptive statistics for the GPT-generated firm-level risk exposure variables and the key dependent variables for our main sample from 2018 to 2021. Variables are defined in Section 3. Panel A reports summary statistics. Panel B reports Pearson correlations among the variables. *PRB* is an abbreviation for *PRiskBigram*, *PRS* for *PRiskSum*, *PRA* for *PRiskAssess*, *CRB* for *CRiskBigram*, *CRS* for *CRiskSum*, *CRA* for *CRiskAssess*, *AIRS* for *AIRiskSum*, and *AIRA* for *AIRiskAssess*.

Panel A. Summary Statistics

	Mean	Std	P25	P50	P75
PRiskSum	0.006	0.009	0.000	0.002	0.011
PRiskAssess	0.013	0.013	0.000	0.011	0.020
CRiskSum	0.002	0.005	0.000	0.000	0.002
CRiskAssess	0.006	0.009	0.000	0.001	0.009
AIRiskSum	0.001	0.004	0.000	0.000	0.000
AIRiskAssess	0.004	0.006	0.000	0.000	0.007
Implied Volatility	0.488	0.268	0.300	0.415	0.591
Abnormal Volatility	-0.124	0.482	-0.399	-0.234	-0.010
Investments	1.083	0.733	0.313	1.238	1.449
1 ($ Lobby Amount > 0)	0.135	0.342	0.000	0.000	0.000
1 (# Green Patent > 0)	0.121	0.326	0.000	0.000	0.000
1 (# AI Patent > 0)	0.104	0.305	0.000	0.000	0.000

Panel B. Correlations

	PRB	*PRS*	*PRA*	*CRB*	*CRS*	*CRA*	*AIRS*	*AIRA*
PRB	1.000							
PRS	0.116	1.000						
PRA	0.120	0.742	1.000					
CRB	0.011	0.139	0.150	1.000				
CRS	0.036	0.349	0.364	0.246	1.000			
CRA	0.047	0.415	0.483	0.260	0.661	1.000		
AIRS	0.004	0.050	0.060	-0.014	0.071	0.049	1.000	
AIRA	-0.013	0.044	0.088	-0.006	0.069	0.055	0.512	1.000

Table 2. Variance Decomposition

This table the variance decomposition results for each risk exposure variable for our full sample from 2018 to March 2023. We report political risk exposure measures in Panel A, climate change risk exposure measures in Panel B, and AI-related risk exposure measures in Panel C. The first sub-panel reports incremental R-squared values by adding time, industry, and time×industry fixed effects. The second sub-panel reports R-squared values by adding firm fixed effects. Variables are defined in Section 3.

Panel A. Political Risk

	PRiskSum	*PRiskAssess*	*PRiskBigram*
Panel A1.	Incremental R^2	Incremental R^2	Incremental R^2
Time FE	1.94%	4.51%	3.20%
Industry FE	3.51%	4.64%	3.50%
Time×Industry FE	2.79%	2.82%	1.96%
Implied Firm Level Variation	91.76%	88.03%	91.34%
Sum	100.00%	100.00%	100.00%
Panel A2.	R^2	R^2	R^2
Firm FE	32.09%	35.58%	30.79%
Remaining Variation	67.91%	64.42%	69.21%
Sum	100.00%	100.00%	100.00%

Panel B. Climate Change Risk

	CRiskSum	*CRiskAssess*	*CRiskBigram*
Panel B1.	Incremental R^2	Incremental R^2	Incremental R^2
Time FE	4.05%	5.82%	0.30%
Industry FE	5.82%	8.44%	7.86%
Time×Industry FE	4.10%	3.48%	2.36%
Implied Firm Level Variation	86.03%	82.26%	89.48%
Sum	100.00%	100.00%	100.00%
Panel B2.	R^2	R^2	R^2
Firm FE	24.13%	32.28%	32.36%
Remaining Variation	75.87%	67.72%	67.64%
Sum	100.00%	100.00%	100.00%

Panel C. AI Exposure Risk

	AIRiskSum	*AIRiskAssess*	*AIRiskBigram*
Panel C1.	Incremental R^2	Incremental R^2	Incremental R^2
Time FE	1.07%	5.00%	-
Industry FE	3.37%	9.67%	-
Time×Industry FE	2.99%	2.73%	-
Implied Firm Level Variation	92.57%	82.60%	-
Sum	100.00%	100.00%	
Panel C2.	R^2	R^2	R^2
Firm FE	21.31%	31.31%	-
Remaining Variation	78.69%	68.69%	-
Sum	100.00%	100.00%	-

51

Table 3. Political Risk Exposure and Volatility

This table reports the association between firm-level political risk exposure measures and volatility using our main sample period from 2018 to 2021. We use industry and time fixed effects in Panel A, and industry×time fixed effects in Panel B. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Industry and Time Fixed Effects

Dep Var = Implied Volatility

	(1)	(2)	(3)	(4)	(5)
PRiskSum	0.888***		0.857***		0.160
	(4.50)		(4.37)		(0.73)
PRiskAssess		0.777***		0.757***	0.676***
		(5.08)		(4.97)	(3.72)
PRiskBigram			0.015	0.014	0.014
			(1.54)	(1.43)	(1.40)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time & Ind	Time & Ind	Time & Ind	Time & Ind	Time & Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.477	0.477	0.477	0.477	0.477

Dep Var = Abnormal Volatility

	(1)	(2)	(3)	(4)	(5)
PRiskSum	1.576***		1.574***		-1.031***
	(6.55)		(6.54)		(-2.96)
PRiskAssess		1.987***		1.995***	2.523***
		(11.42)		(11.44)	(10.00)
PRiskBigram			0.001	-0.006	-0.004
			(0.07)	(-0.48)	(-0.34)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time & Ind	Time & Ind	Time & Ind	Time & Ind	Time & Ind
N	39276	39276	39276	39276	39276
Adjusted R^2	0.353	0.355	0.353	0.355	0.355

Panel B. Industry×Time Fixed Effects

Dep Var = Implied Volatility

	(1)	(2)	(3)	(4)	(5)
PRiskSum	0.737***		0.704***		0.154
	(3.64)		(3.49)		(0.69)
PRiskAssess		0.635***		0.614***	0.535***
		(4.07)		(3.95)	(2.92)
PRiskBigram			0.016	0.015	0.015
			(1.60)	(1.52)	(1.50)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time × Ind	Time × Ind	Time × Ind	Time × Ind	Time × Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.477	0.477	0.477	0.477	0.477

Dep Var = Abnormal Volatility

	(1)	(2)	(3)	(4)	(5)
PRiskSum	1.559***		1.559***		-0.659**
	(6.62)		(6.61)		(-1.96)
PRiskAssess		1.808***		1.815***	2.152***
		(10.51)		(10.52)	(8.75)
PRiskBigram			0.000	-0.005	-0.004
			(0.02)	(-0.43)	(-0.34)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time × Ind	Time × Ind	Time × Ind	Time × Ind	Time × Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.477	0.477	0.477	0.477	0.477

Table 4. Climate Change Risk Exposure and Volatility

This table reports the association between firm-level climate change risk exposure measures and volatility using our main sample period from 2018 to 2021. We use industry and time fixed effects in Panel A, and industry×time fixed effects in Panel B. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Industry and Time Fixed Effects

Dep Var = Implied Volatility

	(1)	(2)	(3)	(4)	(5)
CRiskSum	0.807**		0.769**		-0.576*
	(2.51)		(2.42)		(-1.72)
CRiskAssess		1.075***		1.074***	1.281***
		(4.51)		(4.53)	(4.74)
CRiskBigram			0.069	-0.316	-0.251
			(0.07)	(-0.34)	(-0.27)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time & Ind	Time & Ind	Time & Ind	Time & Ind	Time & Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.476	0.477	0.476	0.477	0.477

Dep Var = Abnormal Volatility

	(1)	(2)	(3)	(4)	(5)
CRiskSum	1.693***		1.502***		-1.401**
	(3.72)		(3.28)		(-2.41)
CRiskAssess		2.332***		2.269***	2.752***
		(9.32)		(8.85)	(8.49)
CRiskBigram			1.678**	0.807	0.967
			(2.27)	(1.08)	(1.29)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time & Ind	Time & Ind	Time & Ind	Time & Ind	Time & Ind
N	39276	39276	39276	39276	39276
Adjusted R^2	0.353	0.354	0.353	0.354	0.354

Panel B. Industry×Time Fixed Effects

Dep Var = Implied Volatility

	(1)	(2)	(3)	(4)	(5)
CRiskSum	0.676**		0.660**		-0.500
	(2.00)		(1.99)		(-1.45)
CRiskAssess		0.915***		0.924***	1.102***
		(3.70)		(3.76)	(3.97)
CRiskBigram			-0.077	-0.396	-0.344
			(-0.08)	(-0.41)	(-0.35)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time × Ind	Time × Ind	Time × Ind	Time × Ind	Time × Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.489	0.489	0.489	0.489	0.489

Dep Var = Abnormal Volatility

	(1)	(2)	(3)	(4)	(5)
CRiskSum	1.262***		1.019**		-1.906***
	(2.91)		(2.33)		(-3.45)
CRiskAssess		2.216***		2.116***	2.768***
		(9.20)		(8.58)	(8.90)
CRiskBigram			2.253***	1.355*	1.557**
			(3.11)	(1.87)	(2.15)
Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Time × Ind	Time × Ind	Time × Ind	Time × Ind	Time × Ind
N	35003	35003	35003	35003	35003
Adjusted R^2	0.455	0.457	0.456	0.457	0.457

Table 5. AI-related Risk Exposure and Volatility

This table reports the association between firm-level AI-related risk exposure measures and volatility using our main sample period from 2018 to 2021. We use industry and time fixed effects in Panel A, and industry×time fixed effects in Panel B. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Industry and Time Fixed Effects

	Implied Volatility			Abnormal Volatility		
	(1)	(2)	(3)	(4)	(5)	(6)
AIRiskSum	-0.457		-0.444	1.738***		1.853***
	(-1.38)		(-1.34)	(3.19)		(3.01)
AIRiskAssess		-0.170	-0.019		0.459	-0.166
		(-0.60)	(-0.07)		(1.22)	(-0.39)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Time*Ind FE	No	No	No	No	No	No
N	35003	35003	35003	39276	39276	39276
Adjusted R^2	0.476	0.476	0.476	0.353	0.352	0.353

Panel B. Industry×Time Fixed Effects

	Implied Volatility			Abnormal Volatility		
	(1)	(2)	(3)	(4)	(5)	(6)
AIRiskSum	-0.520		-0.477	1.153**		1.853***
	(-1.56)		(-1.42)	(2.25)		(2.36)
AIRiskAssess		-0.224	-0.061		0.166	-0.292
		(-0.79)	(-0.20)		(0.46)	(-0.72)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No	No
Industry FE	No	No	No	No	No	No
Time*Ind FE	Yes	Yes	Yes	Yes	Yes	Yes
N	35003	35003	35003	39276	39276	39276
Adjusted R^2	0.489	0.489	0.489	0.455	0.455	0.455

Table 6. Out-of-Sample Analysis: Capital Market Consequences

This table repeats the analysis of Tables 3, 4, and 5 in the post-GPT training sample from 2022 to March 2023. We report political risk exposure measures in Panel A, climate change risk exposure measures in Panel B, and AI-related risk exposure measures in Panel C. We use industry and time fixed effects throughout the analyses. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Political Risks

	Implied Volatility		Abnormal Volatility	
	(1)	(2)	(3)	(4)
PRiskSum	1.716***		0.430	
	(5.04)		(1.09)	
PRiskAssess		1.246***		0.941***
		(5.12)		(3.44)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	9923	9923	9246	9246
Adjusted R^2	0.423	0.423	0.115	0.117

Panel B. Climate Change Risks

	Implied Volatility		Abnormal Volatility	
	(1)	(2)	(3)	(4)
CRiskSum	2.072***		0.276	
	(4.26)		(0.46)	
CRiskAssess		2.135***		1.494***
		(6.14)		(3.76)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	9923	9923	9246	9246
Adjusted R^2	0.421	0.424	0.115	0.117

Panel C. AI-Related Risk

	Implied Volatility		Abnormal Volatility	
	(1)	(2)	(3)	(4)
AIRiskSum	0.775*		2.281***	
	(1.85)		(3.69)	
AIRiskAssess		1.166***		1.259***
		(3.53)		(2.76)
Controls	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	9923	9923	9246	9246
Adjusted R^2	0.420	0.421	0.116	0.116

Table 7. Risk Exposure and Investments

This table reports the association between firm-level risk exposure variables and capital investments using our main sample period from 2018 to 2021. We report political risk exposure measures in Panel A, climate change risk exposure measures in Panel B, and AI-related risk exposure measures in Panel C. We use capital expenditure scaled by recursive total capital as the dependent variable. We use industry and time fixed effects. All continuous variables are winsorized at 1% and 99% levels. Variables are defined in Section 3 and Section 4. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Political Risk

Dep Var	Investment				
	(1)	(2)	(3)	(4)	(5)
PRiskSum	-0.792**		-0.792**		0.317
	(-2.08)		(-2.06)		(0.58)
PRiskAssess		-0.918***		-0.922***	-1.081***
		(-3.33)		(-3.32)	(-2.72)
PRiskBigram			-0.000	0.003	0.002
			(-0.01)	(0.13)	(0.11)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
N	35615	35615	35615	35615	35615
Adjusted R^2	0.279	0.280	0.279	0.280	0.280

Panel B. Climate Change Risk

Dep Var	Investment				
	(1)	(2)	(3)	(4)	(5)
CRiskSum	-0.606		-0.524		0.924
	(-0.90)		(-0.77)		(1.08)
CRiskAssess		-1.077***		-1.059**	-1.375***
		(-2.62)		(-2.54)	(-2.64)
CRiskBigram			-0.865	-0.392	-0.502
			(-0.86)	(-0.39)	(-0.50)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
N	35615	35615	35615	35615	35615
Adjusted R^2	0.279	0.280	0.279	0.279	0.279

Panel C. AI-Related Risk

Dep Var	Investment				
	(1)	(2)	(3)	(4)	(5)
AIRiskSum	0.694		-		0.231
	(0.84)				(0.25)
AIRiskAssess		0.751		-	0.674
		(1.28)			(1.02)
AIRiskBigram			-	-	-
Controls	Yes	Yes			Yes
Time FE	Yes	Yes			Yes
Ind FE	Yes	Yes			Yes
N	35615	35615			35615
Adjusted R^2	0.279	0.280			0.279

Table 8. Risk Exposure and Firm Responses

This table reports how firms respond to different risk exposures using our main sample period from 2018 to 2021. We report political risk exposure measures in Panel A, climate change risk exposure measures in Panel B, and AI-related risk exposure measures in Panel C. For Panel A, we use a lobbying activity indicator as the dependent variable. For Panel B, we use a green patent filing indicator and, for Panel C, we use an AI patent filing indicator as the dependent variable. We use industry and time fixed effects throughout the analyses. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Political Risk

Dep Var	$\mathbf{1}$ ($ Lobby Amount > 0)				
	(1)	(2)	(3)	(4)	(5)
PRiskSum	0.057		0.142		-1.258***
	(0.18)		(0.46)		(-4.19)
PRiskAssess		0.489**		0.437*	1.079***
		(2.04)		(1.84)	(4.21)
PRiskBigram			0.041**	0.037*	0.039*
			(2.04)	(1.83)	(1.94)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
N	39937	39937	39937	39937	39937
Adjusted R^2	0.191	0.191	0.191	0.191	0.192

Panel B. Climate Change Risk

Dep Var	$\mathbf{1}$ (# Green Patent > 0)				
	(1)	(2)	(3)	(4)	(5)
CRiskSum	0.271		0.089		-0.757*
	(0.64)		(0.22)		(-1.77)
CRiskAssess		0.636**		0.542**	0.802***
		(2.25)		(2.01)	(2.70)
CRiskBigram			2.141	1.847	1.934
			(1.47)	(1.28)	(1.34)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes
N	39937	39937	39937	39937	39937
Adjusted R^2	0.254	0.253	0.253	0.254	0.254

Panel C. AI-Related Risk

Dep Var	$\mathbf{1}$ (# AI Patent > 0)				
	(1)	(2)	(3)	(4)	(5)
AIRiskSum	4.085***		-		1.723***
	(5.83)				(2.69)
AIRiskAssess		3.984***		-	3.405***
		(7.72)			(6.96)
AIRiskBigram			-	-	-
Controls	Yes	Yes			Yes
Time FE	Yes	Yes			Yes
Ind FE	Yes	Yes			Yes
N	39937	39937			39937
Adjusted R^2	0.231	0.233			0.233

Table 9. Out-of-Sample Analysis: Economic Outcomes

This table repeats the analysis of Tables 7 and 8 in the post-GPT training sample from 2022 to March 2023. We replicate the findings of Table 7 in Panel A and Table 8 in Panel B. We use industry and time fixed effects. Variables are defined in Section 3 and Section 4. All continuous variables are winsorized at 1% and 99% levels. Standard errors are clustered at the firm-level. ***, **, and * denote significance at 1%, 5%, and 10% levels, respectively.

Panel A. Risk Exposure and Investments

	Capital Expenditure					
	(1)	(2)	(3)	(4)	(5)	(6)
PRiskSum	-1.838**					
	(-2.50)					
PRiskAssess		-1.533***				
		(-3.04)				
CRiskSum			-1.304			
			(-1.16)			
CRiskAssess				-1.659**		
				(-2.24)		
AIRiskSum					2.115*	
					(1.89)	
AIRiskAssess						2.604***
						(3.25)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	9949	9949	9949	9949	9949	9949
Adjusted R^2	0.340	0.340	0.339	0.340	0.340	0.340

Panel B. Risk Exposure and Firm Responses

	1($ Lobby Amount > 0)		**1**(# Green Patent > 0)		**1**(# AI Patent > 0)	
	(1)	(2)	(3)	(4)	(5)	(6)
PRiskSum	0.633*					
	(1.75)					
PRiskAssess		0.935***				
		(3.69)				
CRiskSum			0.457			
			(1.31)			
CRiskAssess				0.473**		
				(2.49)		
AIRiskSum					0.272	
					(0.90)	
AIRiskAssess						0.547***
						(2.89)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11148	11148	11148	11148	11148	11148
Adjusted R^2	0.197	0.198	0.089	0.090	0.067	0.068

Table 10. Aggregated Risk Exposure Index

This table reports the association between aggregated political risk exposure index PRA^{Agg}, economic policy uncertainty index EPU, and economic variables. We construct our aggregate political risk measure by taking the simple average of PRA scores of all firms in the given quarterly. Panel A examines volatility and Panel B examines real effects. All variables are averaged at the quarter level and are defined in Section 8. All continuous variables are winsorized at 1% and 99% levels. t-values are clustered at the quarter-level.

Panel A. Volatility

Dep Var =	Realized Volatility ($t+1$)			Implied Volatility		
	(1)	(2)	(3)	(4)	(5)	(6)
EPU	0.002		-0.001	0.072***		0.028***
	(1.64)		(-0.42)	(4.11)		(2.95)
PRA^{Agg}		0.005***	0.005***		0.092***	0.076***
		(6.98)	(4.48)		(14.14)	(9.63)
Controls	No	No	No	No	No	No
Fixed Effects	No	No	No	No	No	No
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter	Quarter
N	20	20	20	21	21	21
Adjusted R^2	0.025	0.316	0.282	0.434	0.749	0.789

Panel B. Real Effects

Dep Var =	Capital Expenditure			Unemployment ($t+1$)		
	(1)	(2)	(3)	(4)	(5)	(6)
EPU	-0.001		0.001	1.568***		0.620**
	(-0.88)		(0.75)	(3.75)		(2.45)
PRA^{Agg}		-0.003**	-0.003**		2.030***	1.678***
		(-2.72)	(-2.45)		(7.47)	(3.80)
Controls	No	No	No	No	No	No
Fixed Effects	No	No	No	No	No	No
Cluster	Quarter	Quarter	Quarter	Quarter	Quarter	Quarter
N	21	21	21	20	20	20
Adjusted R^2	-0.034	0.098	0.059	0.427	0.751	0.791

Table 11. Asset Pricing on Firm-Level Risk Assessment

This table shows asset-pricing implications for the annualized firm-level risk exposure measures. In Panel A, we present Fama and MacBeth (1973) regression results and Newey-West t-values (with lag=3). We regress monthly stock returns on our firm-level risk exposure measures and the following control variables: stock return for the prior month, stock return for the prior year (skipping the most recent month), log of the market value of equity, log of book-to-market ratio, operating profitability, and investment. Variables are defined in 9. All continuous independent variables are trimmed at 1% and 99%. In Panel B, we form portfolios on March 31 of the subsequent year and delete observations with zero annualized risk exposure. We report quintile portfolio alphas using Fama and French (2015). We then report monthly high-minus-low alphas and their corresponding t-values. t-values are Newey-West adjusted with a lag 3. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Panel A. Fama and MacBeth (1973) Regressions

Dep Var	Return		
	(1)	(2)	(3)
$r_{0,1}$	-0.009	-0.001	-0.010
	(-0.67)	(-0.67)	(-0.72)
$r_{2,12}$	0.002	0.002	0.002
	(0.47)	(0.49)	(0.33)
$\log(ME)$	-0.000	-0.000	-0.000
	(-0.34)	(-0.33)	(-0.32)
$\log(BE/ME)$	-0.001	-0.001	-0.001
	(-0.54)	(-0.56)	(-0.44)
$Profitability$	0.001	0.001	0.001
	(0.54)	(0.60)	(0.62)
$Investment$	-0.006	-0.008	-0.005
	(-0.65)	(-0.80)	(-0.55)
$PRiskAssess^{ann}$	0.077		
	(1.36)		
$CRiskAssess^{ann}$		0.211*	
		(1.98)	
$AIRiskAssess^{ann}$			0.317*
			(1.72)
Adjusted R^2	0.054	0.054	0.055

Panel B. Single Sorts

Sorts On =	Fama and French (2015) Five-Factor Alphas		
	$PRiskAssess^{ann}$	$CRiskAssess^{ann}$	$AIRiskAssess^{ann}$
	(1)	(2)	(3)
Low	-0.06	-0.14	-0.11
2	-0.05	-0.06	0.23
3	0.25	0.08	0.05
4	0.30	0.21	0.16
High	0.38	0.42	0.42
High – Low	0.44	0.56*	0.53**
t-value	(1.51)	(1.90)	(2.31)

60