Real-Time Discovery of Corporate Risks

Mickenzie Bass and Aakash Kalyani

Abstract

We propose a new methodology to discover emerging corporate risks in real time by analyzing the text of quarterly earnings conference calls from 2008 to 2025. Our approach identifies bigrams (two-word phrases) within risk-related sentences whose usage surges significantly and then groups them into thematic topics. The method successfully recovers a timeline of major economic events, from the credit crisis in 2008 to macroeconomic and tariff uncertainty in 2025. We find that firms manage these risks differently. While macroeconomic uncertainty is associated with reductions in investment and employment, a rise in trade uncertainty is associated with capital expenditures and hiring. These expansions, however, are also associated with higher inflation: Higher trade uncertainty is, on average, followed by significant increases in producer prices. Our findings demonstrate that not all uncertainty is alike and suggest that the recent rise of macroeconomic and trade uncertainty together poses a stagflationary risk.

JEL codes: E32, E31, E60, F1

Federal Reserve Bank of St. Louis *Review*, Fourth Quarter 2025, Vol. 107, No. 16, pp. 1-17.
https://doi.org/10.20955/r.2025.16

1. INTRODUCTION

From sudden tariff hikes and disease outbreaks to flare-ups in geopolitical tension, firms confront an ever-shifting menu of risks. Recent research shows that they navigate each threat differently—for example, economic policy uncertainty (Baker, Bloom, and Davis, 2016), COVID-19 uncertainty (Tarek Alexander Hassan et al., 2023), and Brexit uncertainty (Hassan, Hollander, Lent, et al., 2024)—so the macroeconomic impact of risk is anything but uniform. Yet at the aggregate level, economists and practitioners track a single measure of risk, such as the Chicago Board Options Exchange's CBOE Volatility Index (VIX).

In this article, we propose a new methodology to identify emerging risks facing firms by applying natural language processing (NLP) techniques to earnings conference calls. Our methodology is based on identifying terminology that is increasingly used by executives in the context of describing risk. We uncover relevant risk topics over time and also identify how firms manage these risks in a unified framework.

We begin by identifying "risk sentences" in earnings conference calls. Earnings calls are quarterly meetings where executives of publicly traded companies discuss the financials and current operations of their firms and factors that may impact future performance. We parse 209,915 earnings-call transcripts from 6,934 U.S. firms between the second quarter of 2008 and the second quarter of 2025. Following Tarek A Hassan et al. (2019), we denote a sentence as a risk sentence if it contains any of the 136 risk keywords identified in Tarek A

Mickenzie Bass is a research associate and Aakash Kalyani is an economist at the Federal Reserve Bank of St. Louis.

Hassan et al. (2019)'s political risk index, including "gamble," "uncertainty," and "threat." Consistent with their findings, we find that the aggregate incidence of risk sentences in earnings calls is correlated with the VIX and are countercyclical; that is, they spike during recessions. These risks have also recently been at record highs during 2025:Q2.

Next, we represent each "risk sentence" by a vector of one- and two-word combinations (unigrams and bigrams). We prefer bigrams as our unit of analysis because they are less ambiguous than single-word keywords. For example, while words such as "debt" or "government" could have various meanings, "government debt" is much less ambiguous (Tan, Wang, and Lee, 2002). We count each risk bigram as an emerging risk bigram if its incidence in risk sentences increases by at least 200 percent compared with the previous quarter. For each quarter, we find bigrams that refer to important historical risk topics. For example, the top emerging risk phrase in 2008 is "credit crisis," in 2013 is "debt ceiling" and "government shutdown," in 2016 is "Brexit," in 2020 is "virus" and "COVID-19," and so on. More recently in 2025:Q2, we find that these risks are about the "economic environment" and "tariffs."

To systematically identify how firms address these risks, we group the bigrams into topics and add additional keywords using a language model. This language model tells us about the semantic similarity between the bigrams based on past usage by executives in similar contexts. We identify ten major sources of risk for U.S. firms emerging between 2008 and 2025: banking, disasters, disease, financial markets, geopolitics, macroeconomy, monetary policy, commodity prices, policy, and trade uncertainty. More recently, the increase in risk has been driven by concerns about trade uncertainty and macroeconomic uncertainty, both of which are at record highs.

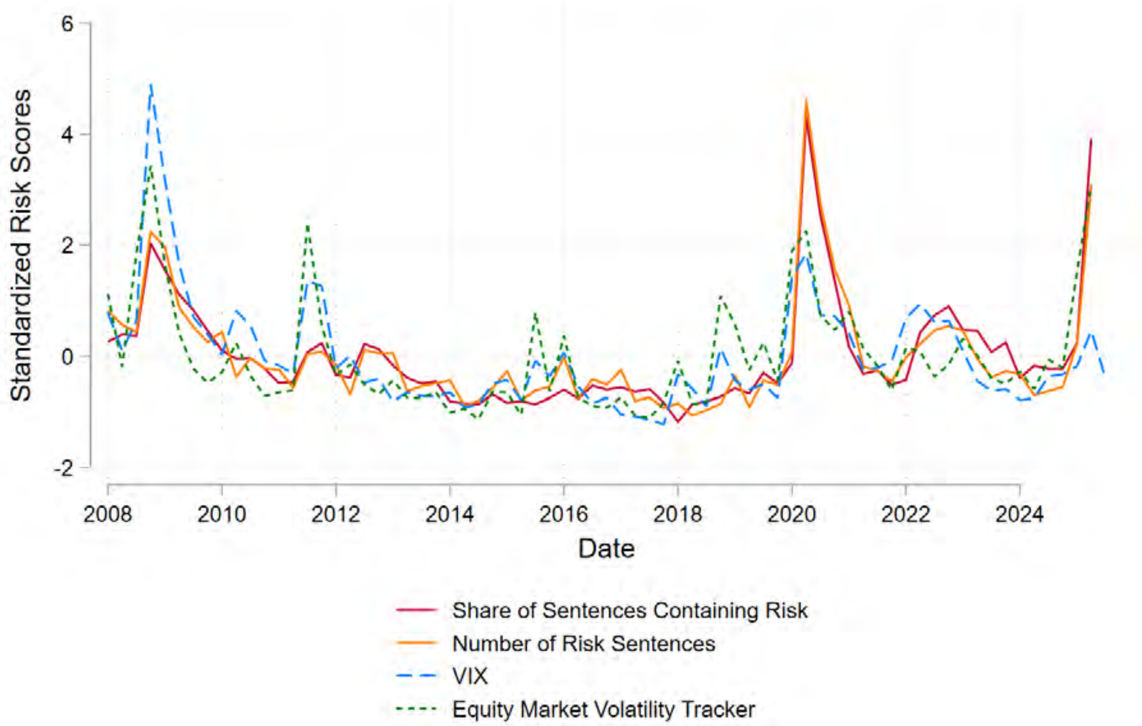
Our results are consistent with those of the vast literature on economic uncertainty, supporting the idea that macroeconomic uncertainty is essentially a contractionary shock. We find that after an increase in complaints about macroeconomic uncertainty, firms tend to reduce investment and hiring. On the other hand, we find contrasting results for trade-related uncertainty. Firms tend to exhibit higher levels of investment and hiring when facing trade risks, possibly to diversify their supply chains. Using variation in trade uncertainty and producer price indices (PPIs) across industries, we also find that an increase in trade uncertainty is associated with an increase in prices, taking about 5–12 quarters for these effects to play out. In our measure, trade uncertainty is a combination of discussions about trade and tariffs. Our results persist when we narrow the analysis to tariff uncertainty rather than considering all sources of trade uncertainty. This evidence supports recent commentary from economists raising concerns about stagflation from the current episode of uncertainty. However, our results suggest that stagflationary risk is a consequence of a combination of rises in macroeconomic and trade uncertainty.

We contribute to two strands of macro-finance literature focused on economic uncertainty. The first strand quantifies economy-wide uncertainty with various proxies—for example, the VIX (a market-based measure of uncertainty), the newspaper-based economic policy uncertainty index of Baker, Bloom, and Davis (2016), the earnings-call-based firm-level political risk index of Tarek A Hassan et al. (2019), or the geopolitical risk index of Caldara and Iacoviello (2022). Building on these aggregates, recent papers construct event-specific proxies—for example, Brexit and the COVID-19 measures of Hassan, Hollander, Lent, et al. (2024, 2023)—yet each one begins with a predefined topic and builds a topic-specific dictionary. We contribute an unsupervised pipeline that flags bigrams whose usage surges, thereby discovering emerging risk narratives as soon as they enter firms' discourse and showing that companies respond to different risks in sharply heterogeneous ways (cutting investment after macro shocks but expanding capacity in the face of supply-chain threats).

A second strand seeks to detect novel topics algorithmically. Kalyani et al. (2023) and Kalyani (2023) detect novel terminology in patents. The closest paper, Hanley and Hoberg (2019), applies latent Dirichlet allocation (LDA) plus semantic–vector aggregation to bank 10-Ks (annual financial reports) and then labels a theme "emerging" if it raises stock-return covariance. However, standard topic-model preprocessing drops infrequent tokens, meaning the newest vocabulary is discarded before analysis, and even large language models like BERT struggle to contextualize rare words (Denny and Spirling, 2018; Schick and Schütze, 2020). By watching relative frequency jumps—and treating entire bigrams as units of analysis—our method considers exactly those rare, newly coined phrases that signal fresh shocks but are dropped by LDA or transformer embeddings.

Finally, several recent reports indicate that uncertainty is significantly higher today. Our real-time indicators translate such qualitative alarms into quantitative forecasts, allowing policymakers and investors to gauge how the current wave of uncertainty is likely to propagate through capital spending, hiring, and credit growth—thereby closing the gap between measured risk and its macro effects.

Figure 1
Percentage of Risk Sentences in a Transcript



NOTE: Following Hassan, Hollander, Kalyani, et al. (2024), the figure displays the standardized share of risk-related sentences in earnings calls for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. It also plots the standardized number of risk sentences in earnings calls. The VIX and the Equity Market Volatility Tracker are both retrieved from FRED and standardized. The share of sentences containing risk is computed as the share of total earnings-call sentences that contain at least one risk keyword.

2. EMERGING RISKS BY YEAR

2.1 Identifying Emerging Risk Language

Our analysis starts with 209,915 quarterly earnings-call transcripts covering 6,934 U.S.-listed firms from 2008:Q2 to 2025:Q2 sourced from S&P Global. Each transcript is sentence-tokenized, and following Hassan, Hollander, Kalyani, et al. (2024), a sentence is labeled a risk sentence whenever it contains at least one of the 136 uncertainty terms in Tarek A Hassan et al. (2019)'s earnings-call-based firm-level political risk index, such as "uncertain," "threat," or "gamble," which explicitly signal risk. Restricting attention to these sentences ensures that we focus on the portions of earnings calls that actually discuss risks.

Figure 1 shows the percentage of risk sentences by quarter. The incidence of risk sentences in earnings calls is almost as high as it was during the beginning of the COVID-19 period (2020:Q2). To show that this increase is not driven by the changing length of earnings calls, the figure also plots the number of risk sentences, showing almost identical patterns. The figure also plots the newspaper-based equity market volatility index by quarter, which shows similar patterns. Both are measures of attention paid to uncertainty by executives and newspapers. In sharp contrast, the VIX does not exhibit the same increase in uncertainty. The VIX is a 30-day equity volatility forecast and essentially captures uncertainty through the expected volatility in equity markets. It does not always capture the level of uncertainty in the economy (Białkowski, Dang, and Wei, 2022) experienced by executives or managers while making investment, employment, and pricing decisions.

From every risk sentence, we extract contiguous, lowercase bigrams (two-word tokens). Bigrams balance context with tractability, but they also contain considerable noise. Therefore, we apply four sequential filters, each chosen to strip away tokens that contribute frequency but no informational value.

We next discard bigrams consisting solely of `spaCy` English stop words—such as "the," "and," or "in"—because stop words carry virtually no semantic content. For the same reason, bigrams containing one stop word are converted into unigrams. Therefore, we treat "of disaster" the same as "a disaster" and convert both to "disaster."

To avoid double-counting the very cues that flagged a sentence as risk-related, we drop bigrams that are made only of those risk-trigger words themselves and stop words (e.g., "risk factor," "the uncertainties"). Finally, we delete purely seasonal markers—month names, quarter labels, and calendar clichés like "first quarter"—since their incidence fluctuates mechanically with the reporting calendar and does not convey new information about uncertainty.

These screens eliminate roughly 41 percent of unique raw bigrams, greatly reducing dimensionality and narrowing our database. For every surviving bigram b in quarter t, we compute its number of transcripts,

$$\#\ \text{transcripts}_{b,t} \ = \ \#\{\text{transcripts that contain } b\},$$

and its quarter-to-quarter log change,

$$\Delta \log \text{Share}_{b,t} \ = \ \left\{\log\left(\#\ \text{transcripts}_{b,t} + 1\right) \ - \log\left(\#\ \text{transcripts}_{t}\right)\right\} - \left\{\log\left(\#\ \text{transcripts}_{b,t-1} + 1\right) \ - \log\left(\#\ \text{transcripts}_{t-1}\right)\right\} , \tag{1}$$

where the $log(x + 1)$ is used to offset zeros. $\log\left(\#\ \text{transcripts}_{t}\right)$ captures normalization by the number of total transcripts in a quarter and ensures that we capture spikes in conversation rather than spikes in coverage.

A bigram is labeled emerging in quarter t if it appears in at least 25 distinct transcripts and its share rises by 200 percent or more relative to $t-1$. The 25-transcript floor screens out idiosyncratic bigrams and those of lower importance in that quarter, whereas the 200 percent surge threshold isolates emerging bigrams.

Table 1 shows the quarter-by-quarter phrases that are identified as "emerging" bigrams and reveals a clear timeline of risks that dominated U.S. corporate discourse from the global financial crisis to mid-2025. The series begins with a cluster of four finance-centric bigrams in 2008:Q4—"credit crisis," "crisis," "global financial," "Lehman"—marking the end of the Great Financial Recession. After a span of no emerging risks, worries about sovereign-debt enters in 2011:Q3–Q4, with "debt ceiling" and the Thailand floods, followed by 2012's fiscal issues ("fiscal cliff") and Hurricane Sandy. From 2013 to 2017, we observe a shift toward policy and geopolitical shocks: repeated debt-ceiling debates, the 2013 government shutdown, the 2014 oil-price collapse, and 2016's Brexit pair; 2017 adds U.S. tax-reform chatter ("border," "tax reform") and severe hurricanes.

The single most dramatic spike occurs in 2020:Q2, where 74 distinct bigrams—dominated by COVID-19–related terms such as "covid pandemic," "social distancing," and "unprecedented times"—signal the pandemic's systemic shock. Subsequent quarters track vaccine rollout ("vaccination" in 2021:Q1), variants ("delta," "omicron"), Russia's invasion ("geopolitical risk," "Russia" in 2022:Q1), renewed macro fears ("recession," "macro uncertainty"), 2023's U.S. regional-bank turmoil ("bank failures"), and labor unrest ("UAW" in 2023:Q4). The most recent entries capture monetary-policy turning points ("rate cuts" in 2024:Q1), election uncertainty ("new administration" in 2024:Q4), and trade tensions ("potential tariffs," "Mexico" in 2025:Q1), culminating in 2025:Q2 with 21 recession- and volatility-oriented bigrams. Overall, the table illustrates how our algorithm repeatedly flags well-known macro, policy, and disaster events yet also surfaces narrower episodes—such as Thai floods or autoworker strikes—that aggregate risk indices typically miss.

2.2 Scoring Earnings Calls by Risk Topic

Having examined the bigrams, we then map each emerging bigram to one of ten intuitive risk buckets—banking, disasters, disease, financial markets, geopolitics, macroeconomy, monetary policy, commodity prices, policy, and trade uncertainty—and reconcile disagreements. Terms that describe managerial responses rather than threats (e.g., "hedge program," "navigate environment") remain unclassified so that the taxonomy captures the risks themselves.

Next, we follow Hassan, Hollander, Kalyani, et al. (2024) and use the bigrams in Table 1 as seed words, and we use an embeddings vectors algorithm, Word2Vec, trained on earnings conference calls to add additional keywords to each topic.[1] See Appendix Table 2 for all topic dictionaries. Using our expanded dictionary of emerging risk topics, we score all risk sentences from earnings calls. For example, a risk sentence containing "political instability" would be counted as a geopolitical-related sentence with risk. We define the share of risk sentences that contain each topic τ as

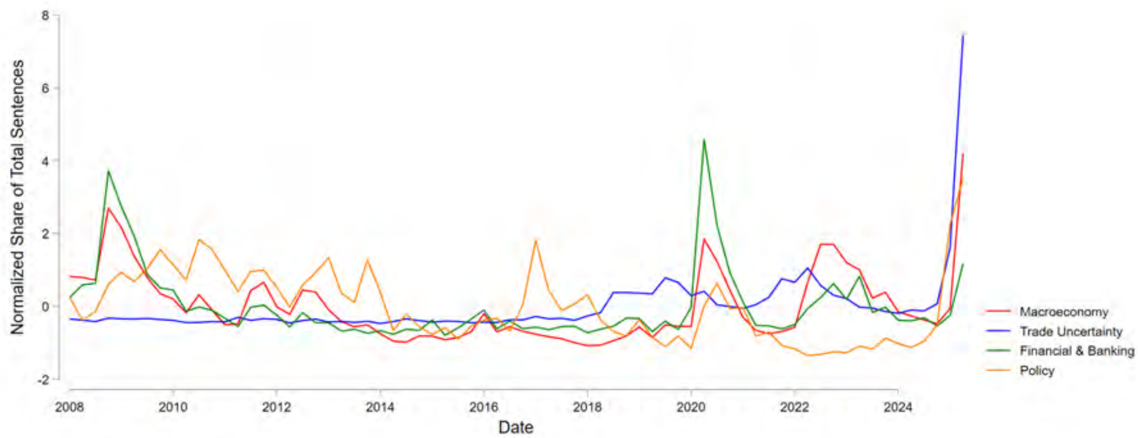
$$\text{TopicRisk}_{\tau,i,t} \ = \ \frac{\#\{\text{risk sentences containing } \tau\}}{\#\{\text{total sentences in earnings call by firm } i \text{ at time } t\}}.$$

1. We complement these keywords with intuitive additions chosen by us. For example, we add tariff and tariffs to the topic trade.

Table 1
Emerging Risk Bigrams by Quarter

Quarter	Count	Emerging Risk Bigram
2008:Q4	4	credit crisis, crisis, global financial, lehman
2011:Q3	2	debt ceiling, macro uncertainty
2011:Q4	1	thailand
2012:Q3	1	fiscal cliff
2012:Q4	1	sandy
2013:Q4	3	debt ceiling, government shutdown, shutdown
2014:Q4	1	oil prices
2016:Q1	1	macro uncertainty
2016:Q3	2	brexit, brexit vote
2017:Q1	3	border, reform, tax reform
2017:Q4	1	hurricanes
2020:Q1	1	virus
2020:Q2	74	adapting, cares, cleaning, covid, covid crisis, covid impact, covid pandemic, covid uncertainty, covidrelated, crisis, current crisis, current economic, current environment, current situation, dedication, deferral, discretionary, downturn, drew, duration, essential, extraordinary, families, financial crisis, financial flexibility, financial guidance, global pandemic, health crisis, liquidity position, longterm impact, midst, navigate, new normal, pandemic, precautionary, precautionary measure, preserve, previously issued, proactive, recover, recovery, remotely, reopen, resilience, responders, revolving, revolving credit, severity, shelterinplace, shut, social distancing, stayathome, stayathome orders, stimulus, strong liquidity, surrounding covid, suspended, suspending, temporarily, uncertain economic, uncertain future, unemployment, unprecedented, unprecedented times, unprecedented uncertainty, wellbeing, withdraw, withdrawing, withdrawn, withdrew, workers
2021:Q1	1	vaccination
2021:Q3	2	delta, variant
2022:Q1	5	geopolitical, geopolitical risk, geopolitical uncertainty, omicron, russia
2022:Q2	3	covid lockdowns, macro uncertainty, recession
2023:Q2	2	bank failures, regional banks
2023:Q4	2	geopolitical environment, uaw
2024:Q1	1	rate cuts
2024:Q4	1	new administration
2025:Q1	2	mexico, potential tariffs
2025:Q2	21	broader economic, current economic, current tariff, elevated uncertainty, global trade, macroeconomic uncertainty, pullback, recession, tariff announcements, tariff impact, tariff policies, tariff policy, tariff uncertainty, tariffrelated, trade uncertainty, uncertain economic

Figure 2
Topic Risk



NOTE: The figure plots normalized topic-risk measures for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. The four emerging risk topics in this plot—Macroeconomy, Trade, Policy, and Finance & Banking—are those with the greatest mean topic-risk score in 2025:Q2. Topic measures are the share of earnings-call sentences that contain at least one topic keyword and one risk keyword. See Appendix Figure 10 for a non-normalized plot of topic risks.

Figure 2 displays the mean share of risk sentences that contain each topic. The figure shows that more recently, there has been a surge in macroeconomic risk and trade risk. Following Loughran and McDonald (2011), we similarly define topic sentiment:

$$\text{Topic Sentiment}_{\tau,i,t} = \frac{\#\{\text{positive sentences containing } \tau\} - \#\{\text{negative sentences containing } \tau\}}{\#\{\text{total sentences in earnings call by firm } i \text{ at time } t\}}.$$

Appendix Figure 9 shows a similar downturn in macroeconomic and trade sentiment in 2025:Q2. In the past, there have been surges in macroeconomic risk (e.g., in 2008:Q4 and 2020:Q2), and trade risks have been elevated since 2018. For the rest of the article, we will focus on the effects of these two risks using the response of variables from historical data.

3. EFFECTS OF EMERGING UNCERTAINTY SHOCKS

The surge in macroeconomic- and trade-related risk in 2024–25 prompts two questions: First, how do firms adjust investment and employment when these two uncertainties spike? Second, do heightened trade risks ultimately filter through to output prices, that is, industry-level PPIs? We address the first question by comparing uncertainty measures to firm-level investment and hiring, and the second with an industry panel that matches our uncertainty measures to Bureau of Labor Statistics PPIs.

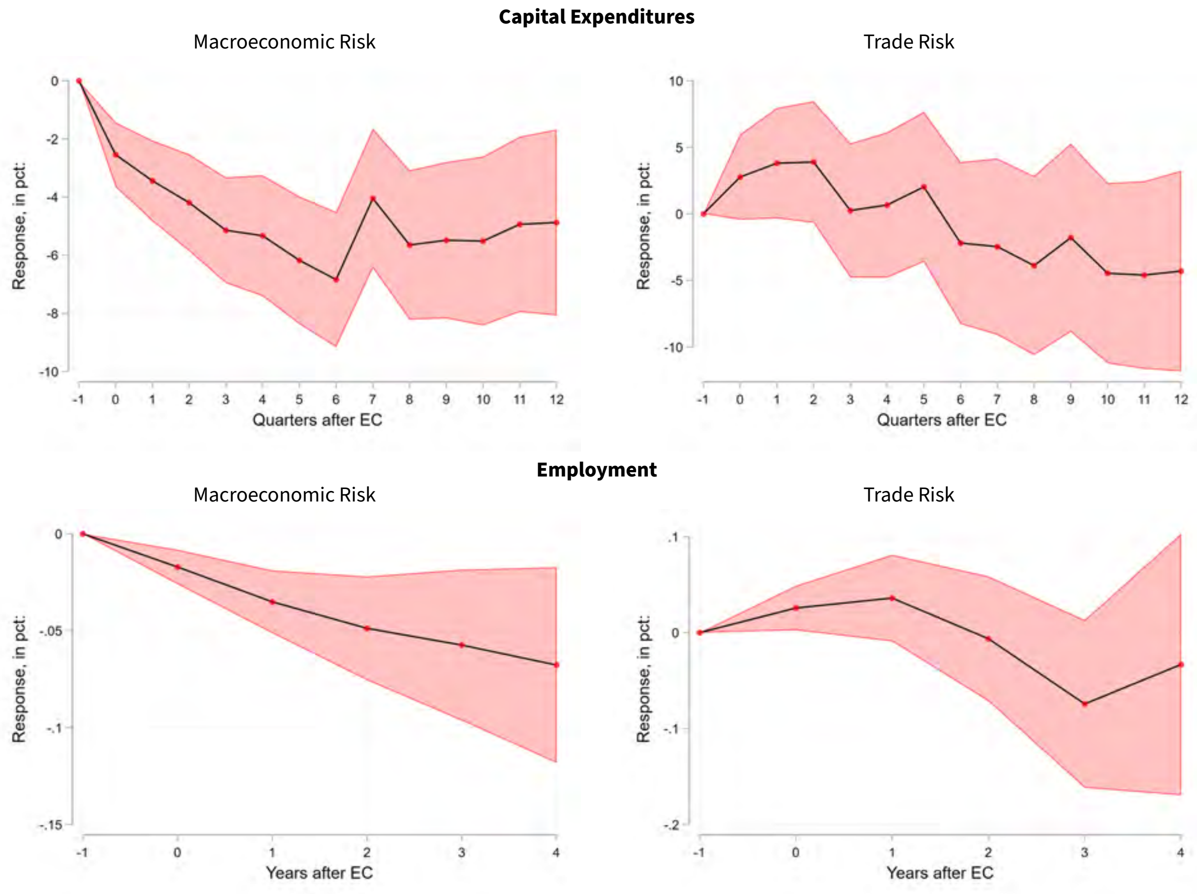
3.1 Firm-Level Specification

Our baseline dynamic equation follows Jordà and Taylor (2025) local projections so that coefficients have a direct cumulative interpretation:

$$\begin{aligned} Y_{i,t+h} &= \delta_t + \sum_{h=0}^{H}\left(\beta_h^1 \text{TopicRisk}_{\tau,i,t}\right) + \sum_{h=0}^{H}\left(\beta_h^2 \text{TopicSentiment}_{\tau,i,t}\right) + \\ &\quad \sum_{h=0}^{H}\left(\beta_h^3 \text{Non-TopicRisk}_{\tau,i,t}\right) + \sum_{h=0}^{H}\left(\beta_h^4 \text{Non-TopicSentiment}_{\tau,i,t}\right) + \mathbf{X}'_{i,t}\gamma_h + \varepsilon_{i,t+h}, \end{aligned}$$

where $Y_{i,t+h}$ is (i) the log change in CapEx, reported quarterly and measured $h = 0, \ldots, 12$ quarters ahead, and (ii) the log change in employment, reported annually and measured 4 years ahead and reported annually. δ_t denote calendar-quarter dummies that absorb aggregate shocks (including the VIX level). $\mathbf{X}_{i,t}$ collects standard controls: three lags of dependent variables, lagged assets, and industry fixed effects. Non-Topic $\text{Risk}_{\tau,i,t}$ is the

Figure 3
Macro and Trade Uncertainty, and Investment and Employment



NOTE: The plots display regression coefficients of risk topic usage on the log of capital expenditures and employment growth. The plots on the left report macroeconomic risk, while those on the right report trade risk. All firms are headquartered in the U.S. Regressions on capital expenditures are at the quarterly level and control for the quarterly lag of firm *i*'s assets, dependent variable lags between time $t-1$ and $t-4$, and quarter fixed effects. Regressions on employment growth are at the annual level and control for a one-year lag of firm *i*'s assets, one-year lag of employment growth, and year fixed effects. All regressions control for non-topic risk, topic sentiment, and non-topic sentiment and include industry fixed effects. Dependent variables are winsorized, and standard errors are clustered by firm.

overall risk for firm i in time t minus Topic Risk$_{\tau,i,t}$, and Non-Topic Sentiment is similarly defined. Standard errors are clustered by firm. Figure 3 reports coefficients with 95 percent confidence bands.

Identification and Interpretation. Because δ_t absorb time-invariant heterogeneity, β^{Macro} and β^{Trade} are identified from across-firm, across-time variation in topic usage relative to other companies in the same quarter. If macro shocks are contractionary, we expect $\beta^{\text{Macro}} < 0$ for investment and employment. Conversely, if trade risk triggers pre-emptive diversification, we anticipate $\beta^{\text{Trade}} > 0$. That is exactly what we find: A 1-percentage-point increase in trade uncertainty is associated with a 0.67 percent increase in capital expenditures and 0.04-percentage-point increase in employment growth, while a similar increase in macroeconomic uncertainty is associated with a 5.33 percent decrease in capital expenditures and a 0.035-percentage-point decrease in employment growth, all one year (four quarters) after the increase in trade uncertainty.

3.2 Industry-Level Pass-Through to Prices

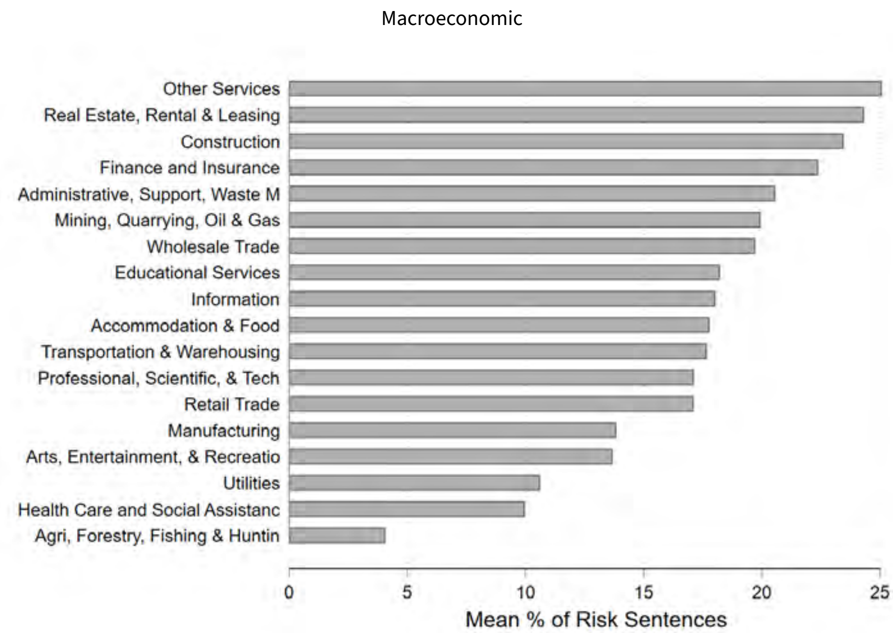
To explore pricing effects, we aggregate the firm-level trade (or macro-related) uncertainty to the three-digit NAICS industry level:

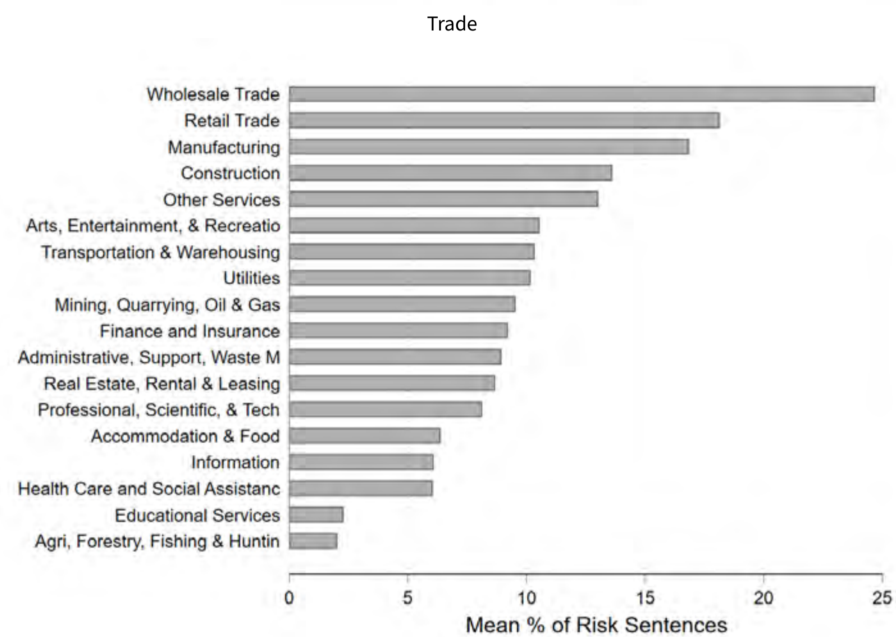
$$TopicRisk_{n,t}^{\text{Trade}} = \frac{1}{N_{j,t}} \sum_{i \in j} TopicRisk_{i,t}^{\text{Trade}}.$$

Figure 4 shows averages of these aggregates across industries. Intuitively, trade as a share of all risk sentences is most prominent in wholesale trade, retail trade, and manufacturing, while macroeconomic uncertainty is most prominent in other services, real estate, rental and leasing, and construction.

Figure 4

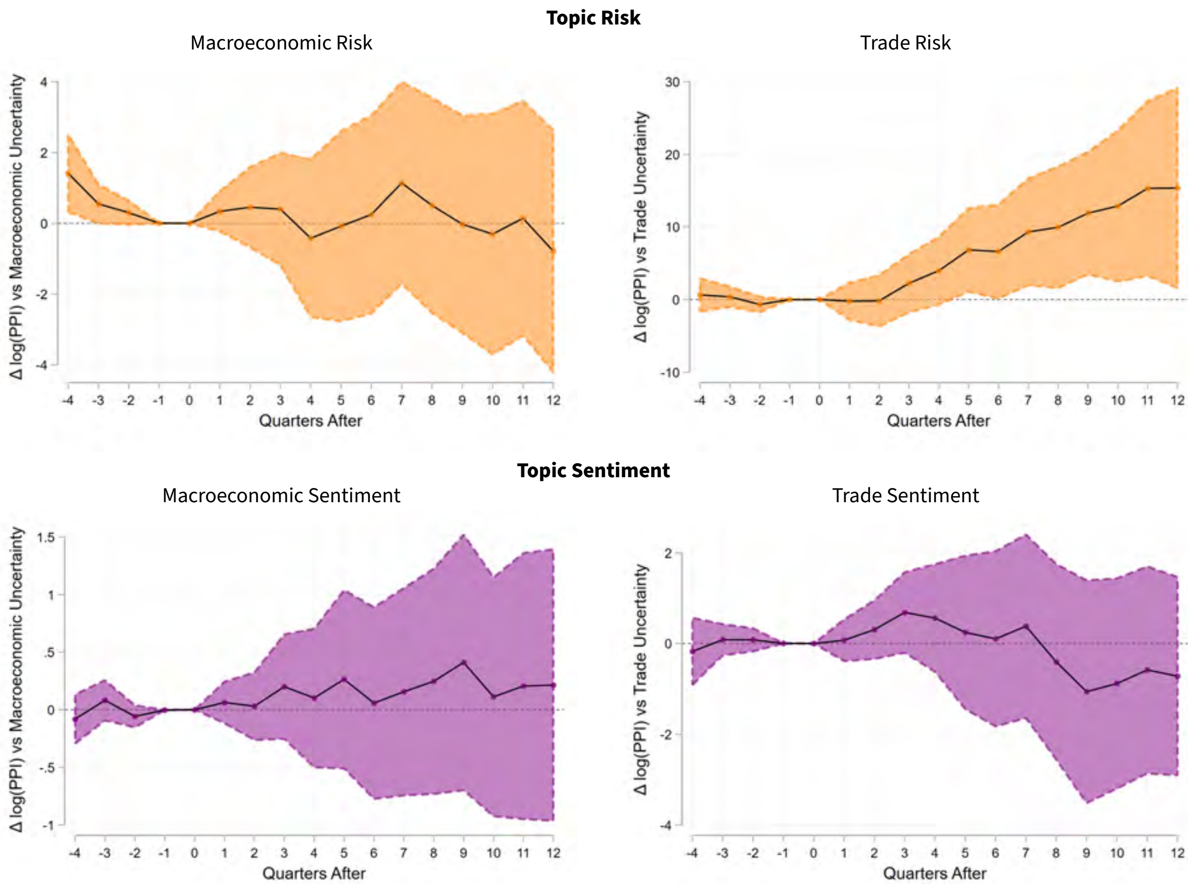
Incidence of Macroeconomic and Trade Uncertainty by Industry in 2025:Q2





NOTE: The figure displays the mean share of risk sentences that contain macroeconomic and trade risk by industry in 2025:Q2. Measures are computed as the share of earnings-call risk sentences that contain at least one topic keyword. Shares are aggregated to the industry level.

Figure 5
Macroeconomic and Trade Risk, and PPI



NOTE: The plots display dynamic regression coefficients of topic risk and sentiment on industry-level PPI growth. The plots on the left report macroeconomic usage, while those on the right report trade usage. All firms are headquartered in the U.S. Regressions include quarter and industry fixed effects, along with three lags of PPI growth, topic risk, non-topic risk, topic sentiment, and non-topic sentiment. The dependent variable is winsorized, and standard errors are clustered by industry.

Similar to firm-level specification, we run the following specification to study PPI effects:

$$Y_{n,t+h} = \delta_t + \sum_{h=0}^{H}\left(\beta_h^1 \text{TopicRisk}_{\tau,n,t}\right) + \sum_{h=0}^{H}\left(\beta_h^2 \text{TopicSentiment}_{\tau,n,t}\right) + \sum_{h=0}^{H}\left(\beta_h^3 \text{Non-TopicRisk}_{\tau,n,t}\right) + \sum_{h=0}^{H}\left(\beta_h^4 \text{Non-TopicSentiment}_{\tau,n,t}\right) + \mathbf{X}'_{n,t}\gamma_h + \varepsilon_{n,t+h}.$$

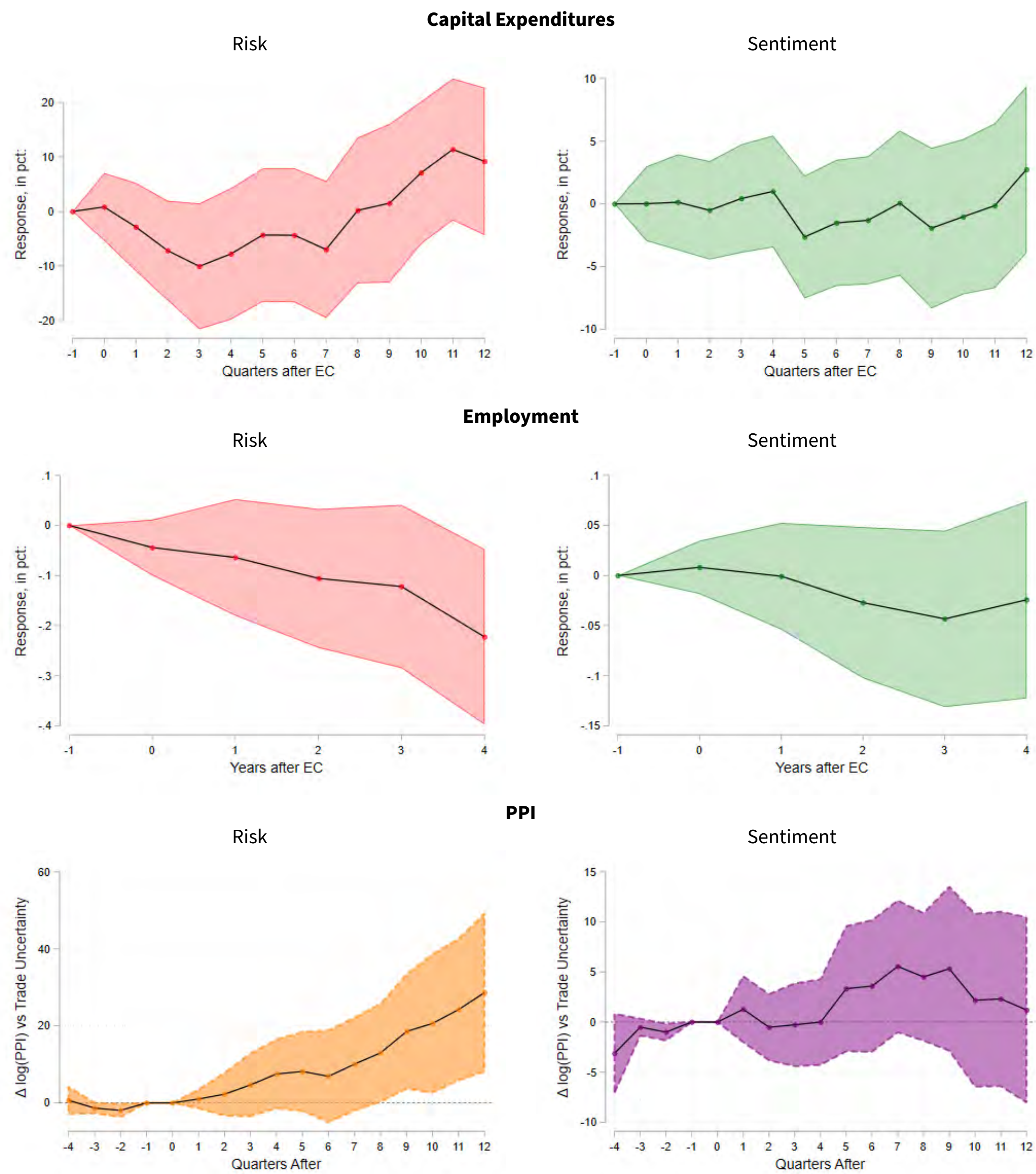
Figure 5 shows the effect of a 1-percentage-point increase in trade uncertainty in earnings calls on log PPI changes. Because the equation above includes time dummies δ_t, β^1 measures whether industries where firms talk more intensely about trade risk tend to subsequently post higher price increases than other industries facing the same aggregate shocks. Positive and statistically significant cumulative effects at horizons 5–12 quarters would corroborate anecdotal evidence that firms eventually pass tariff uncertainty through to customers. Surprisingly, in Figure 5, we find no significant effects of the drop in trade sentiment on PPI changes, suggesting that firms pass on higher costs associated with anticipated increases in tariffs even before the tariffs set in. This could be due to reallocating their production, which causes an increase in marginal costs, or using the uncertainty to increase prices.

3.3 Tariff Uncertainty

Our findings suggest that trade uncertainty is associated with an increase in industry-level PPI. Trade uncertainty captures various types of concerns, including supply chain and tariffs, and the recent rise in trade uncertainty is driven by concerns about tariffs. To investigate tariff risk further, we consider only three bigrams that clearly signal discussion about tariffs: tariff, tariffs, and import dut*. This subset of trade keywords forms a new tariff subtopic. Using the same specifications as before, we estimate the impact of tariff uncertainty on capital expenditures, employment, and inflation.

Figure 6

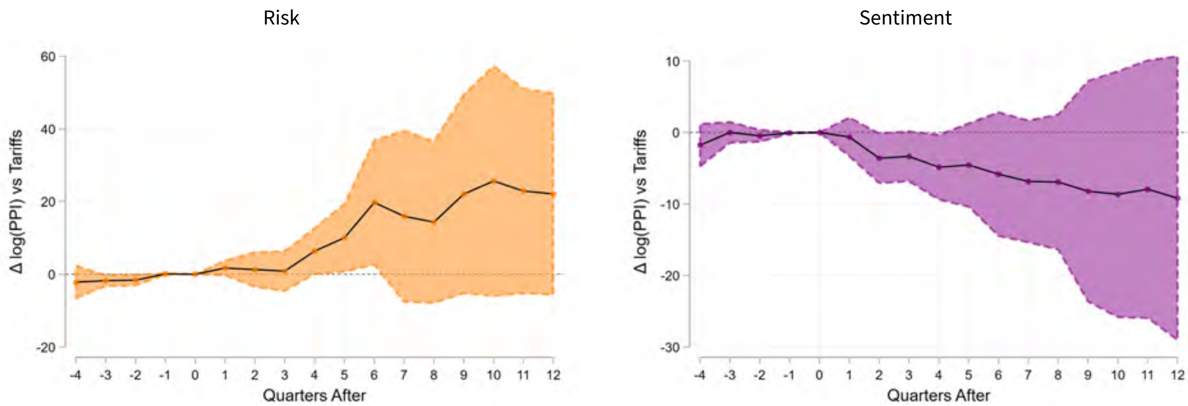
Tariff Risk and Capital Expenditures, Employment, and PPI



NOTE: The plots display regression coefficients of tariff discussions on the log of capital expenditures, employment growth, and industry-level PPI growth. The plots on the left report tariff risk, while those on the right report tariff sentiment. See the notes of Figures 3 and 5 for details of regression specifications. Dependent variables are winsorized, and standard errors are clustered by firm.

Figure 6 presents the results. Even after restricting to three tariff-related keywords, we observe a statistically significant increase in industry-level inflation following increases in tariff uncertainty. A 1-percentage-point increase in tariff uncertainty is associated with a 6.81-percentage-point increase in PPI six quarters later.

Figure 7
Tariff Risk and PPI: 2008–19



NOTE: The plots display dynamic regression coefficients of tariff risk and sentiment on industry-level PPI growth before 2020. All firms are headquartered in the U.S. Regressions include quarter and industry fixed effects, along with three lags of PPI growth, topic risk, non-topic risk, topic sentiment, and non-topic sentiment. The dependent variable is winsorized, and standard errors are clustered by industry.

Thus, with the observed 0.2-percentage-point increase in tariff uncertainty in 2025:Q2, we would expect a 1.36-percentage-point increase in inflation by the end of 2026. Similar to trade uncertainty, the effects of tariff sentiment on PPI are insignificant. These results further suggest that firms pass on costs associated with anticipated increases in tariffs to consumers before those costs are realized.

One may speculate about whether particular events during the sample period impacted the correlation between trade or tariff uncertainty and inflation. For example, could this relationship be the result of trade disruptions during the pandemic and the subsequent inflation in 2021 and 2022? To test this, we estimate our PPI model using data from 2008 through the end of 2019. In Figure 7, we continue to observe a positive relationship between tariff uncertainty and PPI, though the effect is statistically insignificant.

In sum, we link the real-time, topic-level risk indicators uncovered in Section 2.1 to forward-looking firm actions and industry price changes. By contrasting macro and trade uncertainty within the same framework, we quantify how the current spike in trade risk is correlated with inflationary outcomes, while the spike in macroeconomic risk is correlated with contractionary outcomes. After subsetting trade risk to isolate tariff-specific risk, we still find statistically significant increases in inflation. Together, the incidence of these risks raises the prospect of stagflation.

4. CONCLUSION

This article introduces an unsupervised NLP pipeline that turns the full text of 209,915 earnings-call transcripts into a real-time dashboard of emerging corporate risks. By flagging bigrams whose usage suddenly accelerates in risk-related contexts—and grouping them with semantic embeddings—we recover the risk narratives that matter to managers the instant they enter the conversation, from 2008's credit crisis through 2025's renewed tariff uncertainty. Ten intuitive topics—banking, disasters, disease, financial, geopolitics, macroeconomy, monetary policy, commodity prices, policy, and trade uncertainty—trace the contour of these shocks over time.

Linking the text-based signals to outcomes reveals sharp heterogeneity. When firms face a one-standard-deviation jump in macroeconomic uncertainty, they tend to decrease investment and employment for up to two years. In contrast, a comparable rise in trade uncertainty is associated with a rise in CapEx and employment; however, with a four- to eight-quarter lag, producer prices tend to increase, consistent with supply-chain diversification and tariff pass-through. Our evidence points to higher trade uncertainty as a stagflationary shock.

Beyond documenting this tension, our methodology offers a scalable early warning system: It discovers new shocks automatically, quantifies their firm-level exposure, and forecasts growth and inflation effects well before aggregate indicators move.

REFERENCES

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. Measuring economic policy uncertainty. *The quarterly journal of economics* 131 (4): 1593–1636.

Białkowski, Jędrzej, Huong Dieu Dang, and Xiaopeng Wei. 2022. High policy uncertainty and low implied market volatility: an academic puzzle? *Journal of Financial Economics* 143 (3): 1185–1208.

Caldara, Dario, and Matteo Iacoviello. 2022. Measuring geopolitical risk. *American economic review* 112 (4): 1194–1225.

Denny, Matthew J, and Arthur Spirling. 2018. Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it. *Political analysis* 26 (2): 168–189.

Hanley, Kathleen Weiss, and Gerard Hoberg. 2019. Dynamic interpretation of emerging risks in the financial sector. *The Review of Financial Studies* 32 (12): 4543–4603.

Hassan, Tarek A, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun. 2019. Firm-level political risk: Measurement and effects. *Quarterly Journal of Economics* 134 (4): 2135–2202. ISSN: 0033-5533. https://doi.org/10.1093/qje/qjz021. eprint: https://academic.oup.com/qje/article-pdf/134/4/2135/32666210/qjz021.pdf. https://doi.org/10.1093/qje/qjz021.

Hassan, Tarek Alexander, Stephan Hollander, Aakash Kalyani, Laurence van Lent, Markus Schwedeler, and Ahmed Tahoun. 2024. *Economic surveillance using corporate text.* Technical report. National Bureau of Economic Research.

Hassan, Tarek Alexander, Stephan Hollander, Laurence van Lent, and Ahmed Tahoun. 2023. Firm-level exposure to epidemic diseases: Covid-19, SARS, and H1N1. *Review of Financial Studies* 34 (6): 2409–2449.

———. 2024. The global impact of Brexit uncertainty. *Journal of Finance* 79:413–458.

Jordà, Òscar, and Alan M Taylor. 2025. Local projections. *Journal of Economic Literature* 63 (1): 59–110.

Kalyani, Aakash. 2023. The creativity decline: Evidence from US patents. *http://dx.doi.org/10.2139/ssrn.4318158.*

Kalyani, Aakash, Nicholas Bloom, Marcela Carvalho, Tarek Hassan, Josh Lerner, and Ahmed Tahoun. 2023. The diffusion of new technologies. *http://dx.doi.org/10.2139/ssrn.3875827.*

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

Schick, Timo, and Hinrich Schütze. 2020. Rare words: a major problem for contextualized embeddings and how to fix it by attentive mimicking. In *Proceedings of the aaai conference on artificial intelligence,* 34:8766–8774. 05.

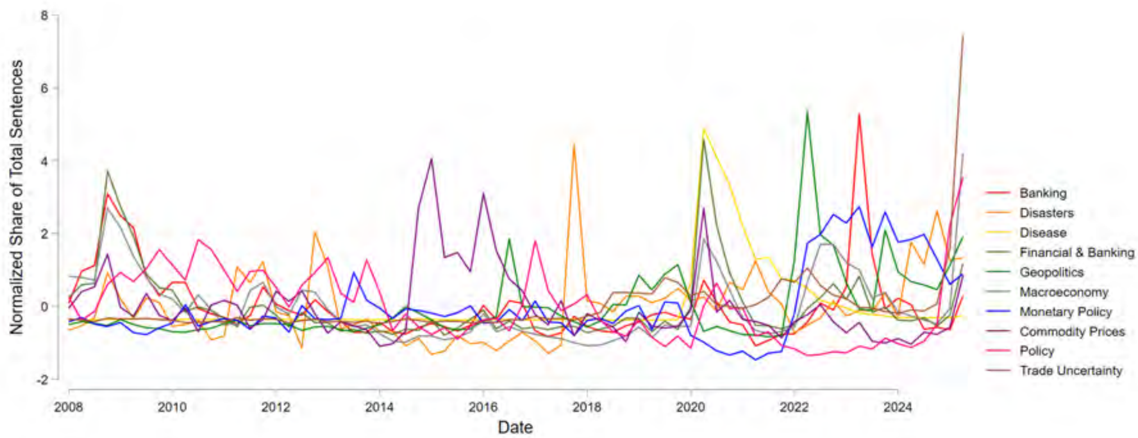
Tan, Chade-Meng, Yuan-Fang Wang, and Chan-Do Lee. 2002. The use of bigrams to enhance text categorization. *Information processing & management* 38 (4): 529–546.

Appendix: Table 2
Emerging Risk Topics and Associated Keywords

Topic	Keywords
Banking	bank, banks, banking, lenders
Disasters	sandy, irma, natural disaster, disaster recovery, extreme weather, flood*, wildfire*, earthquake*, hurricane*, storm*, tsunami*, tornado*
Disease	omicron, variant*, vaccin*, covid*, nonessential, pandemic, ppe, stayathome, responders, lockdown, shelterinplace, new normal, protective equipment, cleaning, distancing, virus, infection rates, booster shots, mask mandates, quarantine, infection control, coronavirus, sarscov, sars cov, delta variant, ebola
Financial	financial crisis, strong liquidity, stress test, stock market, global financial, financial crisis, credit crisis, credit conditions, liquidity, capital ratios, balance sheet, credit availability, capital adequacy, loan portfolios, mortgage lending, consumer lending, credit quality, asset management, investment portfolios, financial markets, trading, derivatives
Geopolitics	international tension*, political instability, global tension*, diplomatic relations, conflict*, war, wars, coups, coup, military, guerrila, warfare, terrorist*, terrorism, air strike, heavy casualties, ukraine, geopolitical, geopolitics, russia, brexit, russiaukraine
Macroeconomy	gdp, inflation*, deflation*, consumer spending, market volatility, recession, econom*, market, macroeconom*, unemployment, furlough, macro uncertainty
Monetary Policy	interest rate*, fed polic*, federal reserve, fed system, quantitative easing, monetary polic*, dovish, hawkish, rate hike*, rate cut*
Commodity Prices	oil price, oil prices, oil pricing, steel price, steel prices, steel pricing, natural gas price, natural gas prices, natural gas pricing, copper price, copper prices, copper pricing, aluminum price, aluminum prices, aluminum pricing, iron ore price, iron ore prices, iron ore pricing, nickel price, nickel prices, nickel pricing, lithium price, lithium prices, lithium pricing, cobalt price, cobalt prices, cobalt pricing, gold price, gold prices, gold pricing, silver price, silver prices, silver pricing, platinum price, platinum prices, platinum pricing, coal price, coal prices, coal pricing, brent price, brent prices, brent pricing, wti price, wti prices, wti pricing, lng price, lng prices, lng pricing, propane price, propane prices, propane pricing, wheat price, wheat prices, wheat pricing, corn price, corn prices, corn pricing, soybean price, soybean prices, soybean pricing, cotton price, cotton prices, cotton pricing, coffee price, coffee prices, coffee pricing, lumber price, lumber prices, lumber pricing, pulp price, pulp prices, pulp pricing, pvc price, pvc prices, pvc pricing, ethylene price, ethylene prices, ethylene pricing, polypropylene price, polypropylene prices, polypropylene pricing, ore price, ore prices, ore pricing, metal price, metal prices, metal pricing, energy price, energy prices, energy pricing
Policy	executive orders, government, administration, policy, regulat*, tax, cares, stimul*, reform*, debt ceiling, super committee, debt ceiling, moratorium
Trade Uncertainty	sanctions, tariff, tariffs, export control*, export restriction*, import restriction*, import dut*, import barrier*, supply chain, treat*, trade, antidumping, dumping, gatt, world trade organization, wto
Tariff Uncertainty	tariff, tariffs, import dut*
Work from Home	remote work, hybrid work, remote employees, distributed workforce, flexible work arrangements, virtual meetings, telecommuting, home office, wfh

NOTE: The table displays final keyword lists for each emerging risk topic. The asterisk "*" denotes a wildcard (e.g., regulat* covers keywords including regulation, regulations, and regulatory). Topics were created by the authors based on unigrams and bigrams identified as "emerging" in earnings calls during the time period. An emerging n-gram appears in at least 25 distinct transcripts in a given quarter, and its share of earnings calls increases by at least 200 percent relative to the previous quarter. We bin these emerging risk bigrams into the topics listed above. We then use an embeddings vectors algorithm, Word2Vec, trained on earnings conference calls, to add additional keywords to each topic. Finally, we complement these keywords with intuitive additions chosen by us (e.g., tariff).

Appendix: Figure 8
Topic Risk: All Topics

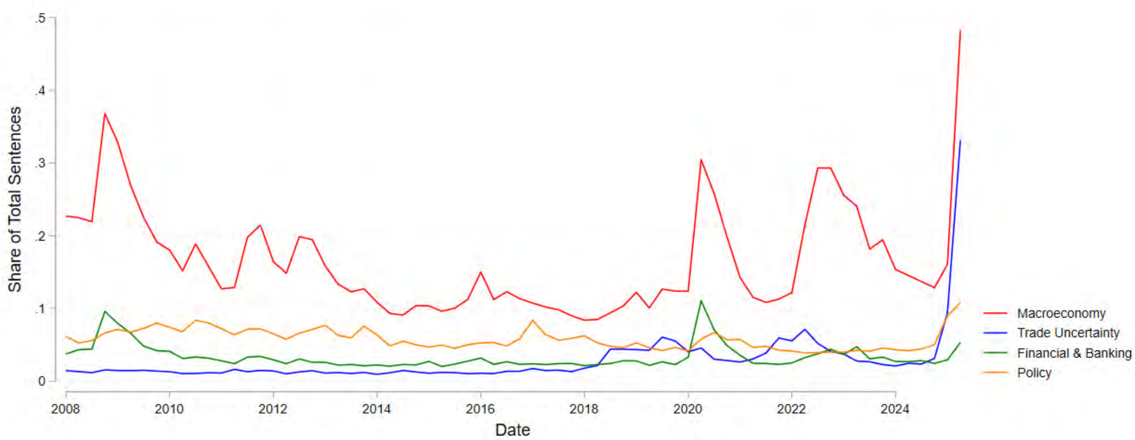


NOTE: The figure plots normalized topic-risk measures for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. Topic measures are the share of earnings-call sentences that contain at least one topic keyword and one risk keyword.

Appendix: Figure 9
Topic Sentiment

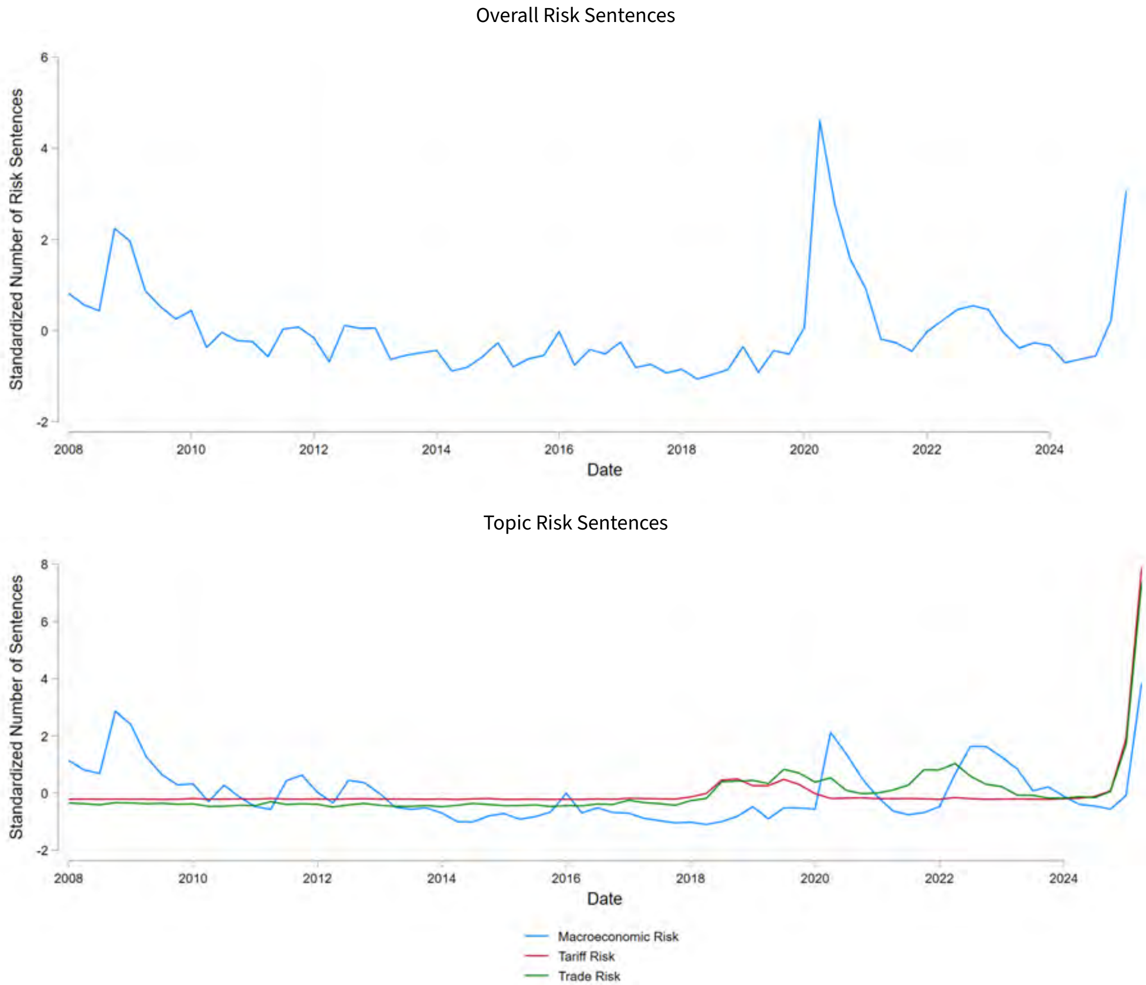
NOTE: The figure plots normalized topic-sentiment measures for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. The four emerging risk topics in this plot—Macroeconomy, Trade, Policy, and Finance & Banking—are those with the greatest mean topic-risk score in 2025:Q2. The topics were created by the authors based on unigrams and bigrams identified as "emerging" in earnings calls during the time period. An emerging n-gram appears in at least 25 distinct transcripts in a given quarter, and its share of earnings calls increases by at least 200 percent relative to the previous quarter. Net sentiment of topic $\tau = 100 \times \frac{\#\{\text{positive sentences containing } \tau\} - \#\{\text{negative sentences containing } \tau\}}{\#\text{ sentences}}$.

Appendix: Figure 10
Non-Normalized Topic Risk



NOTE: The figure plots non-normalized topic-risk measures for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. The four emerging risk topics in this plot—Macroeconomy, Trade, Policy, and Finance & Banking—are those with the greatest mean topic-risk score in 2025:Q2. The topics were created by the authors based on unigrams and bigrams identified as "emerging" in earnings calls during the time period. An emerging n-gram appears in at least 25 distinct transcripts in a given quarter, and its share of earnings calls increases by at least 200 percent relative to the previous quarter. Topic measures are the share of earnings-call sentences that contain at least one topic keyword and one risk keyword.

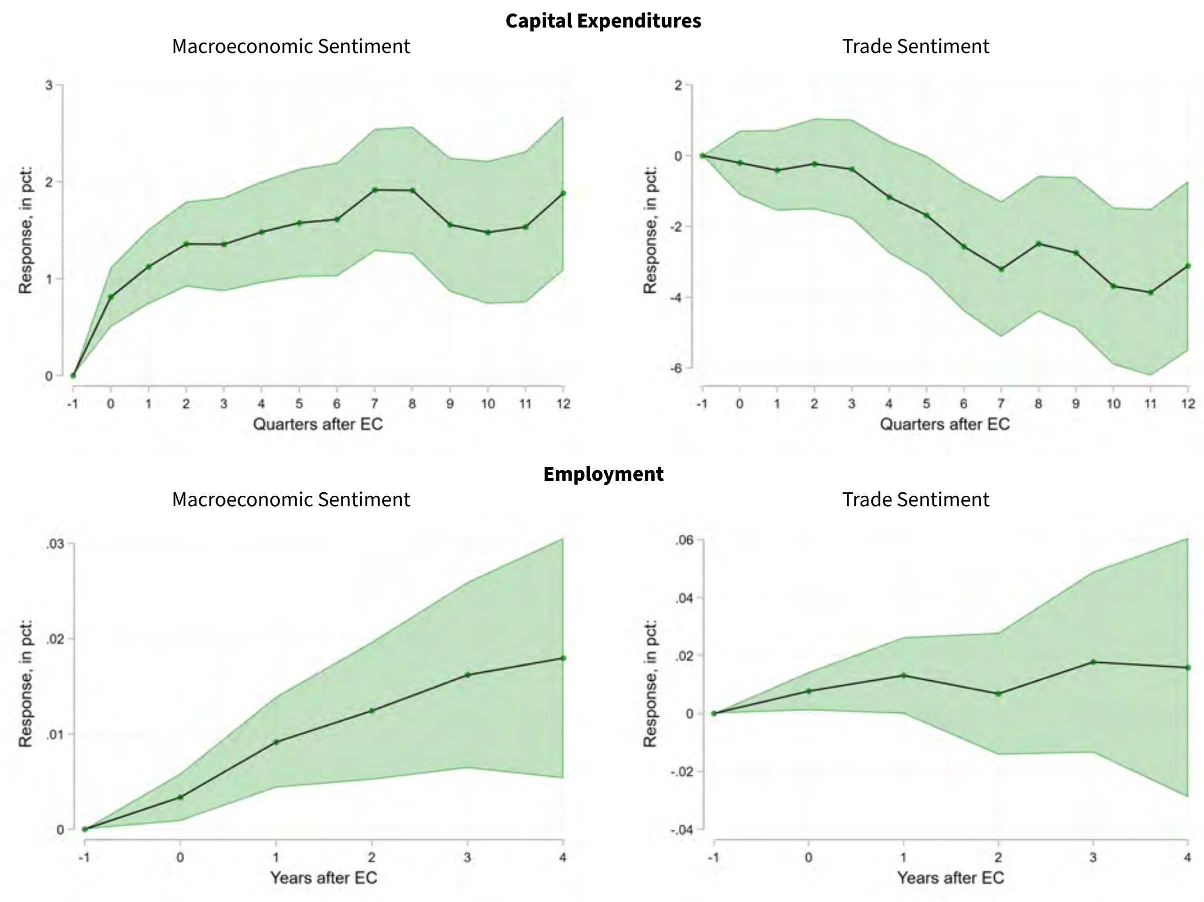
Appendix: Figure 11
Normalized Number of Sentences



NOTE: The figure plots the normalized number of risk sentences and topic-risk sentences for U.S.-headquartered firms from 2008:Q1 through 2025:Q2. Topics were created by the authors based on unigrams and bigrams identified as "emerging" in earnings calls during the time period. Topic-risk sentences are earnings-call sentences that contain at least one topic keyword and one risk keyword.

Appendix: Figure 12

Macro and Trade Net Sentiment, and Investment and Employment



NOTE: The plots display dynamic regression coefficients of topic sentiment on the growth rate of capital expenditures and employment growth. The plots on the left report macroeconomic sentiment, while those on the right report trade sentiment. All firms are headquartered in the U.S. Regressions on capital expenditures are at the quarterly level and control for the quarterly lag of firm *i*'s assets, dependent variable lags between time $t - 1$ and $t - 4$, and quarter fixed effects. Regressions on employment growth are at the annual level and control for a one-year lag of firm *i*'s assets, one-year lag of employment growth, and year fixed effects. All regressions control for topic risk, non-topic risk, and non-topic sentiment and include industry fixed effects. Dependent variables are winsorized, and standard errors are clustered by firm.