Risk from the Inside Out: Understanding Firm Risk through Employee News Consumption*

Fahiz Baba-Yara† Fotis Grigoris‡ Preetesh Kantak§

March 2024

Abstract

We use employee news consumption to characterize firms' exposures to macroeconomic risk, making use of data covering two billion employee-article interactions per day across millions of firms. We find that, in the time-series, employees consume more macroeconomic news following the onset of bad times. In the cross-section, firms whose employees were reading more macroeconomic news ex-ante are more exposed to changing economic conditions ex-post. Consistent with the notion that employee news consumption provides insights into firms' risk exposures, we show that the more exposed firms hedge more, yet have higher costs of capital and subsequently lower investment and hiring rates.

*Previous versions of this paper were circulated under the title "Are Uncertain Firms Riskier?" We are grateful to Hengjie Ai, Xiaohui Bakshi (discussant), Azi Ben-Rephael, Nick Bloom, Greg Brown, Bruce Carlin, Daniel Carvalho, Riccardo Colacito, Tony Cookson, Max Croce, Zhi Da, Jesse Davis, Steven Davis, Andrew Ellul, Joseph Engelberg, Diego Garcia (discussant), Andrei Goncalves, Isaac Hacamo, Christian Heyerdahl-Larsen, Burton Hollifield, Harrison Hong, Ryan Israelsen, Calvin Jia (discussant), Nishad Kapadia, Charles Martineau (discussant), Stijn Van Nieuwerburgh, Stavros Panageas, Anna Pavlova, Jeffrey Pontiff, Seth Pruitt, Jacob Sagi, Gill Segal, John Shim, Kelly Shue, Noah Stoffman, Bo Sun, Andrea Tamoni, Stephen Terry (discussant), Charles Trzcinka, Zhenyu Wang, Isabella Wolfskeil (discussant), Nancy Xu, Mao Ye, Guofu Zhou (discussant), and Kyle Zimmerschied (discussant) for their valuable input. We also thank conference and seminar participants at the ASSA Conference, ASU-Sonoran Winter Finance Conference, BI Norwegian, Bocconi University, Bristol Financial Markets Conference, Dolomite Summer Finance Conference, the Federal Reserve Board Conference on Nontraditional Data and Macroeconomics, Indiana University—Bloomington, the Midwest Finance Association Conference, Millennium Global Investment, the NBER Summer Institute, PanAgora Asset Management, SITE Macroeconomic Uncertainty and Volatility Conference, the Conference on Financial Economics and Accounting, the Financial Research Association Conference, and the University of North Carolina—Chapel Hill. The authors further acknowledge the Indiana University Pervasive Technology Institute for providing HPC database, storage, and consulting resources that have contributed to the research results reported within this paper (Stewart, Welch, Plale, Fox, Pierce, and Sterling, 2017). We are indebted to our data provider, who wishes to remain anonymous, for its generous support and Carter Davis for his help in the early stages of the project. All errors are our own.

†Indiana University (fababa@iu.edu)

‡University of Iowa (fgrigoris@uiowa.edu)

§Indiana University, Corresponding author (pkantak@iu.edu)

Investors, policymakers, and firms face the difficult task of allocating scarce resources in a continuously evolving economic landscape. Recognizing this challenge, a growing literature looks to gain insights by analyzing the timing of news (e.g., Fisher, Martineau, and Sheng, 2022), attention to salient economic announcements (e.g., Ben-Rephael, Carlin, Da, and Israelsen, 2021), social media discourse (e.g., Cookson, Engelberg, and Mullins, 2023), and perceptions of aggregate risks (e.g., Ai and Bansal, 2018). Implicit in this literature is a focus on *investor* considerations. In contrast, the literature has largely ignored attention to economic concerns from the production side of the economy. That is, can we measure *firms*' economic concerns, namely their exposures to risks, in such an environment? We answer this question by probing the minds of each firm's employees and assessing the information content in the news they consume on a daily basis.

At first glance it is not evident that rank-and-file employees' news consumption would provide insights into a firm's growth prospects and risk exposures. On the one hand, a firm worried about its risk exposure could task its employees to learn about and act upon those risks. Similarly, employees may read news about the state of the economy if they are concerned about their own human capital that is tied to firm prospects. In either case, the news that the average employee reads may provide insights into the risks the firm faces. On the other hand, if the content rank-and-file employees read does not vary with their employer's business prospects, then the news they consume will be uninformative.

Our analysis demonstrates that examining patterns in employees' news consumption provide insights into the risk exposures of firms. Utilizing a comprehensive dataset—encompassing over two million firms, 4,000 publishers, and one billion daily user-article interactions—we develop a firm-level "relative attention" measure. The measure captures the extent to which employees prioritize macroeconomic news over other business-related news.[1] Through this, we deduce a firm's exposure to systematic risk and establish two baseline facts.

First, we perform a time-series analysis showing that employees at a typical firm shift their attention towards consuming more macroeconomic news in the days following bad aggregate economic conditions. As published news articles are an equilibrium outcome of readers' preferences and the media's production technology (Mullainathan and Shleifer, 2005) that can affect asset prices (Martineau and Mondria, 2023), it is reassuring but not altogether surprising that aggregate

[1] To access the employee attention measures used in this paper please follow this link.

consumption of macroeconomic news is correlated with measures of economic conditions. Thus, second, we conduct a cross-sectional analysis that establishes that firms whose employees were reading more macroeconomic news prior to an economic shock consume even more macroeconomic news after the shock. These highly exposed firms have higher costs of capital, engage in more risk management, and subsequently invest less in both physical and human capital. In short, our analysis shows that firm-level employee news consumption reflects both the *ex-ante* exposure to macroeconomic risks as well as the *ex-post* reaction to subsequent macroeconomic shocks.

Our motivation for measuring firm-level risk exposures through employee news consumption is rooted in the literature that elicits firm-level expectations from surveys of key corporate personnel (see, e.g., Weber, D'Acunto, Gorodnichenko, and Coibion, 2022, for a review). While the responses to these surveys often reflect the expectations of top executives, they are typically (i) conducted on a low-frequency basis (e.g., quarterly or annually), (ii) reflect only the narrow set of questions asked by the surveyor, and (iii) conducted anonymously such that these expectations cannot be linked to firm actions. Our ability to observe the news consumption of rank-and-file employees at firms can be thought to reflect an almost high-frequency (i.e., daily) survey in which employees express their most salient concerns. We assume that an employee concerned about a particular risk would choose to read news about it, and that, if enough employees are both aware and concerned, then employee-aggregated reading will reflect the high firm-level exposure to that risk.

Our data on employee reading of news articles is drawn from a consortium of over 4,000 online publishers (hereafter referred to as "the Consortium"). Each of these publishers—which span a variety of news publications (e.g., the *Wall Street Journal*, *Forbes*, and *Bloomberg*) and trade periodicals (e.g, *Hart Energy*)—provide the Consortium with data on each user-article interaction (e.g., URL of the article, time of reading, external IP address). This allows the Consortium to link the nearly two *billion* daily interactions from 2016 through 2022 to specific firms. For example, the Consortium would observe the 10 unique users at Company A reading the same article on a given day, while observing only one user at Company B reading two different articles.

The Consortium deploys a machine learning algorithm on the content to decompose each article into its essential topics.[2] Returning to the previous example, the Consortium may determine that

[2]The Consortium's primary business is to generate a signal of user *intent* to purchase an underlying product or service. The objective of this machine learning algorithm is to generate an accurate, topic-specific signal so that a client of the Consortium can better direct sales, marketing and advertising dollars towards a firm whose topic-related

the article read by employees of Company A was 30% related to "inflation" and 70% related to "FOMC." In contrast, the first article read by employees of Company B was 100% related to "CPUs," and the second article was evenly split between "CPUs" and "cloud computing." After decomposing each article into topics, the Consortium aggregates topic interactions within each firm and day to produce a dataset of firm-topic interactions. In our example, the Consortium would uncover that Company A (Company B) was mainly focused on the "FOMC" ("CPU") topic for that day. This dataset forms the basis of our analysis.

While the set of news sources, articles, and topics covered in the Consortium's dataset is vast, we are primarily interested in the degree to which firms are paying attention to macroeconomic news. To define a constrained set of macroeconomic-related topics, we construct a corpus of macroeconomically relevant news in the spirit of the Economic Policy Uncertainty (EPU) index of Baker, Bloom, and Davis (2016). Namely, we gather articles that mention keywords, such as "economy," from the *Wall Street Journal*, *USA Today*, and *New York Times*, among other publishers. We have the Consortium deploy its algorithm onto these articles to identify the subset of approximately 600 topics in their corpus that are related to the macroeconomy. Aggregating each firm's relative attention towards this subset of topics on a given day allows us to obtain a firm-by-day measure of how intensively employees are consuming macroeconomic-related news.

Do the typical firm's employees shift their attention towards macroeconomic-related news as business conditions deteriorate? We answer this question using a proxy that captures the proportion of macroeconomic-relevant topics a firm's employees are paying attention to on a given day relative to other topics. We find that in the time-series, relative attention to macroeconomic news is closely related to the underlying state of the economy. For instance, employees at a typical firm read more macroeconomic news in the days following an increase in the corporate default spread, but less macroeconomic news as the funding conditions of financial intermediaries improve (He, Kelly, and Manela, 2017). Firms' employees also read relative more macroeconomic-related news following a rise in economic uncertainty, measured using either the VIX, EPU, or Bekaert, Engstrom, and Xu (2022) uncertainty indices.

While these time-series results demonstrate that employees re-allocate their attention to macroe-

intent is high. The Consortium is thus economically motivated to fit the large corpus and diverse set of topics as well as possible. At present, the articles are decomposed into nearly 7,000 unique topics.

conomic news in response to fluctuating economic conditions, they do not tell us which firms are most likely to shift their attention in bad times and why. As noted above, the set of news articles published is an equilibrium outcome of the media's production decisions and readers' preferences. As such, the previous measure of attention confounds supply and demand for news. To distinguish between the two, we use a procedure similar to *tf-idf* scores in computational linguistics (see, e.g., Gentzkow, Kelly, and Taddy, 2019, for an application of *tf-idf* scores in finance). In short, terms with high (low) *tf-idf* scores are those that are most (least) useful for differentiating the content of a given document from a corpus of documents. Similarly, we compute time-varying *topic frequency, inverse aggregate frequency* (henceforth, *tf-iaf*) scores for each macroeconomic topic. Topics with a high (low) *tf-iaf* score in a given period are those that are the most (least) useful for differentiating what firms are paying attention to in the cross-section of firms.

As an example of the intuition underlying these scores, consider the monthly release of the Consumer Price Index (CPI). Many employees will read articles about inflation around this time by virtue of the increased supply of such articles. Knowing that employees at two firms—A and B—are reading about inflation will tell us little about which firm is more exposed to the macroeconomy. Consequently, inflation-related topics will have low *tf-iaf* scores during these times. However, if the employees of Firm B more intensely reads other macroeconomic topics (e.g., interest rate swaps or duration management), then these topics will carry relatively high *tf-iaf* scores. We classify Firm B and its employees as being plausibly more exposed to macroeconomic risk than Firm A. This logic also underlies the use of *tf-idf* scores in a wide variety of economic application, such as constructing the climate risk indexes in Engle, Giglio, Kelly, Lee, and Stroebel (2020).

Our primary measure of firm-level exposure to macroeconomic risk captures the proportion of employees' attention towards macroeconomic topics, where topics are now weighted by their *tf-iaf* scores. Using the weighted measure of news consumption, we perform a simple validation exercise. If our cross-sectional measure of a firm's relative attention to macroeconomic news (henceforth, CS-RA$_{i,t}$) reflects firm exposure to macroeconomic risk, then we would expect the macroeconomic news consumption of the employees of high CS-RA$_{i,t}$ firms to be more sensitive to fluctuating economic conditions than that of low CS-RA$_{i,t}$ firms. The results from this exercise are economically and statistically stark. When the level of economic activity declines, or the uncertainty associated with economic conditions rises, the employees of high CS-RA$_{i,t}$ firms consume even more macroeconomic news than those of low CS-RA$_{i,t}$ firms. We further corroborate these results by conducting a high-frequency event study examining how employees of high versus low CS-RA$_{i,t}$ firms shifted their macroeconomic news consumption around the onset of the COVID-19 pandemic, which is the most prominent economic shock in our sample.

The set of topics that are upweighted (downweighted) to maximize cross-sectional differences in macroeconomic reading seems to reflect the systematic risk exposures of firms. Two natural questions follow. First, are the firms with employees who are reading more macroeconomic news actively managing their risk exposures? Second, how do investors perceive the risks of these more exposed firms? In regard to the first question, we find that firms reading the most about macroeconomic risk are about 5 to 10% more likely to partake in hedging activity than low CS-RA$_{i,t}$ firms. This propensity increases by an additional 5% in bad times. We measure hedging activity using the approach of Campello, Lin, Ma, and Zou (2011) and define bad times as periods in which the EPU index increases by more than one-standard-deviation. These findings are robust to a variety of controls and industry-by-date fixed effects, highlighting how high CS-RA$_{i,t}$ firms are more likely to engage in hedging activity for reasons that extend beyond industry-specific risk exposures and key macroeconomic events compared to their low CS-RA$_{i,t}$ peers.

In regard to the second question, we find that the equity market perceives these high CS-RA$_{i,t}$ firms as particularly risky. Notably, the implied cost of capital measure of Gebhardt, Lee, and Swaminathan (2001) shows that high CS-RA$_{i,t}$ firms have costs of capital that are about 0.60% to 1.70% per annum larger than those of low CS-RA$_{i,t}$ firms. This difference in the cost of capital across firms is not driven by common shocks at a given point in time (e.g., the onset of the COVID-19 pandemic) or differences in the cost of capital across industries (i.e., the fact that durable goods manufacturers are riskier than non-durable goods manufacturers). Moreover, the association between the intensity of macroeconomic-related reading and the cost of equity capital survives when we control for both industry-by-time fixed effects, CAPM betas, and a battery of firm-level characteristics typically associated with risk, such as size, leverage, and profitability.[3]

Having shown that reading important macroeconomic news is related to the cost of capital, we next show that it also predicts future firm-level corporate decisions. For instance, the firms with

[3] We find qualitatively similar results when we examine differences in the cost of capital across firms using simple portfolio sorts and realized returns rather than a regression approach based on analyst-implied costs of capital.

the highest relative attention measures invest about 7% (3%) per quarter less in physical capital (inventories) and have hiring growth rates that are 5% per annum lower than firms with the lowest attention to the macroeconomy. We establish these results by estimating firm-level panel regressions that control for (i) a comprehensive set of observable firm characteristics that are known to predict real firm outcomes, (ii) a variety of fixed effects that account for unobservable differences between industries and across time, and (iii) assessments of firm-level risk elicited from both earnings call transcripts and the options market. This shows that employee news consumption provide novel and incremental insights into their firm's risk exposure.

We conclude our analysis by asking two questions about the kinds of firms that allocate their attention towards macroeconomic news. First, since the asset-pricing literature has demonstrated that firms with certain characteristics (e.g., those with high profitability rates and low investment rates) are particularly risky, does a firm's relative attention simply reflect its underlying characteristics? We show that the answer to this question is unequivocally "no." A variance-decomposition analysis shows that variation in observable characteristics explains very little variation in relative attention. Second, we ask whether changing economic conditions have a causal impact on a firm's relative attention to macroeconomic news. We use the instrumental variables approach of Alfaro, Bloom, and Lin (2024) to show that exogenous variation in aggregate uncertainty has a pronounced effect on the attention allocation of a firm's employees.

1 Related Literature

This paper contributes to several strands of literature in economics and finance. Most prominently, we contribute to the nascent literature on news *consumption* by using granular data on employee news reading to infer firm-level risk exposures. Our central premise is that if enough employees are concerned and aware of a given risk faced by their firm, then the average employee's reading will reflect their employers exposure to that risk. Closely related to our study is Cookson, García, and Jarnecic (2024), who examine how the aggregate consumption of articles published by the Australian Financial Review responds to variations in market-wide and firm-specific returns. In contrast to their study, we exploit the fact that we observe the internet domains of each reader, allowing us to study *firm-level* news consumption and show that reading has both reactive and predictive value in understanding a firm's risk exposure.

Our results are closely related to the strand of literature that estimates economic quantities via news *production*. While our study takes the perspective of an article's reader and shows that the types of articles read reflect risk exposures, these other studies examine how news production helps to predict returns and explain fluctuations in aggregate economy activity. Examples include Dougal, Engelberg, Garcia, and Parsons (2012), who estimate the causal effect of news articles on returns; Manela and Moreira (2017), who develop a news-based measure of volatility; Baker et al. (2016), who develop an uncertainty index; Bybee, Kelly, Manela, and Xiu (2023a), who show how news content can predict macroeconomic quantities and financial market returns; and Fisher et al. (2022), who examine publishers' attention to key macroeconomic announcements. Our work also relates to Bybee, Kelly, and Su (2023b), who construct a factor model based on narratives in the media to explain cross-sectional returns. While their study regresses firms' stock returns on news indexes to infer risk exposures, we use our novel data on employee news consumption to obtain a measure of risk exposure that do not rely on often noisy financial market data.

Our motivation for measuring firm-level risk exposures via non-traditional data stems from the growing literature that elicits firm risk from key accounting disclosures. For instance, early work by Israelsen (2014) examines firm-level risks by estimating topic models on SEC filings to measure the risks firms disclose, while more recent work by Mazumder, Pruitt, and Ross (2023) considers text embedding in these disclosures to infer *how* a firm discloses particular risks. In a similar vein, Hassan, Hollander, Van Lent, and Tahoun (2019) uses natural language processing methods to measure political and non-political risk using discussions in earnings call transcripts. While these methods provide insights on how key decision makers (i) view their firm's risk exposures and (ii) choose to disseminate these views to investors, they suffer from the fact that the underlying disclosures are made infrequently, i.e., quarterly or annually. In contrast, our news consumption measure is available on a daily basis (see Goldstein, Spatt, and Ye, 2021, for other uses of big data in finance). While a small fraction of these employees are key decision makers, our results show that general employee news consumption is still highly informative about their employer's risks, even after we control for alternative proxies for risk exposure.

Our focus on measuring each firm's exposure to macroeconomic risk through employee news consumption is motivated by the extensive literature demonstrating that shocks to the level and volatility of macroeconomic variables, such as consumption growth, are key drivers of marginal utility. For instance, Bansal and Yaron (2004); Campbell and Cochrane (1999) and Gabaix (2012) focus on shocks to the level of macroeconomic growth, whereas Bloom (2009) and Colacito, Croce, Liu, and Shaliastovich (2022), among others, focus on time-varying second moments. First- and second-moment shocks also feature prominently in empirical studies (see, e.g., Bali, Brown, and Tang, 2017; Ludvigson, Ma, and Ng, 2021). Finding suitable proxies for firm exposure to macroeconomic risk, however, remains challenging. Return-based measures, such as the covariance between each firm's excess returns and consumption growth or the VIX index, are noisy estimates of investors', as opposed to firms', perceptions of macroeconomic risk. In contrast, our ability to measure the types of news that employees are consuming reflects a real-time measure of the salient concerns that each firm faces. Our premise—which we empirically verify—is that employees of riskier firms allocate more of their attention towards reading macroeconomic news compared to employees of less risky firms. Employee attention to macroeconomic news thus serves as a non-return-based measure of the firm's risk exposure.[4]

Finally, our paper is broadly related to the literature on attention allocations. This literature emphasizes the importance of attention in decision-making as it shapes the information that individuals use to make choices (see, e.g., Gabaix, 2014). However, while much of the literature has focused on how investors, both individual and professional, allocate their attention when making decisions (see, e.g., Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016; Peng and Xiong, 2006; Van Nieuwerburgh and Veldkamp, 2010), little is known about the attention allocation of the employees who are responsible for most of the firm's operations. Notably, it is not a given that the employees of firms that are exposed to more macroeconomic risk will spend more time reading about these topics. Indeed, Chinco, Hartzmark, and Sussman (2022) argue that investors do not appear to consider consumption risk in investment decisions, despite the large academic literature on consumption-based asset pricing. When viewed against this backdrop, our evidence is consistent with the notion of rational attention; specifically, rank-and-file employees of riskier firms do indeed vary their attention more in response to fluctuating business-cycle risks.

[4]Other studies that use non-traditional data to examine firm behavior include Loughran and McDonald (2011), who measure the tone of financial disclosures through textual analyses, and Ben-Rephael et al. (2021), who use Bloomberg terminal usage to characterize the provision of effort of corporate executives.

2 Data

Our proprietary data come from a company—"the Consortium"— that analyzes content in internet articles published across thousands of media sites (members). Although the Consortium's primary business objective is to supply clients with actionable signals of *intent* to purchase specific business-to-business products and services, the scope and variety of topics covered by their text corpus is broad. The topic breadth stems from the Consortium's diverse member pool spanning numerous industries and businesses. Members range from generalist publishers, including The *Wall Street Journal*, *Forbes*, and *Bloomberg*, to more specialized and niche content providers such as *Hart* (energy), *StepStone* (private equity), and *QuinStreet* (consumer products). Figure 1 presents a small fraction of the approximately 4,000 publishers that supply the Consortium with its data.

The Consortium's members are publishers who supply raw readership data; in return, they receive analytics providing insights into user (i.e., reader) engagement with their published content. Members are also part owners of the Consortium and so share profits from the Consortium's core business. On a typical day, the Consortium observes over two billion user interactions, garnering rich perspectives on employees' daily reading habits across public, private, and nonprofit firms. For each interaction, the Consortium logs several data points, including the URL of the specific online content being read, the user's external IP address, and their cookie data. Leveraging the URLs of each online article, the Consortium employs a BERT-based algorithm to decompose each article into topics.[5] The Consortium uses the supplied IP addresses and cookie data to associate users with domains, thus linking topic interactions to specific firms. This process provides us with granular data on the degree to which each firm (i.e., domain) is paying attention to specific topics on a daily basis.

2.1 Topic Decomposition

The Consortium's algorithms distill individual news articles into their key topics—combinations of words and associations that are learned using a set of training corpora and validated by humans. These topics come in two varieties (see, e.g., Gentzkow et al. (2019) for an overview of these textual analysis methods). "Specific" topics are created in order to provide insights to client firms that

[5]BERT, which stands for **B**idirectional **E**ncoder **R**epresentation from **T**ransformers, is a large language model developed by Google Research in 2018 (see Devlin, Chang, Lee, and Toutanova, 2018, for more details).

sell particular products or services. For example, a biotechnology firm may request information on which companies are researching "RNA sequencing" and "cancer genomics," to guide its sales and fundraising teams. On the other hand, "general" topics, which are learned from a broader corpus of articles, are created to enhance the fit of the Consortium's natural language processing (NLP) algorithm, i.e., subsume common variation in reading across users (e.g., articles about politics, vacations, and sports). One would assume that interactions with these "general" topics offer limited insights into firm-specific business operations.

Each article is potentially a combination of many topics. For instance, a piece covering "RNA sequencing" and "cancer genomics" may also reference the "U.S. Food and Drug Administration (FDA)." The Consortium's NLP algorithm generates a proportionality score indicating the significance of each topic in an article. In the previous example, the Consortium's algorithm may determine that 50% of the article relates to "RNA sequencing" and 45% to "cancer genomics," but only 5% to the FDA. Re-training the Consortium's NLP algorithm on the set of daily *user-article* interactions to, e.g., expand the set of topics is both financially expensive and time-consuming. We thus take the topic decomposition as a fixed component of our analysis.

After running its NLP algorithm on the high-dimensional dataset, the Consortium constructs a lower dimensional dataset of daily *domain-topic* interactions. Prior to this, the Consortium applies several filters to streamline the data and focus solely on meaningful content interactions, including bot-detection and proportionality thresholds. For example, topic proportionality must not only be sufficiently high on a per-article basis, but also repeat with high enough historical frequency for the interaction to be included. On a typical day, this *domain-topic* dataset features roughly two million domains and 7,000 topics. Finally, as we are interested in understanding what the allocation of employee time in the cross-section tells us about firm risk, we focus on the number of employees who interact with a given topic each day rather than the number of interactions with a topic. This choice further de-emphasizes anomalous interactions that may come from bots.

Figure 2 illustrates this process and shows a fictional domain (xyz.com) with three users on 11/17/2018. Each user reads the same *Wall Street Journal* (WSJ) article, one user also reads an article on microchip.com, and another user also reads an article by *Bloomberg*. Each publisher feeds this user-article interaction data to the Consortium, which applies its NLP algorithm to determine that the microchip.com article is entirely about CPUs, the *WSJ* article is a 30%/70%

split between inflation and the Federal Open Market Committee (FOMC), and the *Bloomberg* article is a 20%/80% split between inflation and politics. The Consortium then aggregates these interactions across users and topics. On this day, three users at xyz.com were focused on FOMC and inflation-related news, whereas only one was focused on CPU and Politics.

This dataset of domain-topic interactions provides us with details on the set of topics that each firm is paying attention to daily. However, directly using this daily data in our analyses is problematic because reading activity displays a variety of intra-week effects. For example, there are significantly fewer user-article interactions on weekends than on weekdays, and a different composition of topics read on Mondays versus Fridays. Figure OA.3.1 in the Online Appendix highlights this intra-week pattern in reading by showing that on an average Tuesday, 93 unique users per firm (across all roughly two million domains) interact with the Consortium's data, while on Fridays, the number falls to 83. We address these intra-week patterns in reading by either applying day-of-week fixed effects or aggregating the daily domain-topic interaction data into lower frequencies, such as quarterly or annually, when conducting our analysis.

2.2 Exploring the Data

The Consortium's raw data cover user-level interactions with each topic. This taxonomy, the members of the Consortium, and the number of covered firms, however, have all evolved over time. Panel A of Table 1 provides an overview of this evolution and presents annual summary statistics for (a) the number of unique domains in the data, (b) the number of topics in the Consortium's taxonomy, and (c) the cross-sectional distribution of employee attention across topics.

There are two takeaways from Panel A, which covers both public and private firms. First, there is a substantial degree of cross-sectional and time-series heterogeneity in the Consortium's coverage of firms and topics. The number of domains (topics) covered by the Consortium increased from about 650,000 (2,500) in 2016 to about 2.1 million (7,400) in 2022. This trend reflects the fact that the Consortium has increased both its member base and the set of topics it covers. Second, Panel A also shows that the distribution of the number of topics that firms pay attention to in any given week is highly positively skewed. The mean (median) firm pays attention to about 330 (130) topics per week. A skewness of around four suggests that a majority of firms engage with just a small subset of the topics.

Panel B repeats the exercise after matching the Consortium's dataset to the CRSP/Compustat universe of firms—the set of domains over which we conduct our empirical analyses in Sections 4 and 5.[6] Employees of public firms pay attention to a broader array of topics. The mean (median) number of topics a public firm pays attention to is around 2,800 (2,900). Consequently, the skewness of the firm-topic distribution diminishes among public firms. Comparing the results in Panel A to those in Panel B suggests that the number of topics that a firms' employees pay attention to is inherently correlated with the firm's size.

Table 1 shows that the number of firms that interact with the Consortium's data and the number of topics that the Consortium covers generally increase over time. These changes in the composition of the Consortium's data could potentially distort the time-series dynamics of a firm's attention to any particular topic. We address this issue by focusing on cross-sectional variation in relative attention to topics across firms on a given date. This ensures that our results are immune to problems associated with changes in (i) the Consortium's algorithm used to decompose articles into topics, and (ii) the technological landscape (e.g., the possibility that users read more news simply because a faster version of the *Wall Street Journal* app becomes available).

Visualizing the topics. While the Consortium's raw data typically features over 7,000 individual topics, a careful analysis of the Consortium's taxonomy indicates that some topics are more related to one another than others. To illustrate this point, the Consortium provided us with category labels associated with each topic. For example, the individual topics such as "M&A," "M&A due diligence," and "capital injection" fall under the "corporate finance" category, whereas the "succession planning" and "layoffs" topics fall under the "staff departure" category.

To provide a sense of the types of topics covered by the Consortium's data, we count how many topics are associated with each category on a randomly selected week in the middle of the sample period (the week ending November 17, 2018). We present these counts as a word cloud in Figure 3, with category labels weighted such that those more prominently featured are a larger set of topics. The figure shows that certain categories (e.g., technology and financial services) tend to feature many more topics than other categories (e.g., urban planning). The distribution of topics across

[6] We only include firms with a CRSP share code of 10, 11, or 12 and a CRSP exchange code of 1, 2, or 3. This confines our analyses to the public equity of firms listed on the NYSE, AMEX, and NASDAQ exchanges. Moreover, we link firms in the CRSP/Compustat universe to the Consortium's dataset via the firm's domain(s). While we remove financial firms and utilities from all empirical tests that involve the CRSP/Compustat universe, we still report attention-related statistics for firms in these industries in Tables 1 and 2 for the purpose of completeness.

categories is not uniform; it is comforting that it mirrors industries that are a larger proportion of the economy, assuaging concerns of inherent industry biases in the data.

Summary statistics by industry. As the category coverage tilts towards finance, business services, and technology, we report the quality of firm matches between CRSP/Compustat and the Consortium's data. Table 2 shows summary statistics by industry groups. We match each firm to one of 17 two-digit NAICS industry codes. While we report statistics for financials and utilities, we remove these sectors in our empirical analysis following convention in the literature.

The last row of this table shows that we can successfully match 86% of the 4,198 firms that exist in the CRSP/Compustat universe between 2016 and 2022. The matched firms are typically larger, representing an average of about 88% of the aggregate market capitalization. Non-price-based measures of firm size, such as the number of employees, show that the matched firms comprise approximately 85% of the total number of employees in the CRSP/Compustat universe. The second to last column also indicates that the employees within an industry interact with about 45% of the topics on average in the Consortium's data on a given day.

The preceding rows of this table show the same summary statistics for each of the 17 industry groups. We observe a large degree of heterogeneity: our sample includes only 10 agriculture-related firms but 1,760 manufacturing-related firms. That said, our matched sample still includes the bulk of each industry's market capitalization. For instance, the 10 agricultural firms in our sample represent 99% of the total market capitalization of the agriculture sector. This pattern is generally consistent across industries, with notable exceptions being the education sectors, with only 65% of its total market capitalization matched. Lastly, the table reveals varying levels of attention paid to topics by firms in each sector. While firms in education interact with only about 15% of topics on a given day, the employees of the typical retail-based firms interact with over 68% of topics.

Additional statistics. For the sake of brevity, Section OA.4 in the Online Appendix provides a battery of additional summary statistics and descriptions of the Consortium's data. We thoroughly examine the distribution of *domain-topic* interactions and show that many topics are only marginally informative about a firm's business line(s). Unsurprisingly, a large proportion of the online reading activity of the average firm is spent on current events. For example, in the week ending November 17, 2018, the preponderance of online time was spent on general topics including "South by Southwest," a popular festival and conference related to music and film, and "Call of

Duty," a popular video game which had released a new version shortly before this date. A key insight from this analysis is that rudimentary measures of firm-level attention, such as the total number of interactions scaled by assets or the number of employees, are most likely uninformative about the firm's economic exposures, something we address in the next section.

3 Motivating our measure of exposure to macroeconomic risk

In this section we show statistical evidence that the composition of reading by employees has a strong relationship to a firm's macroeconomic risk. Section 3.1 uses a set of time-series analyses to demonstrates that the average firm's relative attention to macroeconomic conditions is highly correlated with fluctuations in well-known proxies that capture the state of the business cycle, including the corporate default spread, the VIX index, and the EPU measure proposed by Baker et al. (2016). Our study, however, also seeks to understand how firm-level attention to macroeconomic-related news in near real time can be used to measure differences in *exposure* to macroeconomic risk. In Section 3.2 we formulate a data-driven method to differentiate between topics that are informative and uninformative about a firm's economic exposure to macroeconomic risk, notwithstanding the average employee's focus on reading about general news and events.

3.1 Attention to macroeconomic-related topics

As noted above, firms pay attention to a broad array of topics. To constrain their attention to only macroeconomic-related topics, we first create a sample corpus of articles using a procedure similar to that of Baker et al. (2016). This corpus includes more than 2,500 articles published by the *WSJ*, *The Economist* (segmented by six section tabs), *Financial Times*, Federal Reserve Beige Books (segmented in 12 regions), Federal Reserve Notes, and the Bank of International Settlements. We retain articles that include the terms "economic," "economics," "economical," and "economy." We then have the Consortium deploy its proprietary NLP algorithm onto this macroeconomic-related corpus. The topics that emerge from this exercise are an approximate 600-topic subset of the roughly 7,000 total topics in the Consortium's taxonomy.

The key difference between our approach and that of Baker et al. (2016) is that their corpus is generated by further intersecting the search terms with both "uncertainty" and other policy-related terms. This distinction is important for two reasons. First, the Consortium supplies our data at the topic level. As such, we cannot measure how much of a given article is related to economic uncertainty specifically, versus economics generally. Second, firm concerns about macroeconomic conditions could reflect worries about either a slowdown in economic growth (a first-moment effect) or an increase in economic uncertainty (a second-moment effect). It is thus an empirical question as to which of these quantities most closely relate to future variations in reading activity. Rather than distinguishing between mechanisms in this section, we focus instead on characterizing firm-level reading activity, broadly, and defer addressing this question to Section 5.1.

Panel A of Figure 4 shows the types of macroeconomic-related topics that emerge from this analysis (more prominently read topics are represented as bigger words). The set of popular topics that emerge from the articles underlying the corpus are related to a broad range of categories. In the language of Ludvigson et al. (2021), several prominent topics are related to the real side of the economy (e.g., "consumer spending," "economic growth," and "economic inequality"), while others are related to financial markets (e.g., "exchange rate," "interest rate," and "market volatility"). This reflects the fact there is no single facet of macroeconomic risk. Rather, the articles underlying the corpus reflect a wide variety of concerns that market participants face.

To illustrate how average reading of the macroeconomy changes in the time series, we first designate $Total$, $Macro$, and $Other$ as the sets of total topics, macroeconomic-related topics, and the remaining set of other topics, respectively,

$$Macro \cup Other = Total \qquad \text{and} \qquad Macro \cap Other = \varnothing.$$

For each firm, we stack its reading into vectors associated with $Macro$ and $Total$. Each element j is the number of the firm's employees reading that topic. Topics not read by a firm's employees are given a value of zero. For reasons that will become obvious in the next section, we refer to these vectors as a firm's topic-frequency or $\boldsymbol{tf}$ vector. Our measure of attention to the macroeconomy for firm i at time t is the dot product of the macroeconomic and total vector of reading,

$$\text{TS-RA}_{i,t} = \cos\left(\boldsymbol{tf}_{i,t}^{\text{Macro}}, \boldsymbol{tf}_{i,t}^{\text{Total}}\right) = \frac{\boldsymbol{tf}_{i,t}^{\text{Macro}} \cdot \boldsymbol{tf}_{i,t}^{\text{Total}}}{\|\boldsymbol{tf}_{i,t}^{\text{Macro}}\| \times \|\boldsymbol{tf}_{i,t}^{\text{Total}}\|}, \tag{1}$$

where $\|\boldsymbol{v}\|$ are the Euclidean (or ℓ^2) norm of vector $\boldsymbol{v}$. As the elements of $\boldsymbol{tf}_{i,t}^{Macro}$ are members of $\boldsymbol{tf}_{i,t}^{Total}$ and equation (1) is bounded between zero and one, TS-RA$_{i,t}$ effectively represents the proportion of macroeconomic-related reading to total reading by firm i at time t.

We focus on *relative* attention because of the highly non-linear relationship between firm size and attention to topics. For instance, a large corporation such as Microsoft, with a workforce of 150,000 employees, will naturally consume more information than a smaller corporation like Malibu Boats, with 600 employees. However, a smaller, but still large corporation, such as Red Hat, does not necessarily engage with significantly fewer topics than Microsoft, despite being a tenth of its size. We provide a graphical illustration of this relationship by plotting the average log length of $\boldsymbol{tf}_{i,t}^{Macro}$ and $\boldsymbol{tf}_{i,t}^{Total}$ versus log firm size in Figure OA.3.2. Due to the non-linearity, reading cannot be compared across firms by simply dividing $\boldsymbol{tf}_{i,t}^{Macro}$ by size proxies such as the number of employees, assets, or market capitalization. What is also clear in the figure, however, is that both the length of $Macro$ and $Total$ follow similar patterns across the distribution; the dot product of the two vectors thus does not show an obvious relationship with size.

We next show that during poor economic times, employees tend to read more macroeconomic- versus other business-related news. We illustrate this fact by plotting TS-RA$_{i,t}$ aggregated across all public firms at time t,

$$\widetilde{\text{TS-RA}}_t = \frac{\overline{\boldsymbol{tf}}_t^{Macro} \cdot \overline{\boldsymbol{tf}}_t^{Total}}{\|\overline{\boldsymbol{tf}}_t^{Macro}\| \times \|\overline{\boldsymbol{tf}}_t^{Total}\|}, \tag{2}$$

where $\overline{\boldsymbol{tf}}_t = \frac{1}{N_t}\sum_{i=1}^{N_t} \boldsymbol{tf}_{i,t}$ and N_t is the number of public firms at time t.

Figure 5 demonstrates a substantial amount of time-series and cross-section heterogeneity in the degree to which firms pay attention to macroeconomic conditions. The solid blue line in the figure reports the time series of $\widetilde{\text{TS-RA}}_t$ from equation (2). On average, employees allocate about half of their attention to macroeconomic topics across the sample. Their attention to these topics, however, rises substantially during "bad" times, such as the political uncertainty surrounding the 2016 Presidential elections and during the COVID-19 pandemic in mid-2020. The proportion of attention towards macroeconomic-related topics for the average firm peaked around 62% during these periods. The dashed blue lines report the average values of TS-RA$_{i,t}$ among firms with a relative attention measure that is 25% above or below the mean at each point in time. The takeaway from this analysis is that attention to macroeconomic conditions also varies widely across firms.

For instance, while the mean firm allocated about 60% of its attention towards macroeconomic news in mid-2020, a substantial number of firms had proportions greater (less) than 0.75 (0.45).

Given that a firm's reading about macroeconomic conditions is likely to reflect the employees' concerns about the state of the economy, we expect $\widetilde{\text{TS-RA}}_t$ to covary with the business cycle. As motivation, Figure 6 plots the time series of $\widetilde{\text{TS-RA}}_t$ alongside the corporate default spread (top panel) and the EPU index of Baker et al. (2016) (bottom panel). To aid comparability, each variable is standardized. The average firm's employees pay relatively more attention to macroeconomic conditions when either the corporate default spread or the EPU index is relatively high. Although the dynamics of $\widetilde{\text{TS-RA}}_t$ and these business-cycle proxies are highly correlated, they are not perfectly aligned. For instance, $\widetilde{\text{TS-RA}}_t$ spiked prior to the onset of the COVID-19 pandemic in early 2020, and tended to rise, but not as much as either alternative proxy, at the start of the COVID-induced recession of March 2020. The fact that employee reading about macroeconomic news is not perfectly aligned with variables that measure different aspects of aggregate macroeconomic conditions opens the possibility that our measure, which captures the consumption of news, contains additional firm-level information.

To statistically clarify the relationship between macroeconomic conditions and employee reading, we next regress $\text{TS-RA}_{i,t}$ onto the realization of various business-cycle variables,

$$\text{TS-RA}_{i,t} = \text{Day-of-Week}_t + \beta \cdot \text{MacroV}_{t-1} + \epsilon_{i,t}, \tag{3}$$

where MacroV_{t-1} is a proxy for macroeconomic activity at time $t-1$, measured using either the corporate default spread, the intermediary capital ratio of He et al. (2017), the WTI oil returns, the EPU index of Baker et al. (2016), the VIX index, or the macroeconomic uncertainty index of Bekaert et al. (2022). While the first three variables are closely related to the level of economic activity, the latter three are tightly linked to the notion of economic and financial market uncertainty. The regressions are conducted at a daily frequency and include day-of-week fixed effects to control for intra-week seasonality in reading data (see Section 2). All business-cycle variables are standardized. As such, the coefficient values reflect the change in a firm's relative attention after a one-standard-deviation change in each underlying proxy for macroeconomic activity. Finally, standard errors are clustered at both the firm and date level.

Table 3 presents the results of regression (3) in columns (1), (3) and (5). For example, Panel A shows that the typical firm increases its relative attention to macroeconomic news by about 0.015 (i.e., 1.5 percentage points) in the days following a one-standard-deviation increase in the corporate default spread. Moreover, firms generally read less about macroeconomic conditions in the days following improvements in the health of financial intermediaries and increases in oil prices. Panel B shows that firms also read more about macroeconomic conditions in the days following increases in uncertainty. For example, the typical firm's relative attention increases by 0.007 (0.013) when the VIX (EPU) index spikes. Interestingly, the economically strongest relationship between reading and macroeconomic conditions holds with the uncertainty measure of Bekaert et al. (2022), which structurally separates variations in uncertainty from variation in risk aversion.

3.2 Cross-sectional differences in exposure to risk

While the previous section shows that $\widetilde{\text{TS-RA}}_t$ has a positive time-series correlation with several proxies for economic conditions, we are primarily interested in understanding what employee reading activity tells us about cross-sectional differences in firm-level exposures to macroeconomic shocks. A complication, however, arises when exploring this idea: most of the relative attention to macroeconomic topics, which is captured by $\text{TS-RA}_{i,t}$, is common across firms (see Panel A of Figure 4). This common component of reading can relate to both the supply of news (e.g., publishers write more articles and firms' employees read more articles related to macroeconomic conditions in bad times) and the demand for news (e.g., employees worried about the their uninsurable labor risk read more about macroeconomic conditions in bad times). To identify the cross-sectional differences in risk exposure that we are interested in studying, we thus need to differentiate between topics that are more or less informative about a firm's latent risks.

On each day of the sample period, we differentiate between topics by adopting an approach from the field of document retrieval in computer science. Tasks, such as search operations, depend on word sequences within a query to extract related documents from a large corpus. Central to this process is the concept of the *tf-idf* (term frequency-inverse document frequency) score, where each word w within a document d is assigned a score defined as:

$$tf\text{-}idf_{d,w} = \underbrace{\frac{\#\text{ Word in Document}}{\text{Total Words in Document}}}_{\text{Term frequency (tf)}} \times \underbrace{\frac{\#\text{ Documents in Corpus}}{\#\text{Documents featuring Word}}}_{\text{Inverse document frequency (idf)}}. \tag{4}$$

The *tf-idf* scores emphasize the importance of unique words while downplaying the importance of common ones. For example, prepositions frequently appear *within* (i.e., have high *tf*) and *across* (have low *idf*) documents, so this widespread use diminishes their usefulness in distinguishing between documents in the corpus. Consequently, prepositions generally have low *tf-idf* scores.

We build on this idea to propose a novel measure for differentiating between informative and uninformative topics for a firm. We refer to this measure as the *topic frequency-inverse aggregate frequency*, or *tf-iaf*, score. The scoring strategy downweights common topics, such as those associated with current event (since they pervade reading across firms), and upweights topics with high readership within each firm. This enables us to identify topics that distinguish a firm's reading activity from those of other firms. We define this score for topic j and firm i at time t as:

$$\begin{aligned} tf\text{-}iaf_{i,j,t} = & \underbrace{\begin{pmatrix}\text{Number of Employees at Firm } i \\ \text{Interacting with Topic } j \text{ at time } t\end{pmatrix}}_{\text{Topic frequency (tf)}} \times \\ & \underbrace{\begin{pmatrix}\text{Average Fraction of Employees Across All} \\ \text{Firms Interacting with Topic } j \text{ at time } t\end{pmatrix}^{-1}}_{\text{Inverse aggregate frequency (iaf)}} \\ \equiv & \; tf_{i,j,t} \times iaf_{j,t}, \end{aligned} \tag{5}$$

where topic j is either in the macroeconomic or other subset of topics on date t. Here, the *tf* component is the number of total users within a firm interacting with a topic on a given day. This component is identical to the inputs in equation (1), which places a higher importance on topics with a larger number of interactions. In contrast, the *iaf* component downweights the topics that receive a large number of interactions across all firms. This effectively places a lower weight on topics that all firms are reading because of the availability of news at time t and allows us to focus on topics that better reflect a firm's demand for news. For example, if many publishers write about the unemployment rate on the first Friday of each month when the Bureau of Labor Statistics

(BLS) releases its employment data, and most employees at firms are reading about this current event, then employment-related topics will receive low *iaf* scores on those days.

With this logic in mind, we measure firm i's exposure to macroeconomic risk on day t via the degree to which its employees are paying attention to differentiated macroeconomic topics relative to differentiated other business topics,

$$\text{CS-RA}_{i,t} = \frac{\boldsymbol{tf\text{-}iaf}_{i,t}^{Macro} \cdot \boldsymbol{tf\text{-}iaf}_{i,t}^{Total}}{\|\boldsymbol{tf\text{-}iaf}_{i,t}^{Macro}\| \times \|\boldsymbol{tf\text{-}iaf}_{i,t}^{Total}\|}. \tag{6}$$

Here, $\boldsymbol{tf\text{-}iaf}_{i,t}^{Macro}$ is the vector of firm specific *tf-iaf* weights on macroeconomic topics from equation (5) and $\boldsymbol{tf\text{-}iaf}_{i,t}^{Total}$ is the firm-specific vector of *tf-iaf* weights on the total reading vector. The time index t represents attention to each set of topics over a daily (or lower) frequency. We refer to this quantity as our cross-sectional measure of firm i's *relative* attention towards macroeconomic conditions at time t, or $\text{CS-RA}_{i,t}$.

Motivational evidence. To show that $\text{CS-RA}_{i,t}$ meaningfully distinguishes between firms with high and low exposures to macroeconomic risk, we perform a daily event study around the onset of the COVID-19 pandemic in early 2020. Specifically, we estimate $\text{CS-RA}_{i,t}$ on February 26, 2020, which was the day of the first non-travel-related infection of COVID-19 in the United States. This variable remains constant across our event study.

We then map this variable to be uniformly distributed over the interval $[-1, 1]$ (to remove the effect of any potential outliers in reading) and compute a weighted average value of $\widetilde{\text{TS-RA}}_t$ from equation (2) among firms with positive and negative values of $\text{CS-RA}_{i,2/26}$. Within each group, the absolute value of $\text{CS-RA}_{i,2/26}$ determines the weight that is placed on the reading of firms within each portfolio. Looking at the difference in $\widetilde{\text{TS-RA}}_t$ between firms with positive and negative values of the uniformly distributed $\text{CS-RA}_{i,2/26}$ allows us to evaluate how the macroeconomic-related reading of each group changed as concerns over the COVID-19 pandemic evolved.

The results, which are reported in the left-hand side of Figure 7, show that prior to February 26, high $\text{CS-RA}_{i,2/26}$ firms tended to allocate about four percentage points more of their relative attention to macroeconomic topics. As concerns related to the pandemic spread throughout March, these same firms increased their reading of macroeconomic news by a statistically significant amount. By March 17, a few days after the Trump Administration declared a nationwide emergency and issued additional travel bans against non-U.S. citizens, high CS-RA$_{i,2/26}$ firms were allocating about seven percentage points more of their relative attention towards macroeconomic news, a quantity almost double the unconditional difference.

The fact that the relative attention of firms that are highly exposed to the macroeconomy, as determined by CS-RA$_{i,2/26}$, increased upon the onset of a macroeconomic shock is consistent with the notion that this variable reflects a firm's exposure to macroeconomic risk. An additional implication of this risk-based narrative is that the returns of the highly exposed CS-RA$_{i,2/26}$ firms should deteriorate once the macroeconomic shock is realized. The right-hand side of Figure 7 plots the difference in cumulative returns of high versus low CS-RA$_{i,2/26}$ firms and shows that this is indeed the case. The employees of firms that are highly exposed to macroeconomic conditions not only read more macroeconomic news upon the onset of the COVID-19-related recession, but also experienced significantly lower returns than low CS-RA$_{i,2/26}$ firms over the sample. Strikingly, these differences in reading and returns were flat before February 26.

Statistical evidence. To establish that CS-RA$_{i,t}$ is a valid measure of firms' exposures to macroeconomic risk, we hypothesize that if a firm is more (less) exposed to economic conditions, then changes in the firm's relative attention to macroeconomic news should coincide more strongly (weakly) with changes in common proxies for economic conditions. That is, firms with high (low) values of CS-RA$_{i,t}$ in the recent past are likely to read relatively more (less) about macroeconomic news when economic conditions worsen (improve).

$$\text{TS-RA}_{i,k,t} = \alpha_{k,t} + \beta \text{CS-RA}_{i,k,t-1} + \gamma \text{MacroV}_{t-1} \times \text{CS-RA}_{i,k,t-1} + \epsilon_{i,k,t}. \tag{7}$$

Here, TS-RA$_{i,k,t}$ (CS-RA$_{i,k,t-1}$) is the time-series (cross-sectional) measure of the relative attention of firm i in industry k on day t, and MacroV$_{t-1}$ is one of the same six proxies for macroeconomic conditions underlying equation (3). Moreover, $\alpha_{k,t}$ is an industry-by-time fixed effect, in which the index k corresponds to the three-digit NAICS industry of firm i. These fixed effects subsume both the average effect of macroeconomic conditions on relative attention, as quantified by estimating equation (3), and the day-of-week fixed effects included in the aforementioned equation.

The key parameters of interest are β and γ. The former parameter reflects any unconditional differences in relative attention to macroeconomic news between firms in the same industry with different exposures to economic conditions. The latter parameter estimates the degree to which firms with high versus low exposures to macroeconomic conditions change the composition of their reading on day t due to changing macroeconomic conditions on day $t-1$. To minimize the effect of outliers and focus on cross-sectional differences in exposure between firms, we map CS-RA$_{i,k,t-1}$ to lay in the interval $[0,1]$ on each day t. As a result, the estimated value of γ reflects how relative attention changes between firms in the 0^{th} to 100^{th} percentile of CS-RA$_{i,k,t}$.[7] We present the results of this analysis in the second and third rows of Panels A and B in Table 3.

To begin, all specifications show that the employees of firms that are highly exposed to macroeconomic conditions read substantially more macroeconomic news. Moving from the lowest to the highest value of CS-RA$_{i,k,t-1}$ increases the proportion of macroeconomic news consumed by more than 6%. Second, the estimated values of γ in columns (2) and (4) suggest that the employees of high CS-RA$_{i,k,t-1}$ firms also tend to read relatively more macroeconomic news in the days following an increase in the corporate default spread or a decrease in the health of the balance sheets of financial intermediaries, two signals for weak economic conditions. Focusing on intermediary capital shows that the employees of highly exposed firms reduce their relative attention to macroeconomic news by almost 3% in the days following a one-standard-deviation increase in the intermediary capital ratio, an effect that is almost 50% larger than the unconditional effect in Column (3).

While the relation between the composition of reading and risk exposure is corroborated when we examine changes in the corporate default spread in Column (2), Column (6) suggests that firms that read more about macroeconomic risk in the day following an increase in oil prices, which are often considered good times. First, although the sign of this relation is inconsistent with the results in Columns (2) and (4), its economic magnitude is relatively small. Second, as discussed in Alfaro et al. (2024), the effects of a shift in oil prices on firm-level prospects is highly heterogeneous. For example, falling oil prices are unambiguously negative for oil producers, but positive for airlines.[8]

Finally, Columns (2), (4), and (6) in Panel B show that similar results emerge when we include a measure of macroeconomic uncertainty on the right-hand side of equation (7). As economic

[7] Specifically, if CS-RA$_{i,t-1}$ denotes the raw relative attention for firm i between $t-1$ and t, then its rank transformation is $F(\text{CS-RA}_{i,t-1}) = \text{Rank}\,(C_{i,t-1})\,/(N_t+1)$, where N_t is the number of firms at time t, $\text{Rank}\,(\min_{i=1,\ldots,N_t} C_{i,t-1}) = 1$ and $\text{Rank}\,(\max_{i=1,\ldots,N_t} C_{i,t-1}) = N_t$. This transformation implies that the α-quantile of $F(\text{CS-RA}_{i,t-1})$ is α. For notational simplicity, we continue to refer to the rank transformed value as CS-RA$_{i,t-1}$ in our tables and regression specifications.

[8] We exploit this heterogeneity when differentiating the effects of first- and second-moment shocks in Section 5.1.

uncertainty increases, measured using either the VIX, EPU, or Bekaert et al. (2022) uncertainty index, then high CS-RA$_{i,t}$ firms tend to increase their reading of macroeconomic topics by 1.0 to 2.0% more than low CS-RA$_{i,t}$ firms. Consistent with the results in Panel A, this shows that CS-RA$_{i,t-1}$ captures heterogeneity in how firm employees change the composition of their reading in response to a change in macroeconomic conditions—i.e., CS-RA$_{i,t-1}$ on day $t-1$ for firm i is a timely reflection of the firm's exposure to macroeconomic risk at time t.

3.2.1 Discussing the relative attention measures

The preceding analyses delivers two takeaways: (i) firm employees change their attention to macroeconomic news in response to changes in economic conditions, and (ii) firms with high CS-RA$_{i,t}$ shift reading towards macroeconomic news during bad times more so than those with low CS-RA$_{i,t}$. This raises several follow-on questions: (i) is there a change in the set of "salient" topics due to the *tf-iaf* adjustment? and (ii) if so, does this change tell us anything important about the firm? This sections explores the distribution of topics across the TS-RA$_{i,t}$ and CS-RA$_{i,t}$ measures. We then use the insights gleaned to develop a framework to better understand the economic content of the measures.[9]

To illustrate the effect of the *tf-iaf* adjustment, we first focus on non-macroeconomic topics. Consider the topic clouds presented in Figure 8. The figures in the top row display the reading for the chemicals manufacturing (NAICS 334) industry, while those in the bottom row show the reading for the computer and electronics manufacturing (NAICS 325) industry. The word clouds in the first column are constructed using the average topic frequency within each industry for the week ending November 17, 2018. These figures show virtually no differences in reading between the two industries. Prominent topics include "live streaming," "South by Southwest," "blu-ray," "United States Secret Service," and "Call of Duty." In the parlance of the *tf-iaf* scores, these topics, however, will also have low inverse aggregate frequency (*iaf*) scores because all employees seem to interact with these topics often.

The second column of Figure 8 demonstrates how these raw word clouds change when we weight each topic by its *tf-iaf* score from equation (5). The salient topics are now very industry specific.

[9]Similar to how the cross-sectional asset-pricing literature examines how heterogeneity in firm characteristics (e.g., book-to-market) reflect differences in risk exposures across firms, we examine what differences in firm-employee reading, measured using equation (6), tell us about heterogeneity in firm-level risk (see, e.g., Zhang, 2005).

For instance, we see that the "cancer genomics," "drug discovery," and "angiogenesis inhibitors" topics are important for firms in the chemicals sector, which includes pharmaceutical firms, whereas the "Cisco ACI," "remote desktop protocol," and "software defined perimeter" topics are important for firms in the computer industry. The procedure enhances the weight of topics such as "cancer genomics" in the chemicals industry and "Cisco ACI" in the computing industry, because many read these topics within the respective industries, while few read the topics outside those industries. Broadly speaking, the *tf-iaf* adjusted topics that emerge are closely related to the production and investment decisions of firms in each industry.

We also apply the *tf-iaf* adjustment to the set of topics drawn from the macro corpus as move from TS-RA$_{i,t}$ to CS-RA$_{i,t}$. Panel B of Figure 4 shows the *tf-iaf* weighted macro-related word cloud for the week ending November 17, 2018. While the raw word cloud in Panel A of Figure 4 shows that employees in many firms read about "quantitative easing," "consumer spending," and "economic growth," these prominent topics are relatively uninformative in distinguishing reading about macroeconomic risk in the cross-section of firms. Rather, what distinguishes reading among firms are finance-related terms that come in two broad categories: (i) topics such as "credit risk," "exchange rate," and "duration management" that closely relate to firm risk management, and (ii) topics such as "Basel II" and "International Accounting Standards Board" that closely relate to regulatory compliance.[10] Both categories of topics are generally related to corporate hedging and risk mitigation activities. While these finance-related terms emerge as important macro-related topics, we reiterate that both finance and utility firms are excluded from both the main analysis and the set of firms underlying these word clouds.

Theoretical framework. How does CS-RA$_{i,t}$, which seems to be more related to the proportion of mitigation- versus investment-related reading, versus TS-RA$_{i,t}$ link to a firm's risk exposure? And would we expect this measure tell us anything else about firm-level outcomes? A simple two-period framework helps answer these questions. In this framework, the firm's value

[10]Table OA.3.2 reports the top 20 macro-related news topics that received the most attention based on their *tf-iaf* scores from 2016 through 2022. The topics are ranked by their average *tf-iaf* weights, computed at the three-digit NAICS industry level to mitigate extreme weights. The set of topics over this extended period is consistent with the broad risk management and regulatory compliance themes observed in the daily word cloud.

today is

$$
\begin{aligned}
V_t = & \text{Production}_t - \text{Mitigation Cost}_t \\
& + E_t\left[M_t\left(\text{Production}_{t+1} + \text{Mitigation Benefits}_{t+1}\right)\right].
\end{aligned} \tag{8}
$$

That is, the firm's value is a function of output today (Production_t) and the discounted value of output tomorrow ($M_t \cdot \text{Production}_{t+1}$). Firms also have the ability to mitigate risk. While mitigation helps hedge realizations of bad states of the world tomorrow ($\text{Mitigation Benefits}_{t+1}$), it comes with a cost that is incurred today (Mitigation Cost_t). This implies a tradeoff for the firm: higher mitigation expenditure implies fewer resources for the firm to invest in next period's production. Moreover, riskier firms—those whose cash flows covary more with the business-cycle (captured by the stochastic discount factor, M_t)—will be incentivized to hedge (invest) more (less).

In short, if $\text{CS-RA}_{i,t}$ reflects firm-level risk exposure, then this measure should (i) positively predict hedging and mitigation activity, (ii) positively predict the cost of capital, and (iii) negatively predict investments in capital. This intuition is related to that in the literature linking macroeconomic risk to risk mitigation (e.g., Brown, 2001; Hong, Wang, and Yang, 2023) and regulatory exposure (e.g., Kalmenovitz, 2022). The following sections examine the usefulness of $\text{CS-RA}_{i,t}$ for predicting prominent firm-level outcomes and show results that are consistent with the main implications of this framework.[11]

4 Relative attention and firm outcomes

In this section we conduct a set of empirical exercises to substantiate the informativeness of our relative attention measure in quantifying firm-level exposures to changing macroeconomic conditions. Section 4.1 shows that firms with high $\text{CS-RA}_{i,t}$ attempt to mitigate risk through either hedging or regulatory activity more aggressively than firms with low $\text{CS-RA}_{i,t}$. Section 4.2 builds on this result by showing that these firms also tend to be riskier. Consistent with higher cost of capital, Section 4.3 shows that $\text{CS-RA}_{i,t}$ predicts lower sales, and lower investment and employment growth, in subsequent periods.

[11] Section OA.2 of the Online Appendix formalizes this intuition in an economic model. An extension that includes employee learning also generates comparative statics that map to our results from Sections 3.1 and 3.2.

4.1 Risk mitigation and compliance

Section 3.2.1 shows that a consequence of applying the *tf-iaf* weights from equation (5) to cross-sectionally differentiate macroeconomic-related reading across firms is that salient topics shift from aggregate concerns, such as "Quantitative Easing" (QE) and "Consumer Spending," towards topics related to firm efforts to mitigate risk and/or comply with regulatory authorities (e.g., "Duration Management," "Basel II," and "International Accounting Standards Board").

Hedging. To link this fact to firm behavior, we follow the methodology of Campello et al. (2011) to see how a firm's hedging activity correlates with CS-RA$_{i,t}$ over the previous year. The hedging measure is generated by first counting the number of times each of the following hedging-related keywords is mentioned in a 10-K: "derivatives," "hedge," "financial instrument," "swap," "market risk," "expos," "futures," "forward contract," "forward exchange," "option contract," "risk management," and "notional." The total number of occurrences of these keywords is then divided by the total 10-K word count to derive a firm-year proxy of hedging intensity.

As Campello et al. (2011) note, hedging activity plays little role for a large fraction of firms. Consistent with this finding, around 25% of firms mention four or fewer hedging-related words in their 10-Ks. A small number of firms, however, frequently mention these words. To minimize the effects of outliers, we define our measure of hedging using an indicator variable that is denoted by $\mathbb{I}^{\text{hedge}}_{i,t}$. This variable takes on a value of one for firm-year observations that have a hedging intensity measure above the 75th percentile of firms in a given year, and zero otherwise. We use this indicator variable as our dependent variable of interest when examining the relation between hedging intensity and a firm's *tf-iaf* adjusted relative attention to macroeconomic risk:

$$\mathbb{I}^{\text{Hedge}}_{i,k,t} = \psi_{k,t} + \beta_1 \text{CS-RA}_{i,k,t-1} + \beta_2 \text{CS-RA}_{i,k,t-1} \times \Delta\text{EPU}_{t-1} + \boldsymbol{Z}'_{i,k,t-1}\gamma + \varepsilon_{i,k,t}. \tag{9}$$

In this regression, CS-RA$_{i,k,t-1}$ is the average daily value of the relative attention measure of firm i in industry k over the year prior to the firm's fiscal year-end. We average the values over this time period as 10-Ks reflect firm activity over an entire year. Additionally, as discussed in Section 3.2, we also map CS-RA$_{i,t}$ to be between zero and one. The benefit of applying this transformation is twofold. First, the transformed measure of relative attention is less sensitive to the presence of outliers. Second, this transformation allows us to interpret the estimated values of β_1 and β_2 in equation (9) as the difference in the propensity to hedge between a firm paying the least versus the most attention to *tf-iaf* adjusted macroeconomic-related topics.

As a firm's propensity to hedge its risks can depend on factors such as the firm's asset base and indebtedness, the vector $\boldsymbol{Z'}$ contains firm-level controls, including those from Leary and Roberts (2014); the measure of financial constraints from Whited and Wu (2006); and lagged CAPM β, a common measure of a firm's exposure to aggregate risk.[12] The variable $\psi_{k,t}$ reflect industry-by-time fixed effects that subsume not only unobserved heterogeneity in hedging activity, but also common economic shocks each industry may face at a given point in time (e.g., the possibility that hedging differs between technology and manufacturing firms at the onset of the COVID-19 pandemic). Certain specifications also include time- and industry-fixed effects separately. Finally, ΔEPU_{t-1} is the lagged change in the EPU index over the previous year. We standardize this variable for ease of interpretation and focus on its interaction with CS-RA$_{i,k,t-1}$ to examine whether firm hedging activity is especially high during times of high macroeconomic uncertainty (i.e., in "bad times"). Given the short time-series, standard errors are clustered by firm.

Panel A of Table 4 presents the results. The first column displays the regression without controls and fixed effects. The coefficient on CS-RA$_{i,k,t-1}$ indicates that as a firm moves from the 0^{th} to 100^{th} percentile of the CS-RA$_{i,t}$ measure, the probability of the firm being a top quartile hedger increases by approximately 11% (t-statistic of 6.00). In the second and third columns, we add date and industry-by-date fixed effects, respectively. The magnitude and statistical significance of the aforementioned effect drops, but is still extremely strong. Column (4) of Table 4 augments Column (3) by including the interaction between relative attention and the change in EPU.

Furthermore, to show that CS-RA$_{i,t}$ reflects firm-level risk beyond that implied by standard measures, we also include lagged CAPM β as a control. While the unconditional effect of relative attention remains similar to that reported in Column (3), the interaction effect shows that high CS-RA$_{i,t}$ firms are especially likely to hedge their exposure to macroeconomic fluctuations in bad times. The coefficient on this interaction effect is 0.05 (t-statistic of 5.01). In Column (5) we add the control variables from Leary and Roberts (2014) and Whited and Wu (2006). The unconditional effect of relative attention reduces in economic magnitude and statistical significance, yet

[12]The Leary and Roberts (2014) vector of controls include firm size, Tobin's q, profitability, leverage, and asset tangibility. CAPM β is estimated on a 250-day rolling basis. Section OA.1 in the Online Appendix provides details on the construction of these control variables, and Table OA.3.1 in the Online Appendix reports summary statistics.

the interaction between relative attention and economic uncertainty remains economically sizeable and statistically robust. High CS-RA$_{i,t}$ firms are almost 5% more likely to actively hedge their risk in bad times compared to their low CS-RA$_{i,t}$ counterparts in the same industry and at the same point in time.

Compliance. While most topics in Panel B of Figure 4 are related to risk mitigation, several topics are also related to compliance. This implies that firms with high exposures to macroeconomic risk tend to have more onerous regulatory and legal burdens. We formalize this link by examining the relation between CS-RA$_{i,t}$ and the firm-level proxy for spending on regulatory compliance from Kalmenovitz (2022). Specifically, we repeat specification (9), but replace the left-hand side variable with each firm's regulatory intensity. Panel B of Table 4 report the results.

As each column in Panel B shows, the relation between a firm's relative attention to macroeconomic risk and its degree of regulatory intensity is positive and statistically significant. In Column (5), we present our most comprehensive specification that includes industry-by-time fixed effects and the firm-level controls discussed above. The difference in regulatory intensity as a firm moves from the 0^{th} to 100^{th} percentile of the CS-RA$_{i,t}$ is significant at the 1% level (t-statistic of 6.88). The preceding columns show stronger results in specifications that include fewer fixed effects and controls. Overall, this confirms that firms that spend more time reading about macroeconomic-related news tend to be subject to more regulatory oversight.

4.2 Cost of capital regressions

We next examine how our measure of firm-level macroeconomic risk covaries with firm-level discount rates. *If* fluctuations in macroeconomic conditions reflect fundamental risks that firms are differentially exposed to, then those with higher exposures should not only spend relatively more time reading about macroeconomic conditions (i.e., have higher values of CS-RA$_{i,t}$) but should also have higher costs of capital to reflect this increased risk. To shed light on this, we follow the approach of Gebhardt et al. (2001) to infer the cost of equity capital (henceforth, ICC$_{i,t}$) from equity analysts' earnings forecasts. Section OA.1 of the Online Appendix provides details on how we construct the measure. Given our data's short time period, we use implied costs of capital to minimize the effects of noise in realized returns. Table OA.5.5 in the Online Appendix shows that we obtain similar takeaways when we use realized returns instead.

We then regress each firm's quarterly cost of capital, $\text{ICC}_{i,t}$, on the firm's lagged measure of $\text{CS-RA}_{i,t}$ using a similar regression specification to equation (9), but now controlling for characteristics that have been linked to returns. These characteristics include the CAPM β, firm size, Tobin's q (the inverse of value), profitability, and investment tangibility. We estimate these regression using data from the end of each earnings announcement month, because it is around this point in time that analysts make the largest adjustments to earnings estimates for the coming quarters and year.

Results are presented in Table 5. The first column indicates a strong positive relationship between a firm's relative attention to macroeconomic news and $\text{ICC}_{i,t}$. Columns (2) and (3) add combinations of date and industry fixed effects to the regressions and show that the results remain economically and statistically significant at the 1% level even after accounting for fixed differences in $\text{ICC}_{i,t}$ across times and industries, respectively. The magnitude of the point estimates in Columns (1) to (3) suggests that a firm moving from the lowest value to the highest value of $\text{CS-RA}_{i,t}$ sees its cost of capital increase by about 110 to 150 basis points. While the magnitude of this effect diminishes when we consider industry-by-time fixed effects and additional controls in Columns (4) and (5), the basic fact remains the same: firms that pay more attention to macroeconomic-related topics have higher costs of capital than those that pay more attention to other types of news.

4.3 Firm-level real outcomes

Since firms with high $\text{CS-RA}_{i,t}$ are more exposed to macroeconomic risk and have higher cost of capital, one would expect this to translate into lower investment, sales growth, and employment growth rates. That is, we expect that firms that allocate a relatively high proportion of their attention to macroeconomic-related topics will tend to implement more contractionary corporate policies. We establish this fact once again by using the panel regression specification in equation (9) after replacing the dependent variable with various firm-level real outcomes. The point estimate β_1 now tells us how a firm's relative attention is related to corporate decisions that it may take, conditional upon the comprehensive set of fixed effects and control variables discussed in earlier. The variable β_2 captures how these actions may change during especially bad economic times.

Table 6 shows that a higher degree of attention to macroeconomic-related topics is indeed associated with a *contraction* in the average firm's future investment and sales. Importantly, this negative association between a firm's attention and real outcomes holds regardless of whether we include industry-by-time fixed effects, thereby capturing differences in relative attention within each industry, the comprehensive set of firm-level controls from Leary and Roberts (2014), the Whited and Wu (2006) measure of financial constraints, and CAPM β. Column (1) of the table, which features no fixed effects or control variables, shows that increases in relative attention to uncertainty are associated with a 4% decline in one-quarter ahead investment (t-statistic of -11.7), a 14% decline in one-quarter ahead sales (t-statistic of -12.8), and a 6% decline in one-year ahead employment growth (t-statistic of -7.3).[13] Adding time fixed effects to the baseline specification in Column (2) leaves these results largely unchanged and eliminates the concern that the results are driven by a limited number of times when all firms cut investment simultaneously (e.g., at the onset of the COVID-19 pandemic in March 2020). Similarly, the regression results in Column (3) show that augmenting Column (2) with industry fixed effects, thereby accounting for fixed differences in investment opportunities across industries, does little to change the strong negative relation between macroeconomic-related news and investment rates.

Columns (3) through (5) present the results of our most comprehensive empirical specifications. These specifications include industry-by-time fixed effects, thereby controlling for time-varying differences in investment opportunities across industries (e.g., the possibility that the average technology firm benefited from the onset of the COVID-19 pandemic while the average durable goods retailer suffered), the interaction of our measure with bad times (i.e., high uncertainty states), and a large set of firm-level control variables. The results in Column (3) indicate that the robust and negative association between attention to macroeconomic-related topics and investment, sales growth, and employment growth persists even after we account for the unobservable heterogeneity in these real effects across industries in time and CAPM β.

Column (4) indicates that the predictive power of CS-RA$_{i,t}$ is particular strong during times of high uncertainty. Notably, at any given point in time, the average firm in each industry that is allocating most of its relative attention to macroeconomic news has an investment rate that is 3.08% per quarter less than the investment rate of the firm that is paying the least attention to macroeconomic topics. During periods of high economic uncertainty, this difference exceeds

[13] Table OA.3.3 in the Online Appendix considers which components of total assets firms adjust in response to deteriorating macroeconomic conditions; both short-term (i.e., inventory) and long-term (i.e., property, plant, and equipment) assets fall as the relative attention to macroeconomic news rises.

4.50% and is statistically significant at the 1% level. Column (5) repeats this specification after adding the battery of firm-specific controls. The unconditional relationship between CS-RA$_{i,t-1}$ is now smaller for all dependent variables, but is still statistically significant at the 1% level. As additional robustness, Table 8 in Section 5.1 extends these specifications to include alternative measures of firm-level risk elicited from either earnings call transcripts or the options market. CS-RA$_{i,t-1}$ continues to consistently predict these firm-level outcomes.

5 Dissecting the drivers of attention to macroeconomic risk

Our analyses thus far show that firms with employees who pay relatively more attention to macroeconomic conditions hedge more, have higher costs of capital, and subsequently invest less capital and produce less output. While these findings do not differentiate between whether employee concerns are more first- or second-moment related, Section 5.1 shows that fluctuations in economic uncertainty are the primary driver of shifts in attention towards macroeconomic conditions in bad times. We establish this by following the instrumental variables approach of Alfaro et al. (2024). The fact that uncertainty drives a firm's allocation of attention to macroeconomic conditions is also consistent with the theoretical framework in Section OA.2 of the Online Appendix, which links firm risk exposures to the degree to which firms choose to mitigate risk rather than invest. Moreover, Section 5.2 shows that high CS-RA$_{i,t}$ firms have return and accounting characteristics that are associated with risk (e.g., high market betas).

5.1 Relation to uncertainty

Does time variation in firm uncertainty predict changes in employee reading of macroeconomic-related news? We assess this question using the following firm-quarter panel regression:

$$\text{TS-RA}_{i,t} = \psi_{i,t} + \beta_1 \Delta\sigma_{i,t-1} + \beta_2 \text{Ex Returns}_{i,t-1} + \boldsymbol{Z}'_{i,t-1}\boldsymbol{\gamma} + \varepsilon_{i,t}. \tag{10}$$

Columns (1) through (3) of Table 7 present the results of regression (10). Column (1) shows a strong statistical relation between changes in either excess returns (first-moment) or volatility (second-moment) and macroeconomic-related reading. The resulting coefficients have the expected signs: a one-standard-deviation increase in volatility (excess return) relates to a 0.7% increase (0.5% decrease) in the proportion of macroeconomic reading. These magnitudes are similar to those shown in Table 3, which relates reading to fluctuations in aggregate macroeconomic quantities. Column (2) adds both firm and date fixed effects to the regression; while the coefficient associated with volatility remains statistically significant, the coefficient for excess returns becomes insignificant. The same takeaway emerges from Column (3), which includes firm-level controls.

There is, however, an endogeneity concern underlying regression (10). Notably, an omitted variable could affect both *firm-level* volatility and a firm's relative attention to macroeconomic news. For example, higher idiosyncratic risk or lower stock returns could suppress non-macroeconomic reading if the employees of these firms refrain from reading news about the firm's main line of business (e.g., potential investment opportunities). This could cause $\Delta\sigma_{i,t-1}$ to predict an increase in TS-RA$_{i,t}$ due to higher idiosyncratic volatility rather than aggregate economic uncertainty.

We address this concern by implementing the instrumental variables approach proposed by Alfaro et al. (2024) to show that employee attention to macroeconomic news is sensitive to fluctuations in economic uncertainty. While Section OA.1 of the Online Appendix provides full details on the estimation procedure, the crux of this approach involves three steps. First, we use an empirical asset-pricing model, such as the Carhart (1997) model, to remove the common component of stock returns. Next, we examine how each industry's idiosyncratic returns covary with aggregate variables, such as currency returns. Finally, we instrument $\Delta\sigma_{i,t-1}$ using the absolute values of the industry-level exposures from the second step. The identifying assumption is that firm-level unobservables are uncorrelated with non-directional industry-level sensitivities to aggregate economic conditions. Thus, 2SLS yields consistent estimates of the parameters in equation (10).

Columns (4) through (6) of Table 7 present the result and show that plausibly exogenous fluctuations in economic uncertainty continue to lead to increased employee attention to macroeconomic news. These effects are both economically large, in that a one-standard-deviation increase in uncertainty induces employees to allocate between 1.2% and 2.3% more attention to macroeconomic news, and statistically significant. This is in spite of the fact that we control for a battery of date and firm fixed effects in Column (6) as well as a battery of firm-level controls. Moreover, the association between excess returns and readings remains negative and is marginally statistically significant. The results consistently show that employees reallocate their attention to macroeconomic news, especially amid increased economic uncertainty.

Relation to alternative measures of firm risk. As CS-RA$_{i,t-1}$ is more closely aligned to second- than first-moment dynamics, we examine whether cross-sectional variation in CS-RA$_{i,t-1}$ provides any incremental insights into firm risk beyond other proxies in the literature. Notably, Table 8 reconsiders the key regressions from Section 4.3 after controlling for two common proxies of firm-level uncertainty. In Panel A we control for the political and non-political uncertainty measures from Hassan et al. (2019). Critically, these measures are derived from earnings call transcripts and thus reflect not only management perceptions of these risks, but also the degree to which they chose to discuss them with investors. This leaves open the possibility that rank-and-file employee attention to macroeconomic risk, as reflected by CS-RA$_{i,t-1}$, will continue to retain explanatory power for firm outcomes. In panel B we control for firm-level implied volatility, a common measure of risk from the literature, derived from options prices as in Dew-Becker and Giglio (2023).

We add these measures of risk to our two most stringent empirical specifications —i.e., the specification of equation (9) that includes previously highlighted controls and/or fixed effects—and examine the marginal effect of an increase in CS-RA$_{i,t}$ on real outcomes. Supposing CS-RA$_{i,t}$ is simply a rotation of these alternative measures, their inclusion in the regression should subsume the information content of employees' relative attention to macroeconomic news.

Columns (1) and (2) of Table 8 feature asset growth, columns (3) and (4) sales growth, and columns (5) and (6) employment growth as dependent variables. The economic significance of CS-RA$_{i,t}$ remains largely intact for most specifications. The one concerning specification is asset growth where there is substantial attenuation in the coefficient on CS-RA$_{i,t-1}$ going from columns (1) to (2). This attenuation, however, has more to do with the inclusion of Tobin's Q—a common proxy for firm investment potential—as a control rather than the inclusion of either the Hassan et al. (2019) or the implied volatility uncertainty measures. Also of note is that firm-level implied volatility often has the opposite sign than what one would predict if it were a strong proxy of firm exposure to aggregate uncertainty. These regressions suggest that CS-RA$_{i,t}$ contains information about firm-level risk beyond these alternative measures of firm risk exposure.[14]

[14] An additional concern is that our findings are driven entirely by the COVID period. Table OA.5.6 of the Online Appendix runs the same set of regressions after removing the January through December 2020 period from the panel. Our baseline results remain largely unchanged after removing this time period.

5.2 Relation to firm-level characteristics

This section explores the asset-pricing implications of firms' relative attention to macroeconomic conditions. Given our short sample period, we focus on the relation between CS-RA$_{i,t}$ and several firm-level characteristics that the prior literature links to risk exposure: market beta, size, book-to-market ratios, profitability, and investment. These characteristics underlie the Hou, Xue, and Zhang (2014) and Fama and French (2015) models. We find that while firms that pay relatively more or less attention to macroeconomic-related topics feature economically and statistically significant differences in these characteristics, substantial unexplained variation in the drivers of CS-RA$_{i,t}$ remains. This indicates that CS-RA$_{i,t}$ varies for reasons beyond these common characteristics.

We show the relation between CS-RA$_{i,t}$ and these firm-level characteristics by sorting firms into either three or five CS-RA$_{i,t}$-ranked portfolios at the end of each calendar quarter. We then calculate the value-weighted average characteristic of each portfolio. Consistent with information revelation on accounting release dates, we update each firm's CS-RA$_{i,t}$ at the end of each announcement month. Table 9 reports the results of this analysis. Across both the tercile and quintile sorts in Panels A and B, respectively, we find that high CS-RA$_{i,t}$ firms have higher book-to-market ratios, higher profitability, and lower investment. These differences are significant at the 5% level. High CS-RA$_{i,t}$ firms also have higher market betas than low CS-RA$_{i,t}$ firms, although these differences are insignificant. While each of these characteristics is in line with the notion that high CS-RA$_{i,t}$ firms are riskier, the outlying characteristic is the difference in market capitalization, which is positive even though the size effect predicts a negative relation with returns.[15] Collectively, the strong relation between CS-RA$_{i,t}$ and these characteristics bolsters our confidence that CS-RA$_{i,t}$ is related to risk exposures.

The analysis underlying Table 9 focuses on the univariate relation between characteristics and CS-RA$_{i,t}$. This makes it difficult to determine the marginal contribution of each characteristic in explaining variation in a firm's relative attention to macroeconomic conditions. Thus, Table OA.3.4 in the Online Appendix performs a simple variance decomposition of CS-RA$_{i,t}$. While we provide the full details in the Online Appendix, the key takeaway from this analysis is that firm-level characteristics explain relatively little variation in CS-RA$_{i,t}$. Without controlling for unconditional

[15]While the CS-RA$_{i,t}$ measure does not line up correctly with the size characteristic, the size factor has had a close to zero return over the past two decades.

differences in attention between sectors (defined using two- or three-digit NAICS codes) or times, variation in characteristics explains about 12% of the variation in CS-RA$_{i,t}$. However, when we control for sector-by-date fixed effects, these characteristics explain less than 0.5% of the variation in relative attention. In fact, over 90% of the variation in CS-RA$_{i,t}$ is explained by firm-specific factors. This highlights the fact that knowledge of CS-RA$_{i,t}$ provides insights into firm-level risk exposures that are not reflected by traditional return- and accounting-based characteristics.

Finally, Table OA.5.5 in the Online Appendix builds on the evidence that CS-RA$_{i,t}$ is not simply a linear transformation of common asset-pricing characteristics by showing that high CS-RA$_{i,t}$ firms tend to earn larger returns, on average, than low CS-RA$_{i,t}$ firms. The unconditional difference in returns is 14 basis points per week, which is equivalent to about 7.5% per annum. While this result is not statistically significant (t-statistic of 1.36), we note that (i) this result is consistent with the positive relation between CS-RA$_{i,t}$ and implied costs of capital in Table 5, and (ii) noise in the time series of realized returns over our short sample period (i.e., 2016 through 2022) reduces our ability to establish statistically significant differences in returns. Columns (2) and (3) then show that this return is not explained by either the CAPM or the Fama and French (1993) three-factor model, but Columns (4) and (5) show that this spread becomes economically and statistically diminished when we consider the Fama and French (2015) five-factor model plus momentum.

6 Conclusion

This paper uses a high-dimensional dataset that reflects the daily internet news reading of the employees of public firms to characterize firm-level exposures to macroeconomic risk. Notably, and in the time series, we show that employees' attention to macroeconomic risk increases in periods following a deterioration in macroeconomic conditions, measured using either a decline in the level of macroeconomic activity or an increase in economic uncertainty. Second, in the cross-section, we show that firms that spend relatively more of their time reading about macroeconomic conditions are (i) more likely to engage in corporate risk mitigation, (ii) have higher costs of equity capital, and (iii) tend to invest and produce less. These relations between employee reading, firm risk, and firm outcomes extend beyond typical firm characteristics, such as size, leverage, and asset tangibility, and often hold after accounting for differences in risk across industries and time periods.

The fact that employee attention to macroeconomic news is highly informative about firm-level risk exposures and outcomes is intuitive ex-post but not entirely obvious ex-ante. After all, if the average employee spends most of their time reading news on the internet that is unrelated to the risks or prospects of their employer, then a firm's relative attention to macroeconomic conditions will contain no useful information about the economic activity of the firm. In contrast, our results—which consistently find that employee readings respond to aggregate economic conditions and predict firm-specific actions—highlight that the attention of rank-and-file employees is useful for understanding the corporate policies that only a few key corporate personnel may ultimately be responsible for making. In some sense, this could reflect the fact that observing the news consumption of individual employees allows us to elicit the "wisdom of crowds" within each firm.

Our results, which directly link the business-relevant attention of employees to firm risk, stand in contrast to many studies that rely on surveys to infer the key concerns of firms and corporate personnel (e.g., Coibion, Gorodnichenko, and Ropele (2020)). Moreover, we contribute to a nascent literature that examines the kinds of news that individuals choose to consume and why. While this paper focuses on how a firm's attention to macroeconomic news aligns with the firm's exposure to macroeconomic risk, asking whether employees pay attention to other sources of risk, such as climate and regulatory risk, could provide interesting avenues for future research that could use these other kinds of news consumption to infer exposures to non-traded risks.

References

Acharya, V. V., Lochstoer, L. A., Ramadorai, T., 2013. Limits to arbitrage and hedging: Evidence from commodity markets. Journal of Financial Economics 109, 441–465.

Ai, H., Bansal, R., 2018. Risk preferences and the macroeconomic announcement premium. Econometrica 86, 1383–1430.

Alfaro, I., Bloom, N., Lin, X., 2024. The finance uncertainty multiplier. Journal of Political Economy 132, 000–000.

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

Bali, T. G., Brown, S. J., Tang, Y., 2017. Is economic uncertainty priced in the cross-section of stock returns? Journal of Financial Economics 126, 471–489.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The Journal of Finance 59, 1481–1509.

Bekaert, G., Engstrom, E. C., Xu, N. R., 2022. The time variation in risk appetite and uncertainty. Management Science 68, 3975–4004.

Ben-Rephael, A., Carlin, B. I., Da, Z., Israelsen, R. D., 2021. Uncovering the hidden effort problem. Tech. rep., National Bureau of Economic Research.

Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77, 623–685.

Breeden, D. T., Litzenberger, R. H., 1978. Prices of state-contingent claims implicit in option prices. Journal of Business pp. 621–651.

Brown, G. W., 2001. Managing foreign exchange risk with derivatives. Journal of Financial Economics 60, 401–448.

Bybee, L., Kelly, B., Manela, A., Xiu, D., 2023a. Business news and business cycles. Journal of Finance .

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

Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

Campello, M., Lin, C., Ma, Y., Zou, H., 2011. The real and financial implications of corporate hedging. The Journal of Finance 66, 1615–1647.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Chinco, A., Hartzmark, S. M., Sussman, A. B., 2022. A new test of risk factor relevance. The Journal of Finance 77, 2183–2238.

Coibion, O., Gorodnichenko, Y., Ropele, T., 2020. Inflation expectations and firm decisions: New causal evidence. The Quarterly Journal of Economics 135, 165–219.

Colacito, R., Croce, M. M., Liu, Y., Shaliastovich, I., 2022. Volatility risk pass-through. The Review of Financial Studies 35, 2345–2385.

Cookson, J. A., Engelberg, J. E., Mullins, W., 2023. Echo chambers. The Review of Financial Studies 36, 450–500.

Cookson, J. A., García, D., Jarnecic, E., 2024. News consumption in the wild .

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 .

Dew-Becker, I., Giglio, S., 2023. Cross-sectional uncertainty and the business cycle: Evidence from 40 years of options data. American Economic Journal: Macroeconomics 15, 65–96.

Dougal, C., Engelberg, J., Garcia, D., Parsons, C. A., 2012. Journalists and the stock market. The Review of Financial Studies 25, 639–679.

Driscoll, J. C., Kraay, A. C., 1998. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. The Review of Economics and Statistics 80, 549–560.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., Stroebel, J., 2020. Hedging climate change news. The Review of Financial Studies 33, 1184–1216.

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

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

Fisher, A., Martineau, C., Sheng, J., 2022. Macroeconomic attention and announcement risk premia. The Review of Financial Studies 35, 5057–5093.

Gabaix, X., 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macrofinance. The Quarterly Journal of Economics 127, 645–700.

Gabaix, X., 2014. A sparsity-based model of bounded rationality. The Quarterly Journal of Economics 129, 1661–1710.

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

Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as data. Journal of Economic Literature 57, 535–74.

Goldstein, I., Spatt, C. S., Ye, M., 2021. Big data in finance. The Review of Financial Studies 34, 3213–3225.

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

He, Z., Kelly, B., Manela, A., 2017. Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics 126, 1–35.

Hong, H., Wang, N., Yang, J., 2023. Mitigating disaster risks in the age of climate change. Econometrica 91, 1763–1802.

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

Hou, K., Xue, C., Zhang, L., 2014. Digesting anomalies: An investment approach. The Review of Financial Studies 28, 650–705.

Israelsen, R. D., 2014. Tell it like it is: Disclosed risks and factor portfolios. Available at SSRN 2504522.

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L., 2016. A rational theory of mutual funds' attention allocation. Econometrica 84, 571–626.

Kalmenovitz, J., 2022. Regulatory intensity and firm-specific exposure. The Review of Financial Studies .

Leary, M. T., Roberts, M. R., 2014. Do peer firms affect corporate financial policy? The Journal of Finance 69, 139–178.

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

Ludvigson, S. C., Ma, S., Ng, S., 2021. Uncertainty and business cycles: Exogenous impulse or endogenous response? American Economic Journal: Macroeconomics 13, 369–410.

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

Martineau, C., Mondria, J., 2023. News selection and asset pricing implications. Available at SSRN 4194851.

Mazumder, R., Pruitt, S., Ross, L., 2023. Risk exposures from risk disclosures: What they said and how they said it. Available at SSRN 4637697.

Mullainathan, S., Shleifer, A., 2005. The market for news. American Economic Review 95, 1031–1053.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation. Econometrica 55, 703–708.

Pástor, L., Sinha, M., Swaminathan, B., 2008. Estimating the intertemporal risk–return tradeoff using the implied cost of capital. The Journal of Finance 63, 2859–2897.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics 80, 563–602.

Stewart, C. A., Welch, V., Plale, B., Fox, G., Pierce, M., Sterling, T., 2017. Indiana university pervasive technology institute .

Van Nieuwerburgh, S., Veldkamp, L., 2010. Information acquisition and under-diversification. The Review of Economic Studies 77, 779–805.

Weber, M., D'Acunto, F., Gorodnichenko, Y., Coibion, O., 2022. The subjective inflation expectations of households and firms: Measurement, determinants, and implications. Journal of Economic Perspectives 36, 157–184.

Whited, T. M., Wu, G., 2006. Financial constraints risk. The Review of Financial Studies 19, 531–559.

Zhang, L., 2005. The value premium. The Journal of Finance 60, 67–103.

Tables and Figures

Table 1: Summary Statistics
This table presents summary statistics for the number of firms (*domains*) covered by the data, the number of topics in the Consortium's taxonomy, and the cross-sectional distribution of firm-topic interactions for each year in our sample. Panel A reports results for all firms, while Panel B focuses on the CRSP-COMPUSTAT universe of firms. The cross-sectional distribution of firm-topic interactions is characterized by the mean, median, standard deviation (*SD*), and skewness, estimated using data from the last week of each year. The sample period is from June 2016 through July 2022.

(a) All firms

			Cross-sectional Firm-Topic			
Year	Firms	Topics	Mean	Median	SD	Skew
2016	651599	2407	204.27	111.00	274.88	3.02
2017	1935696	3227	477.77	256.00	582.79	2.02
2018	2051067	4804	444.62	176.00	687.69	2.95
2019	1713267	5500	262.84	76.00	567.97	4.83
2020	1960832	5869	306.06	105.00	587.48	4.49
2021	2053800	7392	301.77	86.00	642.63	5.01
2022	2124557	7395	295.76	83.00	630.07	4.98

(b) Public firms

			Cross-sectional Firm-Topic			
Year	Firms	Topics	Mean	Median	SD	Skew
2016	2609	2407	992.04	957.00	653.37	0.17
2017	2749	3227	2406.45	2864.00	964.07	-1.13
2018	2755	4804	3229.05	3820.00	1578.06	-0.71
2019	2786	5500	2933.25	3055.50	1962.48	-0.12
2020	2877	5869	3110.18	3185.00	2026.29	-0.10
2021	3181	7392	3426.36	3212.00	2505.42	0.13
2022	3267	7395	3410.00	3204.00	2465.62	0.15

Table 2: **Summary Statistics and Characteristics by Industry**

This table reports average industry characteristics across 17 North American Industry Classification (NAICS) two-digit industries, providing insights into how firms in each industry interact with the Consortium's data. For each industry, we report summary statistics for all public firms (*Public*) and the fraction of public firms that we match to the Consortium's data (*Match (%)*). The statistics include the number of firms in each industry (*Firms*), the market capitalization of each industry (*Market Cap*), and the number of employees in each industry (*Employees*). Additionally, we report the fraction of topics in the Consortium's data that firms in each industry interact with (*Topics (%)*) and the average value of our cross-sectional relative attention measure (*Attention (%)*), as defined in Section 3.2. The topic- and attention-related statistics are first calculated at the firm level across all years and then averaged across all firms in each industry. The definition of each NAICS is available at the U.S. Census website, `https://www.census.gov/naics/`.

	Firms		Market Cap		Employees		Topics	Attention
Industry	Public	Match(%)	Public	Match(%)	Public	Match(%)	(%)	(%)
Technical Services	116	86.44	4995060	96.40	21420	91.33	46.80	45.52
Information	423	77.75	14884940	67.70	14270	57.24	47.89	48.46
Education	27	56.44	1575222	64.95	8756	63.96	15.40	51.22
Agriculture	10	86.21	3332093	98.65	6524	98.39	42.84	51.78
Arts & Entertainment	24	87.59	2706848	98.07	9351	99.07	28.07	53.86
Wholesale	102	85.50	3729525	84.83	9669	88.11	35.80	56.65
Real Estate	259	81.95	5575583	94.57	2834	86.80	23.32	57.83
Manufacturing	1653	89.72	8080339	93.73	10247	90.75	44.58	58.37
Health Care	53	94.62	4256509	99.02	23369	98.49	45.49	58.71
Construction	54	91.93	2664172	98.26	7192	93.25	30.34	60.06
Waste Management	73	90.83	4996922	86.76	25932	95.56	38.04	61.26
Hotel and Food	66	87.66	7187891	92.72	40124	89.11	39.58	61.50
Utilities	91	85.32	11459666	92.70	7282	83.94	30.83	63.90
Transportation	128	75.23	8141457	91.44	18689	88.60	48.76	66.22
Oil & Gas	219	86.78	4245510	93.25	4305	89.15	28.69	69.27
Retail	132	86.87	17438184	84.17	62614	69.16	68.55	72.86
Finance	735	83.68	7194094	91.44	8132	91.70	48.56	74.70
All	4198	85.62	8212605	87.87	12605	83.66	44.98	61.30

Table 3: **Sensitivity of Relative Attention to Fluctuations in Macroeconomic Conditions**
The table reports how the time-series measure of relative attention to the macroeconomy from equation (1) changes in response to fluctuations in either the level of macroeconomic activity (Panel A) or economic uncertainty (Panel B). Columns (1), (3), and (5) of each panel display the results of estimating equation (3), where the dependent variable is each firm's value of relative attention from equation (1) and the independent variables include one of the daily corporate default spread, the intermediary capital ratio from He et al. (2017), or the return on WTI oil in Panel A, and the daily value of the VIX index, EPU index of Baker et al. (2016), or macroeconomic uncertainty index from Bekaert et al. (2022) in Panel B. These regressions also include day-of-the-week fixed effects to subsume any fixed differences in attention across days of the week. Columns (2), (4), and (6) of these panels report the results of estimating equation (7), whereby we interact each macroeconomic variable with each individual firm's exposure to macroeconomic conditions, as defined by equation (6). The regressions in these columns also feature date-by-industry fixed effects. The sample period runs from May 2016 through June 2022. Standard errors are clustered by firm and date.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First Moment Proxies						
	Def. Spread		Int. Capital		WTI Returns	
MacroV_{t-1}	0.0147***		−0.0176***		−0.0130***	
	[14.38]		[−16.66]		[−15.93]	
$\text{CS-RA}_{i,t-1}$		0.0649***		0.0661***		0.0648***
		[38.54]		[41.49]		[37.82]
$\text{CS-RA}_{i,t-1} \times \text{MacroV}_{t-1}$		0.0157***		−0.0278***		0.0087***
		[9.12]		[−17.46]		[5.99]
Observations	3,660,323	3,435,177	3,660,323	3,435,177	3,660,323	3,435,177
R^2	0.0141	0.1903	0.0197	0.1928	0.0116	0.1896
Panel B: Second Moment Proxies						
	VIX Index		EPU Index		BEX Index	
MacroV_{t-1}	0.0065***		0.0128***		0.0217***	
	[5.61]		[11.33]		[17.09]	
$\text{CS-RA}_{i,t-1}$		0.0651***		0.0651***		0.0649***
		[38.84]		[38.90]		[37.94]
$\text{CS-RA}_{i,t-1} \times \text{MacroV}_{t-1}$		0.0175***		0.0187***		0.0103***
		[9.79]		[9.04]		[7.04]
Observations	3,660,323	3,435,177	3,660,323	3,435,177	3,660,323	3,435,177
R^2	0.0043	0.1906	0.0111	0.1908	0.0286	0.1897
Day of Week FE	+		+		+	
Date × Industry FE		+		+		+

Table 4: Risk Mitigation, Regulatory Compliance, and Firm Attention to Macroeconomic Risk
The table presents regression results that examine the relation between firm-level exposure to macroeconomic risk and either hedging intensity (Panel A) or regulatory intensity (Panel B). These results are obtained by estimating equation (9). In Panel A, the dependent variable measures firm-level hedging intensity that we define using the approach of Campello et al. (2011). Specifically, we first count the number of hedging-related words that each firm mentions in its annual 10-K. We then scale this quantity by the total number of words in the firm's 10-K. To eliminate the influence of outliers, this hedging-intensity ratio is transformed to reflect whether each firm is in the top quartile of the hedging-intensity score in a given year. In Panel B, we measure a firm's regulatory intensity using the data constructed by Kalmenovitz (2022). Regressions include combinations of industry, date, and industry-by-date fixed effects, as well as a battery of control variables. These control variables include the set of variables from Leary and Roberts (2014), Whited and Wu (2006) proxy for financial constraints, and the CAPM β, a common proxy for a firm's exposure to macroeconomic risk. Panel A (Panel B) uses data from June 2016 through July 2022 (December 2020). All standard errors are clustered by firm.

	(1)	(2)	(3)	(4)	(5)
			Panel A: Hedging Activity		
CS-RA$_{i,t-1}$	0.1152***	0.0579***	0.0576***	0.0573***	0.0001
	[6.00]	[3.25]	[3.10]	[3.10]	[0.01]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				0.0479***	0.0474***
				[5.01]	[4.95]
Observations	11,188	11,148	11,075	11,075	10,971
R^2	0.0041	0.2245	0.2496	0.2511	0.2793
			Panel B: Compliance Activity		
CS-RA$_{i,t-1}$	19.5738***	13.6059***	14.1330***	14.6192***	8.8025***
	[10.02]	[9.46]	[9.52]	[9.30]	[6.88]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				1.6933***	0.7610
				[2.70]	[1.27]
Observations	24,483	24,455	24,048	24,048	23,219
R^2	0.0242	0.4261	0.5429	0.5431	0.5869
Date FE		+			
Industry FE		+			
Date × Industry FE			+	+	+
CAPM-beta				+	+
Controls					+

Table 5: Implied Cost of Capital and Firm Attention to Macroeconomic Risk

This table shows the results of estimating equation (11) to examine the relation between a firm's implied cost of capital, measured using the approach of Gebhardt et al. (2001), and the firm's relative attention to macroeconomic risk (CS-RA$_{i,t}$), measured via equation (6). Regressions include combinations of industry, date, and industry-by-date fixed effects. All specifications also control for the CAPM β of each firm. Additionally, specification (5) also includes a set of asset-pricing characteristics related to each underlying firm. These characteristics include firm size, Tobins Q (the inverse of the book-to-market ratio), profitability, and investment tangibility. The data underlying this regression span from 2016 through 2022, and all standard errors are clustered by firm.

	(1)	(2)	(3)	(4)	(5)
CS-RA$_{i,t-1}$	0.0166***	0.0167***	0.0109***	0.0117***	0.0060***
	[8.44]	[8.40]	[6.25]	[6.50]	[3.66]
CAPM $\beta_{i,t-1}$	−0.0010	−0.0009	−0.0023***	−0.0024***	−0.0025***
	[−1.23]	[−1.13]	[−2.70]	[−2.70]	[−2.62]
Size$_{i,t-1}$					0.0037***
					[3.94]
ROA$_{i,t-1}$					0.0084***
					[5.61]
Tangability$_{i,t-1}$					0.0019
					[1.50]
Tobin Q$_{i,t-1}$					−0.0071***
					[−6.12]
Date FE		+	+		
Industry FE			+		
Date × Industry FE				+	+
Observations	37,309	37,309	37,285	36,905	35,527
R^2	0.0087	0.0146	0.1033	0.1312	0.1608

Table 6: Real Outcomes and Firm Attention to Macroeconomic Risk

This table reports the results of estimating equation (9) to examine the relation between a firm's relative attention to macroeconomic risk (measured via equation (6)) and total asset growth (Panel A), sales growth (Panel B), and employment growth (Panel C). Asset growth, sales growth, and employment growth are defined in the Internet Appendix OA.1. The specifications include combinations of industry, date, and industry-by-date fixed effects. Additionally, specifications (4) and (5) control for each firm's CAPM β, a common proxy for a firm's exposure to macroeconomic risk. Specification (5) features both the set of controls from Leary and Roberts (2014) and Whited and Wu (2006) measure of financial constraints. The data underlying this regression span from 2016 through 2022, and all standard errors are clustered by firm.

	(1)	(2)	(3)	(4)	(5)
			Panel A: Asset Growth		
CS-RA$_{i,t-1}$	-0.0406^{***}	-0.0326^{***}	-0.0332^{***}	-0.0308^{***}	-0.0099^{**}
	[−11.78]	[−9.18]	[−9.20]	[−8.61]	[−2.47]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				-0.0188^{***}	-0.0152^{***}
				[−5.70]	[−4.42]
Observations	53,846	53,811	53,329	53,329	41,217
R^2	0.0031	0.0193	0.0513	0.0520	0.1037
			Panel B: Sales Growth		
CS-RA$_{i,t-1}$	-0.1362^{***}	-0.1106^{***}	-0.1134^{***}	-0.1164^{***}	-0.0472^{***}
	[−12.84]	[−10.81]	[−10.82]	[−11.02]	[−4.38]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				-0.0260^{***}	−0.0062
				[−3.05]	[−0.71]
Observations	49,048	49,013	48,557	48,557	40,747
R^2	0.0050	0.0462	0.0869	0.0872	0.1206
			Panel C: Employment Growth		
CS-RA$_{i,t-1}$	-0.0569^{***}	-0.0345^{***}	-0.0327^{***}	-0.0340^{***}	-0.0263^{***}
	[−7.32]	[−4.35]	[−3.98]	[−4.14]	[−3.02]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				0.0020	0.0005
				[0.44]	[0.10]
Observations	10,555	10,516	10,449	10,449	8,615
R^2	0.0061	0.0614	0.0880	0.0894	0.1637
Date FE		+			
Industry FE		+			
Date × Industry FE			+	+	+
CAPM-beta				+	+
Controls					+

Table 7: Firm Relative Attention and Uncertainty Shocks

This table shows the results of estimating equation (10), which examines the relationship between changes in risk and attention to macroeconomic-related news (TS-RA$_{i,t}$). Regressions include a combination of firm and date fixed effects as well as the firm-level controls of Leary and Roberts (2014), Whited and Wu (2006), and CAPM β. The proxy for firm risk in columns (1) through (3) is the firm-level return volatility over a 250-day rolling period. In columns (4) through (6) we follow the approach of Alfaro et al. (2024), instrumenting firm-level return volatility with industry-level estimates of sensitivity with growth of various macroeconomic measures in an 2 stage least square analysis (see Online Appendix OA.1 for details). We t-statistics are estimated using the the HAC-robust approach of Driscoll and Kraay (1998) with a lag length of four quarters. The sample period runs from May of 2016 through June of 2022.

	Realized Volatility			Instrumented Volatility		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔVolatility$_{i,t-1}$	0.0071***	0.0020	0.0023*	0.0234***	0.0146***	0.0116***
	[3.05]	[1.67]	[1.77]	[7.95]	[3.34]	[3.33]
Ex Returns$_{i,t-1}$	−0.0051**	−0.0011	−0.0012	−0.0028	−0.0018	−0.0020*
	[−2.23]	[−1.43]	[−1.24]	[−1.10]	[−1.69]	[−1.77]
Lag Reading	+	+	+	+	+	+
Date FE		+	+		+	+
Firm FE		+	+		+	+
Firm Controls			+			+
Observations	52,968	52,735	41,031	52,966	52,733	41,030

Table 8: Firm Relative Attention Versus Measures of Uncertainty

This table reports the results of estimating equation (9) to examine the relation between a firm's relative attention to macroeconomic risk (measured via equation (6)), its exposure to political and non-political uncertainty (see Hassan et al., 2019) and past implied volatility (see Dew-Becker and Giglio, 2023). The dependent variables are total asset growth (Columns 1 and 2), sales growth (Column 3 and 4), and employment growth (Columns 5 and 6). All variables are defined in the Internet Appendix OA.1. CS-RA$_{i,t}$ is mapped to be between 0 and 1 while the other independent variables are standardized. As controls, specifications include industry-by-date fixed effects, a measure of each firm's CAPM β—a common proxy for a firm's exposure to macroeconomic risk—and both the set of controls from Leary and Roberts (2014) and the Whited and Wu (2006) measure of financial constraints. The data underlying this regression span from 2016 through 2022, and all standard errors are clustered by firm.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Hassan, Hollander, Van Lent, and Tahoun (2019)						
	Asset Growth		Sales Growth		Employment Growth	
CS-RA$_{i,t-1}$	-0.0264^{***}	-0.0063	-0.0884^{***}	-0.0328^{***}	-0.0437^{***}	-0.0244^{**}
	[−6.86]	[−1.50]	[−8.40]	[−3.17]	[−4.65]	[−2.53]
PRisk$_{i,t-1}$	-0.0011	-0.0016^{*}	-0.0034	-0.0058^{**}	-0.0032	-0.0038^{*}
	[−1.11]	[−1.69]	[−1.37]	[−2.12]	[−1.48]	[−1.75]
NPRisk$_{i,t-1}$	-0.0012	0.0002	-0.0021	-0.0009	-0.0071^{***}	-0.0036^{*}
	[−1.38]	[0.21]	[−0.98]	[−0.42]	[−3.45]	[−1.73]
Observations	41,250	34,203	39,946	33,988	8,192	7,100
R^2	0.0616	0.1090	0.1041	0.1395	0.0983	0.1812
Panel B: Dew-Becker and Giglio (2023)						
	Asset Growth		Sales Growth		Employment Growth	
CS-RA$_{i,t-1}$	-0.0207^{***}	-0.0015	-0.0734^{***}	-0.0334^{***}	-0.0494^{***}	-0.0290^{***}
	[−5.46]	[−0.35]	[−6.25]	[−2.94]	[−5.36]	[−3.08]
Implied Vol$_{i,t-1}$	0.0063^{***}	0.0079^{***}	0.0459^{***}	0.0180^{***}	0.0052	0.0077
	[3.90]	[3.07]	[9.70]	[2.67]	[1.56]	[1.49]
Observations	40,083	32,832	38,065	32,579	8,086	6,932
R^2	0.0597	0.1145	0.1037	0.1357	0.1097	0.1915
Date × Industry FE	+	+	+	+	+	+
CAPM-beta	+	+	+	+	+	+
Controls		+		+		+

Table 9: Firm Attention to Macroeconomic Conditions: Portfolio Characteristics
The table presents the characteristics of portfolios formed by sorting firms on the basis of their relative attention to macroeconomic conditions, (CS-RA$_{i,t}$), measured via equation (6). Panel A (Panel B) reports the resulting of sorting the cross section of firms into tercile (quintile) portfolios at the end of each quarter from 2016 through 2022. For each portfolio, we report the value-weighted average values of five prominent asset-pricing characteristics: CAPM β, market capitalization, the book-to-market ratio, gross profitability, and asset growth. We also report the average difference in characteristics between the high CS-RA$_{i,t}$ portfolio and the low CS-RA$_{i,t}$ portfolio. The underlying data span from 2016 through 2022, and brackets report t-statistics that are constructed using Newey and West (1987) standard errors.

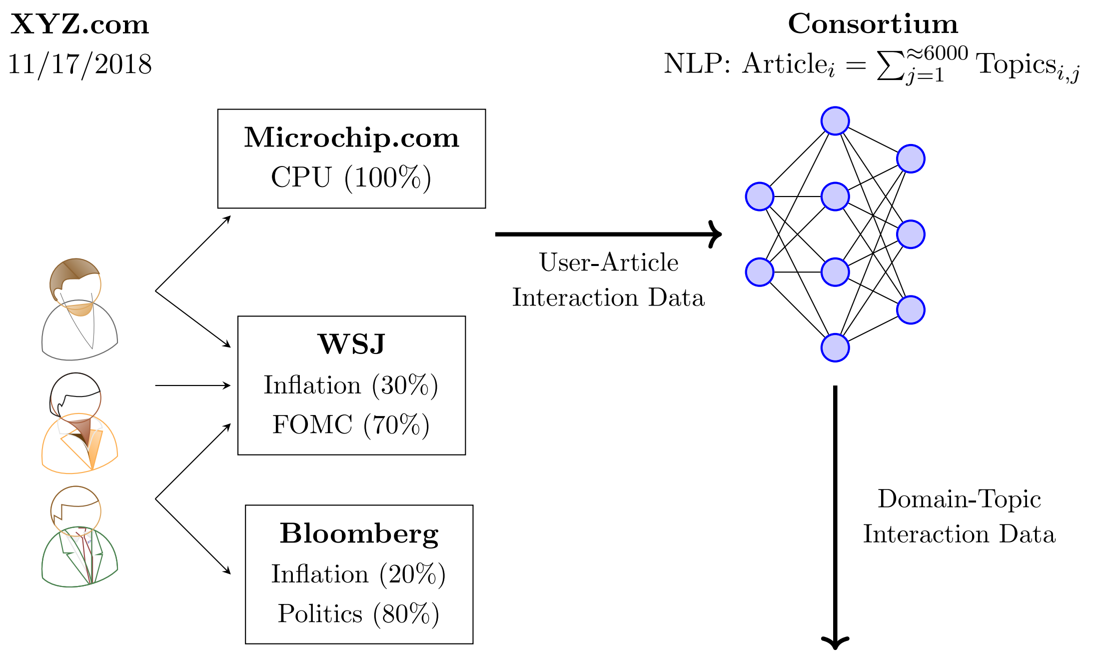
(a) Three CS-RA Portfolios

	Beta	Market Cap	Book to Market	Gross Profit	Asset Growth
Low CS-RA	0.9139	5.7593	0.9451	0.1380	0.3316
Medium	1.0184	6.7104	0.6591	0.2426	0.1969
High CS-RA	1.0428	7.5378	0.7058	0.2637	0.1236
High-Low	0.1289	1.7785	-0.2393	0.1257	-0.2080
t-stat	[8.00]	[9.12]	[-4.28]	[6.46]	[-3.67]

(b) Five CS-RA Portfolios

	Beta	Market Cap	Book to Market	Gross Profit	Asset Growth
Low CS-RA	0.8999	5.6863	1.0392	0.1277	0.3510
2	0.9520	6.0229	0.7629	0.1721	0.2821
Medium	1.0209	6.7290	0.6510	0.2447	0.1943
4	1.0461	7.2068	0.6710	0.2698	0.1437
High CS-RA	1.0398	7.7002	0.7265	0.2593	0.1162
High-Low	0.1400	2.0139	-0.3128	0.1316	-0.2347
t-stat	[8.52]	[9.67]	[-4.31]	[8.13]	[-3.88]

Figure 1: Visualizing the Consortium's Member Base
This figure illustrates a subsample of the Consortium's more than 4,000 members.



Date	Domain	Score	Topic
11/17/2018	xyz.com	3	Inflation
11/17/2018	xyz.com	3	FOMC
11/17/2018	xyz.com	1	CPU
11/17/2018	xyz.com	1	Politics

Figure 2: Visual Representation of the Consortium's data
The figure presents a stylized example of the Consortium's data generation process. Initially, users read various articles from a range of online publishers. The online publishers then provide this user-interaction data to the Consortium. Utilizing machine learning and natural language processing algorithms, the Consortium decomposes each article into a combination of its core topics (illustrated beneath each article). Subsequently, the Consortium aggregates the analyzed user-article interaction data across users and firms (depicted as domains) to generate domain-topic interaction data. This data encompass several variables, such as the date of interactions, the domain of interactions, the user engagement with a specific topic, and an associated topic label.

Figure 3: Category Visualization

This figure presents a word cloud illustrating the prevalence of topic categories across the Consortium's dataset. Each topic in the dataset is associated with a specific category. Category labels are then weighted by the number of associated topics, making categories with more topics appear larger.

(a) Raw topic cloud

(b) *tf-iaf*-weighted topic cloud

Figure 4: Raw and *tf-iaf*-Weighted Topic Clouds for Macroeconomic Topics
This figure presents two topic clouds for the same set of macroeconomic-related topics. Panel A displays the raw topic cloud, showing the most read topics in aggregate. Panel B presents the *tf-iaf*-weighted topic cloud, highlighting the topics that are more informative in distinguishing reading about macroeconomic conditions across firms in the cross-section. We define the corpus of macroeconomic-related articles following the approach outlined in Section 3.2. Likewise, the *tf-iaf* weights are given by equation (5). The data underlying these topic clouds are from the week ending November 17, 2018.

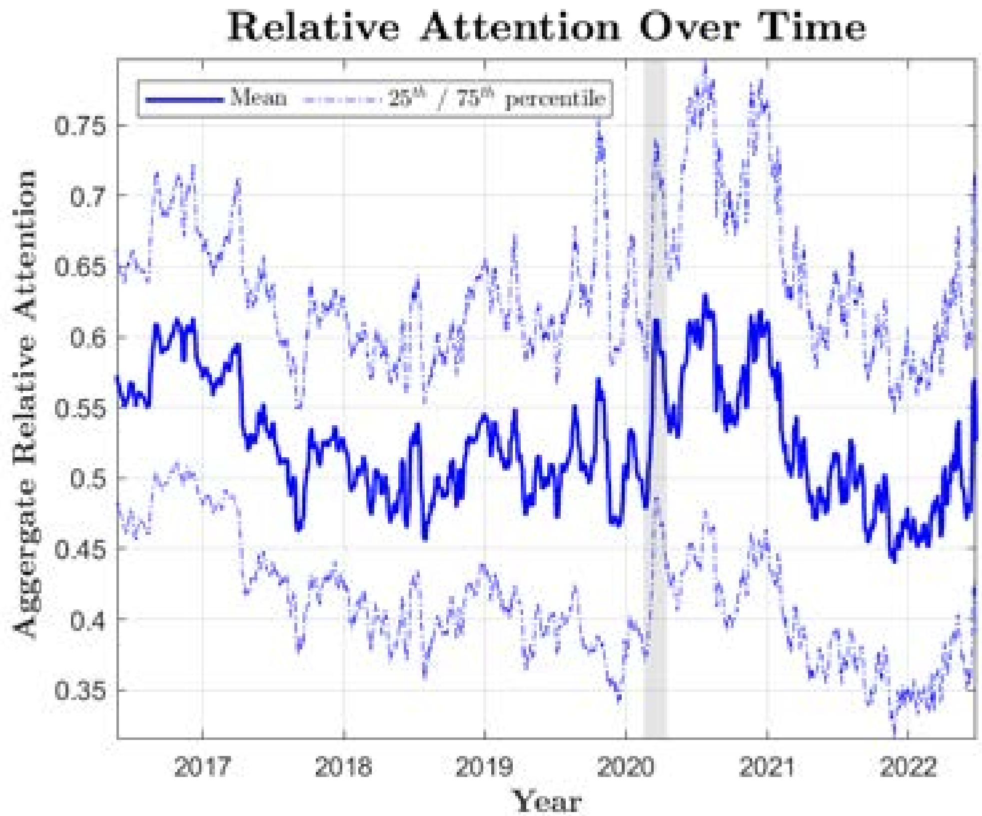


Figure 5: Aggregate Relative Attention ($\widetilde{\textbf{TS-RA}}_t$) and Macroeconomic Conditions
This figure displays the time series of raw relative attention to macroeconomic conditions, as defined in equation (2). The solid blue line displays the average value of TS-RA$_{i,t}$ across all firms (i.e., the time series of $\widetilde{\text{TS-RA}}_t$ from equation (1)), while the dashed blue lines report the average value of TS-RA$_{i,t}$ across all firms with a value of TS-RA$_{i,t}$ that is greater (less) than the 75^{th} (25^{th}) percentile of TS-RA$_{i,t}$ at each time t. The underlying data are recorded at the daily frequency and range from 2016 through 2022. For visual clarity, the figure reports the five-day moving average value of each quantity. The sample runs from May 2016 through June 2022.

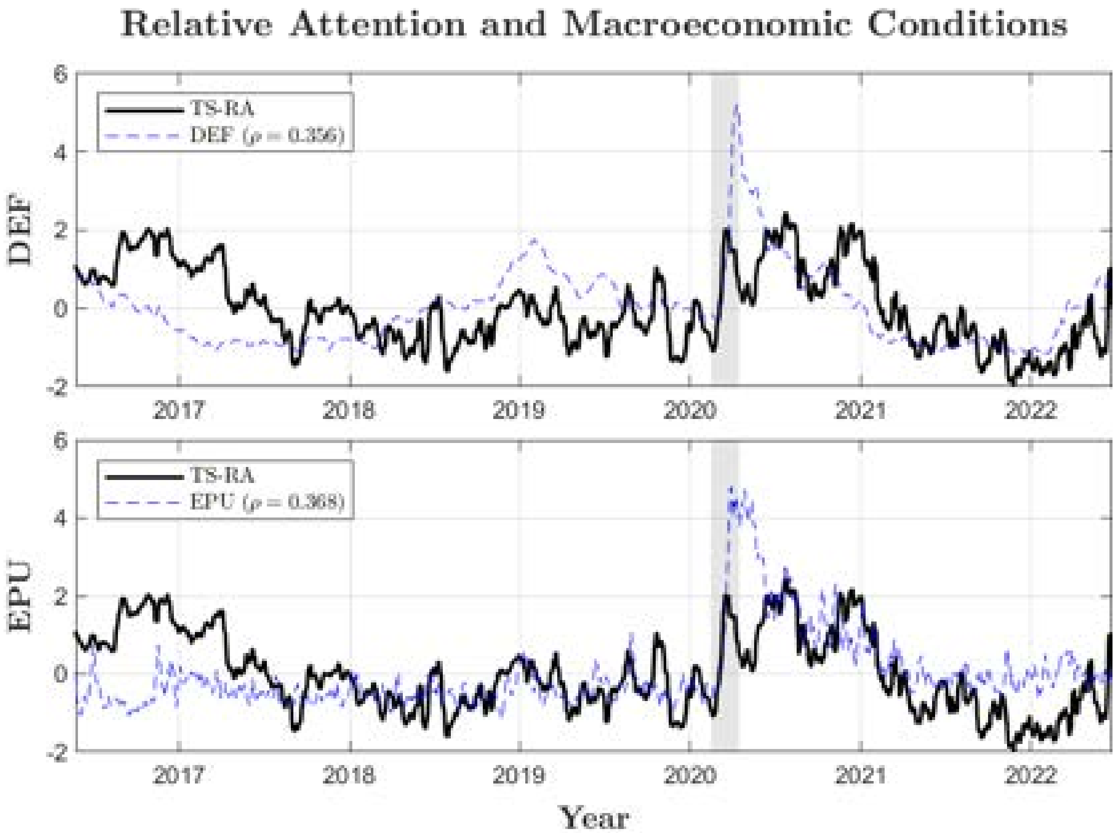
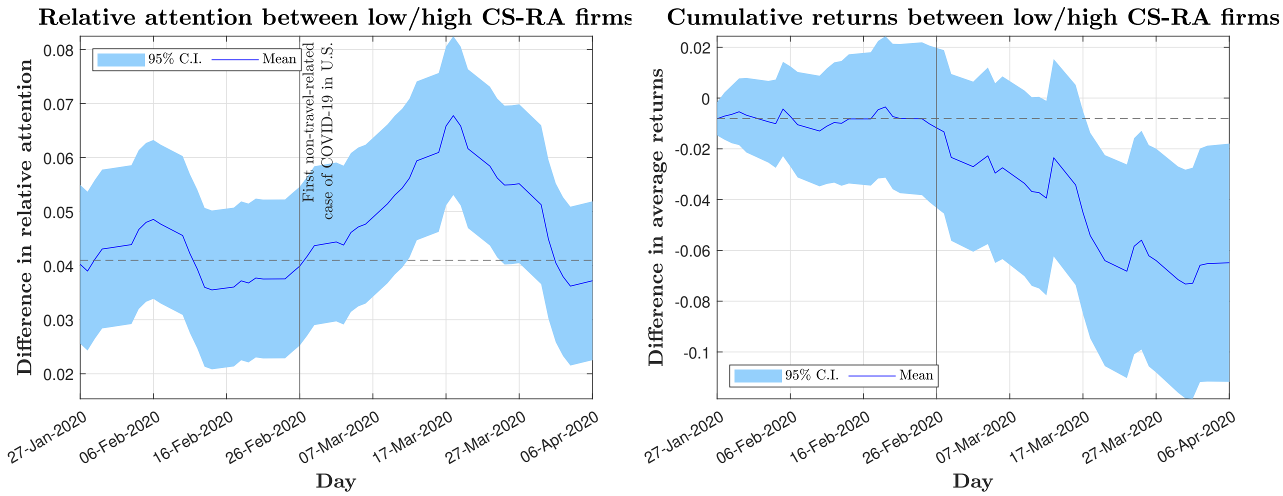


Figure 6: Aggregate Relative Attention ($\widetilde{\text{TS-RA}}_t$) and Macroeconomic Conditions

This figure displays the time series of raw relative attention $\widetilde{\text{TS-RA}}_t$ from equation (2) against the time series of two prominent variables that reflect the state of macroeconomic conditions: (i) the corporate default spread (DEF) and (ii) the Economic Policy Uncertainty index from Baker et al. (2016) (EPU). The underlying data are recorded at the daily frequency and range from 2016 through 2022. For visual clarity, the figure reports the five-day moving average value of each quantity as well as the time-series correlation between $\widetilde{\text{TS-RA}}_t$ and each macroeconomic variable of interest. The sample runs from May 2016 through June 2022.



(a) Reading activity **(b) Returns**

Figure 7: Firm Relative Attention ($\text{TS-RA}_{i,t}$) Around the Onset of the COVID-19 Pandemic
The figure reports the results of a high-frequency event study in which we examine how the average firm's relative attention to macroeconomic conditions, as measured using ($\text{TS-RA}_{i,t}$) from equation (2), changes upon the onset of the COVID-19 pandemic in early 2020. The figure is constructed as follows. First, we estimate the value of $\text{CS-RA}_{i,t}$ for each firm on February 26, 2020, which was the day the first nontravel-related case of COVID-19 was confirmed in the United States (depicted by the vertical black lines). Next, we transform $\text{CS-RA}_{i,t}$ to be uniformly distributed over the interval of $[-1, 1]$. We then compute the value-weighted average value of $\widetilde{\text{TS-RA}}_{i,t}$ from equation (2) across all firms with a positive (negative) value of these transformed $\text{CS-RA}_{i,t}$ scores. For each group, greater weights are placed on firms whose relative attention is more extreme (i.e., close to either negative one or one). The left-hand side of the figure then plots the difference in $\widetilde{\text{TS-RA}}_{i,t}$ between firms with high and low amounts of relative attention to macroeconomic conditions. The right-hand side of the figure then plots the difference in the cumulative returns of these two groups, starting on January 27, 2020. Each sub-figure also displays the following two quantities: (1) the shaded blue region depicts the 95% confidence interval associated with relative attention on the left-hand side and cumulative returns on the right-hand side; and (2) the horizontal dashed line represents the sample average of each difference between January 27 and February 25, 2020, the period preceding the onset of concerns related to the COVID-19 pandemic.

(a) **Raw topic cloud for Chemicals Manufacturing industry**

(b) ***tf-iaf*-weighted topic cloud for Chemicals Manufacturing industry**

(c) **Raw topic cloud for Computer and Electronics Manufacturing industry**

(d) ***tf-iaf*-weighted topic cloud for Computer and Electronics Manufacturing industry**

Figure 8: Raw and *tf-iaf*-weighted topic clouds for two industries
This figure compares the raw and *tf-iaf*-weighted number of interactions across topics for the Chemicals Manufacturing (NAICS 325) and Computer and Electronics Manufacturing (NAICS 334) industries. The figure highlights the difference in topic relevance between the two weighting methods. The *tf-iaf* weights employed for each industry are given by equation (5). The data underlying these topic clouds are from the week ending November 17, 2018.

Online Appendix

OA.1 Variable Definitions

Asset growth. The asset growth rate is computed as the change in total assets (Compustat Quarterly item ATQ) between fiscal quarter t and $t-1$.

Asset tangibility. In analyses that feature quarterly data, asset tangibility is defined as the ratio of a firm's net property, plant, and equipment (Compustat Quarterly item PPENTQ) to the firm's total assets (Compustat Quarterly item ATQ) in fiscal quarter t. In analyses that feature annual data, asset tangibility is defined as the ratio of a firm's net property, plant, and equipment (Compustat Annual item PPENT) to the firm's total assets (Compustat Annual item AT) in fiscal year t.

CAPM $\boldsymbol{\beta}$. The CAPM beta is estimated on a firm-by-firm basis by running a regression of each firm's daily excess stock returns from CRSP onto the excess market return over the preceding 252 trading days.

Employment growth. The firm's employment growth rate is computed as the growth rate in the total number of employees (Compustat Annual item EMP) between fiscal years t and $t-1$.

Implied cost of capital (ICC). To construct this measure, we extract consensus estimates of earnings-per-share before extraordinary expenses ($FEPS_{t+i}$) at time $t+i$ from IBES, the book-value-per-share (B_t) at time t from Compustat, and the month-end stock price (P_t) at time t from CRSP. Our estimate of the ICC is then the solution to the following internal rate of return calculation:

$$P_t = B_t + \frac{FEPS_{t+1}/B_t - ICC_t}{(1+ICC_t)} \times B_t + \frac{FEPS_{t+2}/B_{t+1} - ICC_t}{(1+ICC_t)^2} \times B_{t+1} + TV_t. \tag{11}$$

The solution to this equation provides us with an estimate of a stock's ICC for each month, which is the frequency at which analysts update or reiterate their earnings forecasts in IBES.[16]

In general, only one- and two-year ahead earnings forecasts are reliably available. Hence, we project current earnings forward so as to estimate the terminal value term (i.e., TV_t) in equation (11). Following Pástor, Sinha, and Swaminathan (2008), we define the time-series of all analyst

[16] We rely on analyst forecasts rather than using predictive regressions (see, e.g. Hou, Van Dijk, and Zhang, 2012) due to our short sample and the large number of unanticipated events that occurred between 2016 and 2021 (e.g. the 2016 election results and the economic consequences of the COVID-19 pandemic).

estimated implied return-on-equity (i.e., $FROE_{t+\tau} = FEPS_{t+\tau}/B_{t+\tau-1}$) such that they converge to their long-run, pre-sample industry means (3-digit NAICS from 1995-2015) over a 15 year period,

$$TV_t = \sum_{\tau=3}^{14} \frac{FROE_{t+\tau} - ICC_t}{(1 + ICC_t)^\tau} \times B_{t+\tau-1} + \frac{FROE_{t+15} - ICC_t}{ICC_t \times (1 + ICC_t)^{14}} \times B_{t+14}, \tag{12}$$

where $FROE_{t+15}$ is the firm's industry (NAICS 3-digit) historical mean return-on-equity between 1995 and 2015 (i.e., before our sample). We follow the clean surplus methodology to estimate the future book values, where $B_{t+\tau} = B_{t+\tau-1} + FEPS_{t+\tau} \times (1 - \text{Payout Ratio})$. Like before, we estimate the historical long-run payout ratio using a firm's industry mean payout between 1995 and 2015.

Inventory growth. The inventory growth rate is computed as the change in inventories (Compustat Quarterly item INVTQ) between fiscal quarters t and $t-1$.

Investment growth. The physical investment growth rate is computed as the change in a firm's net property, plant, and equipment (Compustat Quarterly item PPENTQ) between fiscal quarters t and $t-1$.

Leverage. In analyses that feature quarterly data, leverage is defined as the sum of a firm's debt in current liabilities (Compustat Quarterly item DLCQ) plus long-term debt (Compustat Quarterly item DLTTQ) scaled by the firm's total assets (Compustat Quarterly item ATQ) in fiscal quarter t. In analyses that feature annual data, leverage is defined as the sum of a firm's debt in current liabilities (Compustat Annual item DLC) plus long-term debt (Compustat Annual item DLTT) scaled by the firm's total assets (Compustat Annual item AT) in fiscal year t.

Profitability. In analyses that feature quarterly data, profitability in fiscal quarter t is defined as the ratio of net income in fiscal quarter t (Compustat Quarterly item NIQ) to total assets in fiscal quarter $t-1$ (Compustat Quarterly item ATQ). In analyses that feature annual data, profitability in fiscal year t is defined as the ratio of net income in fiscal year t (Compustat Annual item NI) to total assets in fiscal year $t-1$ (Compustat Annual item AT).

Size. In analyses that feature quarterly data, firm size is measured using the natural logarithm of total sales (Compustat Quarterly item SALEQ) in fiscal quarter t. In analyses that feature annual data, firm size is measured using the natural logarithm of total sales (Compustat Annual item SALE) in fiscal year t.

Tobin's Q. In analyses that feature quarterly data, Tobin's q in fiscal quarter t is defined as the sum of the book value of total assets (Compustat Quarterly item ATQ) minus the book value of common equity (Compustat Quarterly item CEQQ) plus the market value of common equity (Compustat Quarterly item CSHOQ multiplied by the end of fiscal quarter stock price given by Compustat Quarterly item PRCCQ), all divided by the book value of total assets. In analyses that feature annual data, Tobin's q in fiscal year t is defined as the sum of the book value of total assets (Compustat Annual item AT) minus the book value of common equity (Compustat Annual item CEQ) plus the market value of common equity (Compustat Annual item CSHO multiplied by the end of fiscal year stock price given by Compustat Annual item PRCC), all divided by the book value of total assets.

Instrumented Variables. In analysis that uses instrumented measures of firm volatility and returns we largely follow the methodology of Alfaro et al. (2024). First, at a daily frequency we de-risk each firm's excess returns on a 5-year rolling basis assuming a Carhart (1997) four factor model. Second, we estimate the contemporaneous exposure of the residuals ($r_{i,t}^{\text{risk-adj}}$) to shifts in 9 macroeconomic variables: the price growth of energy (WTI), exchange rate growth of 7 major currencies (AUD, CAD, EUR, JPY, SEK, CHF, and GBP), and fluctuations in economic uncertainty (the Baker et al. (2016) EPU). The sensitivities are estimated at 2-digit NAICS level

$$r_{i,k,t}^{\text{risk-adj}} = \alpha_k + \sum_c \beta_k^c \cdot r_t^c + \epsilon_{i,k,t},$$

where r_t^c is the growth rate of the macroeconomic variable c and βs are estimated on a 5-year rolling basis using daily data. The subscript k designates the industry of firm i.

Third we then estimate the returns and volatilities of the individual securities and those of the 9 instruments using annual compounding. We, however, pull the data as of each quarter; given our that our data is of a much shorter time-series than that Alfaro et al. (2024) our instrumented regressions are done on a quarterly rather than annual basis.[17] We then create two sets of instruments. Set S instrument changes in the volatility of a firm's excess returns (second-moments) and are defined by

$$s_{i,k,t-1}^c = |\beta_{k,t-3}^c| \cdot \Delta\sigma_{t-1}^c.$$

[17] In order to control for the overlapping estimates, we apply the panel HAC-robust standard errors of Driscoll and Kraay (1998) with a 4 quarter lag.

In words, we take the three-quarter lagged sensitivity of firm i's industry j to fluctuations in instrument c and multiply its absolute value by the quarter-over-quarter change in the volatility of c over the quarter. Set F, in contrast, is meant to instrument for changes in the firm's excess returns (first-moments) and are defined by

$$f^c_{i,k,t-1} = \beta^c_{k,t-3} \cdot r^c_{t-1}.$$

As discussed in Alfaro et al. (2024), our hope is that directional heterogeneity in exposure to first-moment shocks (e.g., oil (airlines) stocks having a negative (positive) β to oil stocks) contrasts directional commonality in exposure to second-moment shocks (i.e., the absolute value of β to oil is always greater than zero for both industries) such that an instrumented variable approach will better identify which moment (either first or second) drives fluctuations in firm-level macroeconomic-related reading.

Non-political risk. We measure each firm's non-political risk via the method developed by Hassan et al. (2019), who apply tools from computational linguistics to analyze the earnings call transcripts of each firm. This data is made available from `https://www.firmlevelrisk.com/`.

Political risk. We measure each firm's political risk via the method developed by Hassan et al. (2019), who apply tools from computational linguistics to analyze the earnings call transcripts of each firm. This data is made available from `https://www.firmlevelrisk.com/`.

Implied Volatility. We measure each firm's implied volatility on a given date t, Implied $\text{Vol}_{i,t,\tau}$, using τ-day-to-maturity volatility surface files from Option Metrics where the τ-surface used is determined by the frequency of the regression. That is, 91-day-to-maturity volatility surfaces are used for quarterly data and 365-day-to-maturity volatility surfaces are used for annual data. We follow the approach of Breeden and Litzenberger (1978) such that

$$\text{Implied Vol}^2_{i,t,\tau} = \frac{2}{R_{f,t} S^2_{i,t}} \left[\int_0^{F_{i,t,\tau}} \text{Put}_{i,t,\tau}(K)\, dK + \int_{F_{i,t,\tau}}^{\infty} \text{Call}_{i,t,\tau}(K)\, dK \right], \tag{13}$$

where $S_{i,t}$ and $F_{i,t,\tau}$ denote the spot and forward (from day t to $t+\tau$ days) prices of stock i, and $\text{Put}_{i,t,\tau}$ and $\text{Call}_{i,t,\tau}$ denote the day t prices of European puts and calls on the market that expire at time $t+\tau$ days with strike K, respectively. The option prices are derived using the Black-Scholes formula and the implied volatilites from the volatility surface files.

OA.2 Model

In Section 3.2.1 we introduced a framework through which we motivate the economic content in CS-RA$_{i,t}$. Specifically, after applying equation (5) to firm-level reading we see that prominent macroeconomic-related topics shift from those related to the real economy to those related to hedging and mitigation activity. Similarly, topics that are not macroeconomic-related shift from those that are more general in nature to topics that are specific to a firm's production. CS-RA$_{i,t}$ can thus be thought to represent a noisy proxy of firm-employee reading dedicated to mitigation versus investment activity. The framework motivates that the degree to which a firm mitigates risk versus invests can be linked to a firm's aggregate risk exposure. By extension, CS-RA$_{i,t}$ in the cross-section of firms then becomes an ex-ante measure of firm exposure to macroeconomic shocks. This section formalizes this intuition in a one-period model in which the firm takes the exposure of its output prices to macroeconomic fluctuations *tomorrow* as exogenous and chooses the degree to which they hedge versus invest *today*.

The objective function of the firm is that of a price-taker facing a variance penalty,

$$\max_{h,i} \quad \pi_0 + E[\overline{m}\pi_1] - \frac{\beta}{2}\mathrm{Var}[\pi_1], \tag{14}$$

where $\overline{m}$ is the expected discount rate, π_1 (π_0) is next (this) period's profit, and β can be thought to represent the exposure of a firm's output price to an aggregate shock. This is the standard approach taken in the hedging literature (see, e.g., Acharya, Lochstoer, and Ramadorai, 2013). All decisions are made at time $t = 0$, and the model resolves at time $t = 1$. There are two choices for the firm to make—h and i—which impact the profitability of the firm in each period. The variable h denotes how much the firm chooses to hedge or mitigate risk, whereas the variable i reflects how much the firm chooses to invest in next period's production. Intuitively, hedging activity is costly for the firm to undertake, but a larger choice of h also reduces the firm's uncertainty about the future output price. Similarly, increasing i reduces profits at $t = 0$, but increases the amount of output the firm has available to sell at the prevailing output price at $t = 1$.

The key measure of risk in the model is σ, which is the expected variation in next period's aggregate output price. As discussed above, for a specific firm, their prices are scaled by risk exposure, β. Below we derive a comparative static that shows, ceteris paribus, that an increase in

β corresponds to an increase in h/i. In other words, firms with higher macroeconomic exposure attempt to mitigate that risk more aggressively as a fraction of firm investment.

As in Hong et al. (2023), we assume an Ak model for production, whereby the output the firm produces at time t is $y_t = Ak_t$. Here, k_t is capital available for the purpose of production at time t and $A > 0$ is the productivity of capital. The law of motion of capital is given as:

$$k_1 = i + (1 - \delta)k_0, \tag{15}$$

where δ is a depreciation rate.

We assume the following structure for the firm's profits in each period:

$$\pi_0 = p_0(y_0 - i) - \frac{\kappa_p}{2}h^2, \quad \text{and,} \tag{16}$$

$$\pi_1 = p_1 y_1 + (\mu_p - p_1)\, h. \tag{17}$$

In words, the firm jointly chooses to invest and/or mitigate risk at period 0. The choice to mitigate, however, carries a quadratic cost with scaling parameter $\kappa_p > 0$. While hedging reduces the firm's profits in period zero, the benefit of risk mitigation is that the firm reduces the variance of its profit in period one. One could microfound this cost in a general equilibrium setup; for example, Acharya et al. (2013) emphasize limits to arbitrage from the perspective of the suppliers—e.g., financial intermediaries—of hedging instruments (e.g., futures contracts). As noted above, the output price in period one, p_1, is risky. This price has a mean μ_p and a standard deviation $\sigma > 0$.

Proposition 1. *There is a positive relationship between firm risk exposure β and the ratio h/i:*

$$\frac{\partial(h/i)}{\partial\beta} > 0. \tag{18}$$

The proof is in the next section of this appendix.

With this framework in mind, we take CS-RA$_{i,t}$ to reflect a high-frequency and news-based proxy for the quantity h/i. Consequently, CS-RA$_{i,t}$ is closely tied to a firm's risk exposure. The positive relationship between the ratio of hedging to investment (and thus CS-RA$_{i,t}$) and the firm's exposure to risk is intuitive: as exposure to risk escalates, firms are inclined to allocate a greater proportion of their resources towards risk management relative to their investments. This behavior is contingent on (i) mitigation activity being costly for the firm and (ii) the cost function (i.e., both the quadratic functional form and κ_p in our idealized setup) being on average the same across firms (see, e.g., Hong et al., 2023, for similar assumptions underlying firm-level mitigation activity).

Finally, a corollary follows from Proposition 1.

Corollary 1. *h/i in the cross section of firms will relate*

1. *positively to hedging activity as* $\frac{\partial(h/i)}{\partial h} > 0$*, and*
2. *negatively to investment activity as* $\frac{\partial(h/i)}{\partial i} < 0$.

These results can also be rationalized from an NPV perspective. Higher exposure to aggregate risk (i.e., β) imply higher higher hurdle rates and thus lower investment activity.

OA.2.1 Extension to employee attention allocation

In the main text many of our empirical results make use of the TS-RA$_{i,t}$ measure (see Sections 3.1 and 5.1). The measure captures the proportion of total employee reading that is spent on general macroeconomic-related topics. We find that (i) in a time-series sense, TS-RA$_{i,t}$ increases substantially after a macroeconomic shock and (ii) that this increase is persistent in such a way that it is more closely related to a firm's exposure to aggregate economic uncertainty (shock to the second-moments) rather than the shock itself (shock to the first-moments). As the measure captures the reading of largely rank-and-file employees, we assume that it may reflect the employees trying to understand or learn about how the macroeconomic environment impacts them—e.g., their human capital or employment risk—rather than concerns about only the firm's profits. Based on this intuition, TS-RA$_{i,t}$ is always a dependent (or reactive) variable in our predictive regressions. A question, however, arises: given the firm's problem above, how would the relationships we find with TS-RA$_{i,t}$ naturally occur? In this section we extend the model to include employee learning to answer this question.

The employees seek to maximize a similar problem to that of the firm:

$$\max \quad c_0 + E[\overline{m}c_1] - \frac{\beta}{2}\mathrm{Var}[c_1], \tag{19}$$

where c_t is consumption at time t, and the expectations and variances are subjective. The consumption of the employee at time 1 is given by $w + b\pi_1 + (y_1 - h)\,\xi$, where w is a constant wage component, and b measures the degree of exposure that employee earnings are exposed to firm profits (e.g., pay increases tied to success of the firm), and ξ is the residual riskiness of the employee's pay. This residual component can be thought to capture the human capital risk discussed above. The employee is uncertain about this term, but can, from their perspective, reduce this uncertainty by learning about the aggregate environment. ξ is distributed $\sim N\left(0,(\sigma_\xi/\tau)^2\right)$, where τ is thus the strength of the signal the employee chooses (and for which TS-RA$_{i,t}$ is a proxy). In addition, p_1 and ξ have a correlation of ρ. Intuitively, when aggregate risk increases the employee feels even greater uncertainty. This separates the choice of reading by rank-and-file employees from that of decision makers at the firm, building off the intuition that TS-RA$_{i,t}$ is a function of employee *self*-interested reading. The employees uncertainty is scaled by the firm's residual risk after the firm chooses its hedging ratio. Employee consumption in the two periods is then:

$$c_0 = w + b\pi_0 - \frac{\kappa_\tau}{2}(y_1 - h)^2(\tau - 1)^2. \tag{20}$$

$$c_1 = w + b\pi_1 + (y_1 - h)\xi. \tag{21}$$

Note that obtaining the signal incurs quadratic costs controlled by κ_τ. The $\tau - 1$ term above means that at $\tau = 1$ the employees are both not reducing uncertainty nor incurring the associated cost. The employee chooses τ to optimize their objective function.

Proposition 2. *If $\rho > 0$, there is a positive relationship between the precision of the signal acquired, τ, and both the size of the aggregate volatility, σ, and exposure to aggregate risk, β:*

$$\frac{\partial \tau}{\partial \sigma} > 0, \tag{22}$$

$$\frac{\partial \tau}{\partial \beta} > 0, \tag{23}$$

$$\frac{\partial^2 \tau}{\partial \sigma \partial \beta} > 0. \tag{24}$$

The proof of this is below in the next section of this appendix.

Our model reveals the response of employees in the face of heightened risk for the firm either in the form of higher aggregate risk or firm-specific exposure. In short, employees become motivated to acquire more accurate and precise information in order to better smooth their consumption profile. This behavior aligns with rational risk-aversion strategies, where individuals seek to reduce uncertainty by improving their knowledge of the factors affecting their well-being.

Finally, a corollary follows from proposition 2.

Corollary 2. *As $\frac{\partial(h/i)}{\partial \beta} > 0$ and $\frac{\partial \tau}{\partial \beta} > 0$ then also $\frac{\partial \tau}{\partial(h/i)} > 0$.*

In closing, if TS-RA$_{i,t}$ is a valid proxy for τ and CS-RA$_{i,t}$ is a valid proxy for h/i then these theoretical results map to our empirical findings from Sections 3.1 and 5.1.

OA.2.2 Model Proofs

Proof of Proposition 1. We calculate the expected period 1 discounted profits:

$$E[\overline{m}\pi_1] = E[\overline{m}(p_1(y_1 - h) + \mu_p h)] = E[\overline{m}p_1 y_1],$$

where we use the fact that $E[\overline{m}p_1] = \overline{m}\mu_p$. So the producer problem is to solve:

$$\max_{i,h}\ p_0 y_0 - i p_0 - \frac{\kappa_p}{2}h^2 + E[\overline{m}p_1 y_1] - \frac{\beta}{2}\text{Var}[p_1(y_1 - h)] \tag{25}$$

$$\text{subject to } y_1 = A(i + (1-\delta)k_0). \tag{26}$$

Then we should have:

$$\max_{i,h}\ p_0 y_0 - i p_0 - \frac{\kappa_p}{2}h^2 + E[\overline{m}p_1 A(i + (1-\delta)k_0)] - \frac{\beta}{2}\text{Var}[p_1(A(i + (1-\delta)k_0) - h)]$$

Given that $p_1 \sim N\left(\mu_p, \sigma^2\right)$, we can write the above problem as:

$$\max_{i,h}\ p_0 y_0 - i p_0 - \frac{\kappa_p}{2}h^2 + E[\overline{m}p_1 A(i + (1-\delta)k_0)] - \frac{\beta}{2}\sigma^2\left(A(i + (1-\delta)k_0) - h\right)^2$$

This yields the first order conditions (FOC):

$$[i]:\quad -p_0 + E[\overline{m}p_1 A] - \beta\sigma^2(A^2(i + (1-\delta)k_0) - Ah) = 0 \tag{27}$$

$$[h]:\quad -\kappa_p h + \beta\sigma^2\left(A(i + (1-\delta)k_0) - h\right) = 0 \tag{28}$$

To verify that this represents a global maximum:

$$V_{ii} = -\beta\sigma^2 A^2 < 0$$

$$V_{hh} = -\kappa_p - \beta\sigma^2 < 0$$

$$\text{and}\quad V_{ih} = V_{hi} = \beta\sigma^2 A$$

Such that the determinant of the hessian is:

$$D(H) = (-\beta\sigma^2 A^2)(-\kappa_p - \beta\sigma^2) - (\beta\sigma^2 A)^2$$

$$= \beta\sigma^2 A^2 \kappa_p + \beta^2\sigma^4 A^2 - \beta^2\sigma^4 A^2$$

$$= \beta\sigma^2 A^2 \kappa_p > 0.$$

These results hold as long as $\kappa_p > 0$. We then solve the $[i]$ FOC w.r.t. i:

$$i = \frac{E[\overline{m}p_1 A] - p_0}{\beta\sigma^2 A^2} - (1-\delta)k_0 + \frac{h}{A}.$$

Plugging this into the $[h]$ FOC:

$$h = \frac{E[\overline{m}p_1 A] - p_0}{A\kappa_p},$$

which means

$$i = \frac{E[\overline{m}p_1 A] - p_0}{\beta\sigma^2 A^2} - (1-\delta)k_0 + \frac{E[\overline{m}p_1 A] - p_0}{A^2\kappa_p},$$

and

$$\frac{h}{i} = \frac{\frac{E[\overline{m}p_1 A] - p_0}{A\kappa_p}}{\frac{E[\overline{m}p_1 A] - p_0}{\beta\sigma^2 A^2} - (1-\delta)k_0 + \frac{E[\overline{m}p_1 A] - p_0}{A^2\kappa_p}}.$$

The derivative w.r.t. β:

$$\frac{\partial(h/i)}{\partial\beta} = \frac{\frac{(E[\overline{m}p_1 A] - p_0)^2}{\sigma^2 A^3 \kappa_p}}{\left(\frac{E[\overline{m}p_1 A] - p_0}{A^2\sigma^2} - \beta(1-\delta)k_0 + \beta\frac{E[\overline{m}p_1 A] - p_0}{A^2\kappa_p}\right)^2} > 0. \qquad \square \tag{29}$$

Proof of Proposition 2. Using the distribution of p_1 and ξ, we write the employee problem as:

$$\max_{\tau} \quad w + b\pi_0 - \frac{\kappa_\tau}{2}(y_1 - h)^2(\tau - 1)^2 + E[\overline{m}(w + bp_1 y_1 + (y_1 - h)\xi)]$$
$$-\frac{\beta}{2}\left(b^2(y_1 - h)^2\sigma^2 + 2b(y_1 - h)^2\rho\sigma\sigma_\xi/\tau + (y_1 - h)^2\sigma_\xi^2/\tau^2\right) \tag{30}$$

The FOC is:

$$-\kappa_\tau(\tau - 1) + \beta(b\rho\sigma\sigma_\xi/\tau^2 + \sigma_\xi^2/\tau^3) = 0$$

Calculating the comparative static w.r.t. β we have:

$$\frac{\partial \tau}{\partial \beta} = \frac{b\rho\sigma\sigma_\xi/\tau^2 + \sigma_\xi^2/\tau^3}{\kappa_\tau + \beta\left(2b\rho\sigma\sigma_\xi/\tau^3 + 3\sigma_\xi^2/\tau^4\right)} > 0. \tag{31}$$

Similarly, calculating the comparative static w.r.t. σ_ξ we have:

$$\frac{\partial \tau}{\partial \sigma} = \frac{\beta b\rho\sigma_\xi/\tau^2}{\kappa_\tau + \beta\left(2b\rho\sigma\sigma_\xi/\tau^3 + 3\sigma_\xi^2/\tau^4\right)} > 0. \tag{32}$$

$$\frac{\partial^2 \tau}{\partial \sigma \partial \beta} = \frac{\kappa_\tau b\rho\sigma_\xi/\tau^2 + \beta b\rho\sigma_\xi^3/\tau^6}{\left(\kappa_\tau + \beta\left(2b\rho\sigma\sigma_\xi/\tau^3 + 3\sigma_\xi^2/\tau^4\right)\right)^2} > 0. \qquad \square \tag{33}$$

That is, all three are positive as long as $b > 0$ and $\rho > 0$.

OA.3 Additional Tables and Figures

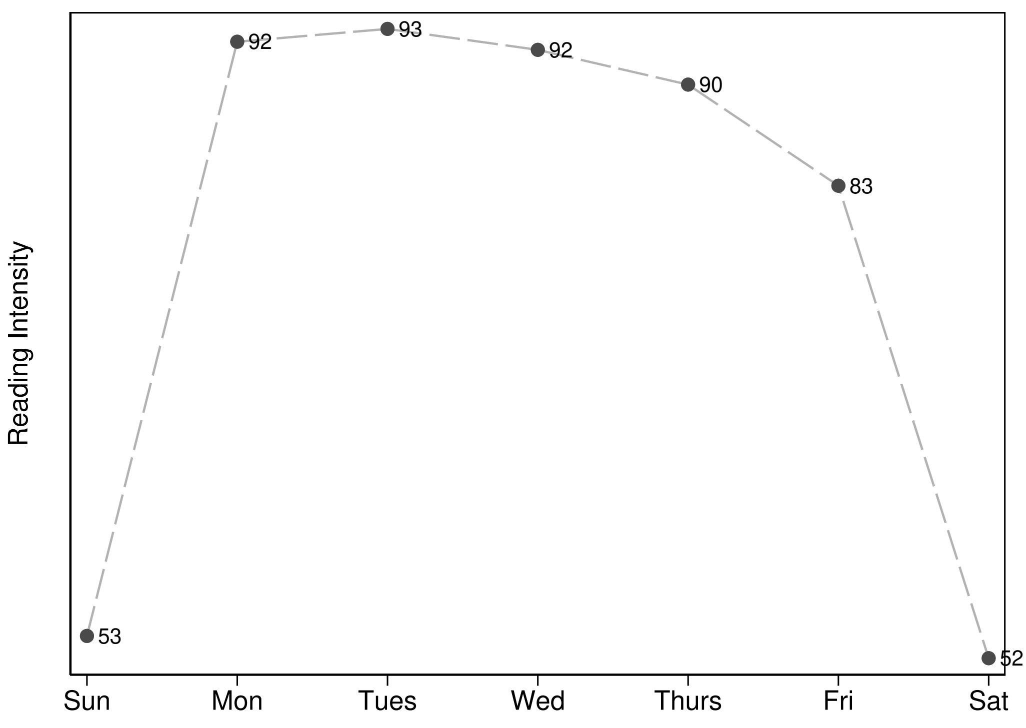


Figure OA.3.1: Average number users per day of the week
The figure plots the average number of unique users at each firm that interact with the Consortium's data on each day of the week. The set of domains over which these numbers are averaged cover both private and public firms from 2016 to 2022.

Table OA.3.1: Summary Statistics (Controls).
The table presents the summary statistics associated with the key real outcomes and control variables employed in Section 4. Here, "N" refers to the total number of observations for each variable, "SD" denotes the variable's standard deviation, and "p25" ("p75") refers to the 25^{th} (75^{th}) percentile of the variable's distribution. All data are quarterly, except for the employment growth rate, which is annual, and spans 2016 through 2022.

	N	Mean	SD	p25	Median	p75
Asset growth	52764	0.0363	0.1985	-0.0315	0.0042	0.0405
Employment growth	11554	0.0680	0.2949	-0.0476	0.0215	0.1224
Sales growth	48320	0.0882	0.4823	-0.0576	0.0244	0.1153
Log(Sales)	48640	4.7980	2.6002	3.3425	5.1538	6.5673
Tobin's q	52764	2.5606	2.2098	1.2267	1.7824	3.0016
Profitability	49521	-0.0296	0.1021	-0.0376	0.0032	0.0188
Leverage	52454	0.2695	0.2432	0.0529	0.2353	0.4084
Tangibility	51965	0.2378	0.2417	0.0605	0.1437	0.3355

Table OA.3.2: Top 20 Macro-related Topics

This table reports the top 20 macro-related news topics that received the most attention (measured by the topic frequency, inverse aggregate frequency (*tf-iaf*) scores) using data from 2016 through 2022. The topics are ranked by their average *tf-iaf* weights over the sample period. There *tf-iaf* weights are computed at the three-digit NAICS industry level to mitigate extreme weights. The table includes three columns: "Name," which denotes the topic's name in the Consortium's taxonomy, "Rank," which denotes the rank of the topic's *tf-iaf* weight over the sample period, in descending order of average attention received, and "Description," which describes the concept behind each topic according to the Consortium's taxonomy.

Name	Rank	Description
Letter of Credit	1	A letter issued by a bank to another bank (typically in a different country) to serve as a guarantee for payments made to a specified person under specified conditions
Enterprise Asset Management	2	Optimal lifecycle management of the physical assets of an organization.
Inventory Turnover	3	A measure of the number of times inventory is sold or used in a time period such as a year.
Mezzanine Finance	4	A hybrid of debt and equity financing that gives the lender the rights to convert to an ownership or equity interest in the company in case of default, after venture capital companies and other senior lenders are paid
Solvency II Directive 2009	5	Risk-focused fundamental framework for the insurance industry
Federal Financial Institutions Examination Council (FFIEC)	6	The Federal Financial Institutions Examination Council (FFIEC) is a formal U.S. government interagency body composed of five banking regulators that is empowered to prescribe uniform principles standards and report forms to promote uniformity in the supervision of financial institutions
Basel II	7	Basel II is the second of the Basel Accords, (now extended and partially superseded[clarification needed] by Basel III), which are recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision.
Home Mortgage Disclosure Act (HMDA)	8	The Home Mortgage Disclosure Act (or HMDA, pronounced HUM-duh) is a United States federal law that requires certain financial institutions to provide mortgage data to the public. Congress enacted HMDA in 1975.
Hedge Accounting	9	Hedge accounting is a method of accounting where entries for the ownership of a security and the opposing hedge are treated as one.
Fixed Asset Management	10	An accounting process that seeks to track fixed assets for the purpose of financial accounting, preventive maintenance, and theft deterrence.
Treasury / Cash / Risk Management	11	The forecasting and evaluation of financial risks together with the identification of procedures to avoid or minimize their impact.
Capacity Utilization	12	Capacity utilization is a measure of the extent to which the productive capacity of a business is being used.
Committee of Sponsoring Organizations (COSO)	13	Organization that aims to combat corporate fraud
Automatic Exchange of Information (AEOI)	14	Automatic Exchange of Information (AEOI) involves the systematic and periodic transmission of bulk¡9d¿ taxpayer information by the source country to the residence country concerning various categories of income (e.g. dividends, interest, etc.).
Managed Portfolios	15	Portfolios that are managed by an investment professional who frequently buys and sells investments
EMV (Payment System)	16	Global standard for credit and debit payment cards based on chip card technology.
Bytecoin (BCN)	17	Bytecoin (ticker symbol BCN[2]) is the first cryptocurrency based on the CryptoNote technology with an open source code designed for anonymous cash settlement. BCN protects the user's privacy with impassive and anonymous transactions.
Capital Allowance	18	Capital allowances is the practice of allowing a company to get tax relief on tangible capital expenditure by allowing it to be expensed against its annual pre-tax income. Generally, the capital allowances will exist for only specified items of tangible capital expenditure, and the expensing is usually spread over a fixed period of years.
Internal Rate of Return (IRR)	19	The internal rate of return (IRR) is a method of calculating rate of return. The term internal refers to the fact that its calculation does not involve external factors, such as inflation or the cost of capital.
Commercial Paper	20	Unsecured, short-term debt instrument

Table OA.3.3: Real Outcomes and Firm Attention to Macroeconomic Risk: Additional Evidence
This table reports the results of estimating equation (9) to examine the relation between a firm's relative attention to macroeconomic risk (measured via equation (6)) and inventory growth (Panel A) and the growth rate of net property, plant, and equipment (Panel B). Each of these variables is defined in the Online Appendix OA.1. The specifications include combinations of industry, date, and industry-by-date fixed effects. Additionally, specifications (4) and (5) control for each firm's CAPM β, a common proxy for a firm's exposure to macroeconomic risk. Additionally, specification (5) features both the set of controls from Leary and Roberts (2014) and Whited and Wu (2006) measure of financial constraints. The data underlying this regression spans from 2016 through 2022, and all standard errors are clustered by firm and time.

	(1)	(2)	(3)	(4)	(5)
Panel A: Inventory growth					
CS-RA$_{i,t-1}$	-0.0370^{***}	-0.0353^{***}	-0.0351^{***}	-0.0362^{***}	-0.0190^{***}
	[−8.11]	[−7.63]	[−7.45]	[−7.65]	[−3.71]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				-0.0104^{***}	-0.0044
				[−2.78]	[−1.11]
Observations	37,648	37,621	37,216	37,216	31,527
R^2	0.0025	0.0223	0.0743	0.0746	0.0948
Panel B: PPENT Growth					
CS-RA$_{i,t-1}$	-0.0834^{***}	-0.0712^{***}	-0.0736^{***}	-0.0723^{***}	-0.0301^{***}
	[−13.23]	[−11.58]	[−11.90]	[−11.69]	[−4.41]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$				-0.0203^{***}	-0.0182^{***}
				[−4.10]	[−3.47]
Observations	53,002	52,967	52,513	52,513	41,196
R^2	0.0044	0.0945	0.1690	0.1693	0.2036
Date FE		+			
Industry FE		+			
Date × Industry FE			+	+	+
CAPM-beta				+	+
Controls					+

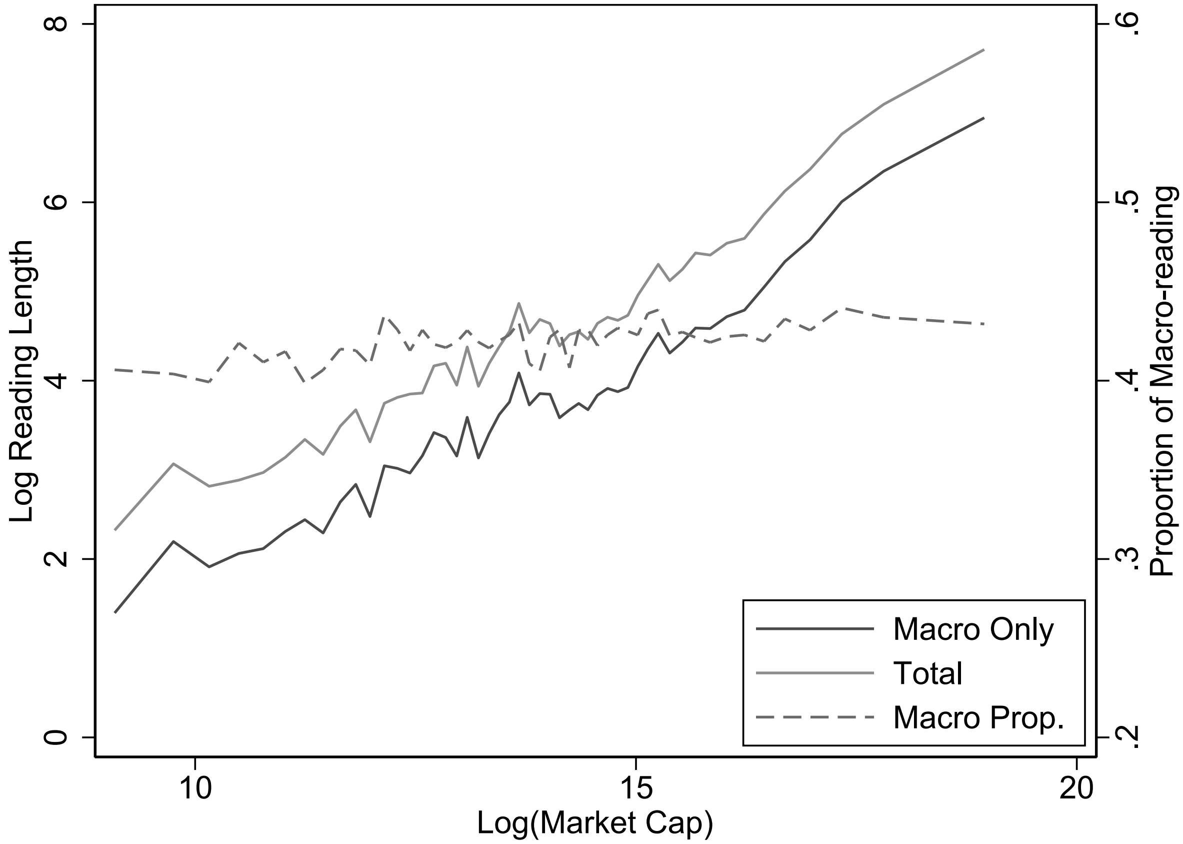


Figure OA.3.2: Relationship between Firm Size and Reading.
The figure presents the relation between firm reading and firm size, measured by the natural logarithm of a firm's market capitalization. We measure the time a firm allocates towards macroeconomic (total) news by the length of a firm's macroeconomic (total) reading vector, $\boldsymbol{tf}_{i,t}^{\text{Macro}}$ ($\boldsymbol{tf}_{i,t}^{\text{Total}}$). We then plot the average log length of each of these vectors across 200 size buckets. The figure also plots the cosine similarity between the two vectors as a function of firm size as represented by equation (1). The data underlying this analysis ranges from 2016 through 2022.

Table OA.3.4: Variance Decomposition by Fixed Effects and Firm Characteristics
The table presents a variance decomposition of the firm's relative attention to macroeconomic risk (i.e., CS-RA$_{i,t}$ from equation (6)) based on projections of CS-RA$_{i,t}$ on various combinations of fixed effects and firm characteristics. The variance of CS-RA$_{i,t}$ is decomposed into the variation attributable to firm characteristics, sector fixed effects, sector-by-date fixed effects, and firm fixed effects. The firm characteristics we consider include CAPM β, size, the book-to-market ratio, gross profitability, and asset growth. The sector fixed effects we consider reflect either 2-digit NAICS codes or 3-digit NAICS codes. The column denoted "No Fixed Effect" considers the possibility that all variation in CS-RA$_{i,t}$ is attributed to the firm-level characteristics. The data underlying this analysis ranges from 2016 through 2022.

	2-digit NAICS	3-digit NAICS	No Fixed Effect
Sector FE	2.42%	5.84%	
Sector × Date FE	24.60%	24.96%	
Firm-specific	72.98%	69.20%	
Permanent difference across firms, within sector-date	42.67%	39.69%	
Across firm-time residual	30.31%	29.51%	
Characteristics:			
Beta	1.75%	1.28%	1.43%
Size	4.99%	4.69%	5.15%
Book-to-Market	0.04%	0.05%	0.14%
Gross Profitability	0.81%	0.66%	0.80%
Asset Growth	0.27%	0.26%	1.04%
Characteristic Total	7.86%	6.94%	8.57%
Number of Sectors	64	247	

OA.4 Detailed Description of the Raw Attention Data

In this section we provide an overview of the Consortium's *domain-topic* dataset, which is used in the bulk of our empirical analyses. The main takeaway from this section is that while the Consortium's dataset covers a wide variety of topics (more than 7,000 in 2022), the majority of these topics are generally uninformative about each firm's business line(s). With this in mind, in Section 3.2 we propose a data-driven method for dealing with these uninformative, general, topics that encompass topics related to politics, entertainment, and sports.

To illustrate the fact that the vast majority of domain-topic interactions are only marginally informative about what a firm does, Figure OA.4.3 explores the distribution of topics for the Computer and Electronics Manufacturing sector (3-digit North American Industry Classification System (NAICS) code 334) during the week ending on November 17, 2018, a week that falls roughly in the midpoint of our sample period. The y-axis in this figure is a normalized measure of the intensity with which the employees of the firms in this sector are interacting with each topic in the given week. This normalized measure of the attention allocated to topic t is defined as

$$\text{NormInteractions}_t = \frac{\sum_{i \in \mathcal{I}} \text{Interactions}_{i,t}}{\max_t \left(\sum_{i \in \mathcal{I}} \text{Interactions}_{i,t} \right)}, \tag{34}$$

where $\text{Interactions}_{i,t}$ corresponds to the number of unique users at each firm i interacting with topic t in the given week. As this measure is scaled by the maximum number of unique interactions across all topics in a given week, $\text{NormInteractions}_{i,t}$ is a scalar that ranges from one (for the topic with the most interactions in a given week) to zero (for any topics with zero interactions in the given week). Numbers between these two extremes represent the amount of attention a given topic receives *relative* to the topic with the highest number of interactions in that week. The x-axis in this figure is then the rank associated with each topic's normalized attention score. These ranks are ordered from the topic with the most interactions, which has a rank of one, to the topic with the least number of interactions. In presenting these normalized interactions we truncate the rank at 250 to highlight the steep decline in attention as we move from the most popular topic to the less popular topics in a given week.

The topics with the most interactions towards the left of the figure are related a group of topics we consider to be current events. These high-interaction topics, which include "South by Southwest," "Call of Duty," and "US Secret Service," were highly relevant topics in the week underlying this exercise and appeared, for example, as headlines on the splash page of major publishers, such as USA Today. Figure OA.4.4 shows an example of three article headlines published by the Consortium's members for three of the top 10 topics underlying Figure OA.4.3.

The first headline highlights a pitch submission deadline for an event that takes place in March 2019 at the South By Southwest music festival, an event with a heavy tech presence that attracts hundreds of thousands of attendees each year. The second headline is regarding a new multiplayer map in the newest version of Call of Duty, a popular online video game. Finally, during the week in question, former President Trump was in Europe. A controversy erupted when the US Secret Service suggested that then President Trump avoid an event due to inclement weather. This goes to show that many of the topics with a high number of interactions in any given week very likely reflect news and current affairs. Untabulated analyses result in similar takeaways when we focus on sectors other than technology and different points in time.

As we move to the right of Figure OA.4.3, and the rank increases along the x-axis, we see a steep decline in the relative amount of attention paid to the topics with a rank between 150 to 250. Although these topics still draw more user interactions than the 5,750 or so other topics in the Consortium's dataset in the given week that we do not plot, these topics with a rank between 150 and 250 still only attract about 10% of the interactions dedicated to the more popular current events described above. Yet, these topics are still very general in nature and cover "Miami, Florida," a popular retirement destination, "Traditional IRA," a common retirement savings account, and "Environment for Aging," an event on senior living design. These results once again highlights how many of the topics with a high number of interactions are very general in nature and do not necessarily reflect details on the business line(s) of the underlying firms.

This begs the question, can we use the Consortium's data to glean any novel insights about a firm's attention to economically relevant news, such as firm risk mitigation, when most employees' attention is concentrated on common and current events? To demonstrate that the answer to this question is "yes," Figure OA.4.5 presents a histogram of interactions across the entire distribution of topics for firms in the computer and electronics sector (NAICS 334). Here, the x-axis reports the degree of topic interaction from topics with the least (zero) to most (one) interactions. The y-axis displays the proportion of topics that fall within some topic interaction interval.

Figure OA.4.5 indicates that the vast majority of the mass of the topic distribution is concentrated among topics that have a relatively low number of interactions. In fact, about 85% of the topics that firms in the computer and electronics sector interacted with during the week of 11/17/2018 received less than than 20% of the interactions dedicated to the 10 most popular topic in the sector that week (i.e., "Live Streaming," "Google +," "South By Southwest," "Call of Duty," " US Secret Service," etc.). For instance, topics inherently related to firms in the computer and electronics sector such as "disk-based backup and storage," "circuit design" and "cloud access security broker" each received about 5% or less of the interactions dedicated to "South by Southwest" and "Call of Duty" that week. Likewise, while topics related to business-relevant risks that we want to focus on, such as "credit risk," "exchange rate," and "cost of capital," received relatively more attention than "circuit design," these risk-related topics still received far less attention than many current events.

Overall, Figure OA.4.3 confirms that the Consortium's data can indeed help us to glean novel insights about a firm's attention to economic uncertainty, provided that we are careful to account for the fact that the bulk of the average employee's attention is, unsurprisingly, dedicated to reading about current events. A corollary from this figure is that simple metrics of firm attention, such as the amount of total reading per employee, are very likely uninformative about the economic environment in which the firm is operating.

Consequently, Section 3.2 develops an intuitive measure of a firm's attention to the risk that immunizes against the aforementioned concern in two steps. First, we define a set of topics that reflect uncertainty-related news, articles, and events. Second, we develop a measure of attention to these uncertainty-related topics that implicitly down weights interactions with topics that are common across all firms. This weighting scheme is motivated by the large literature on natural language processing (e.g., Gentzkow et al. (2019)) and essentially down weights a firm's attention to topics that all other firms are also reading about (e.g., "Call of Duty") and up weights topics that are more likely economically relevant and firm-specific (e.g., "Credit Risk" for all firms and "Circuit Design" for firms in the computer and electronics sector).

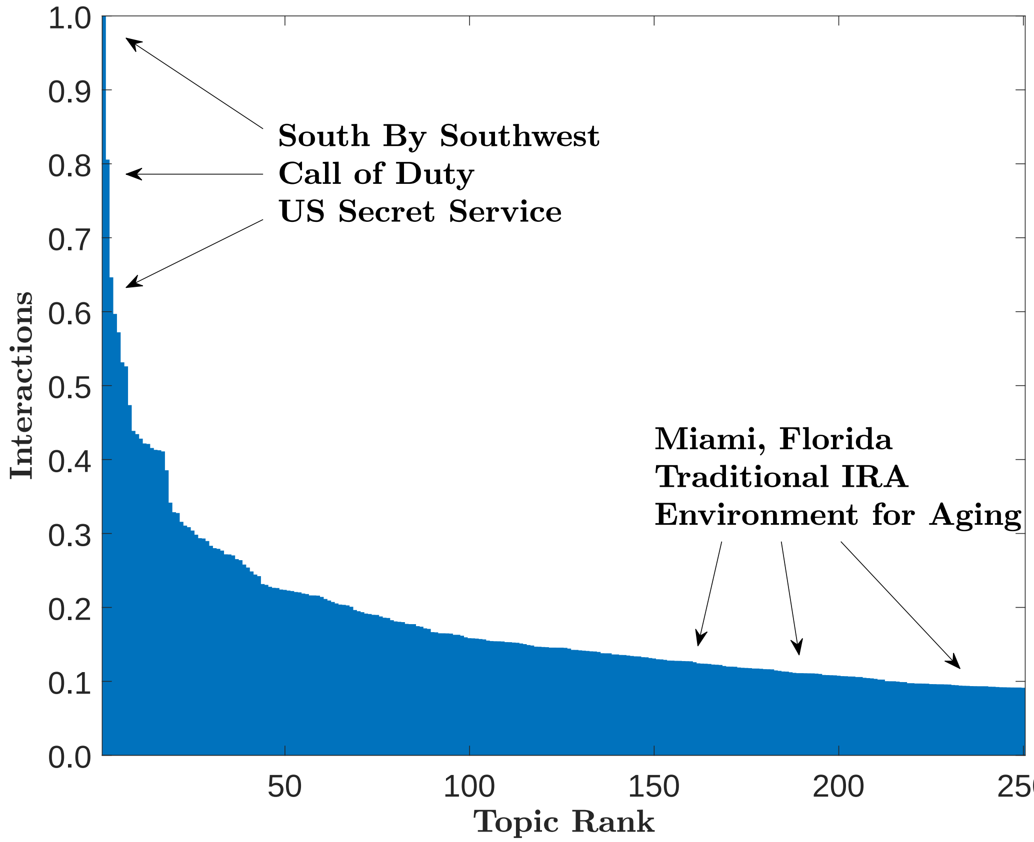


Figure OA.4.3: Bar chart of normalized attention to topics.
This bar chart represents the normalized number of interactions, defined according to equation (34), on the y-axis and rank of various topics across all firms in the North American Industry Classification (NAICS) code 334 industry — the Computer and Electronics Manufacturing industry — during the week ending 11/17/2018 on the x-axis.

USA TODAY

POLITICS

President Trump blames Secret Service for canceling cemetery trip in France

David Jackson
USA TODAY

Published 9:02 a.m. ET Nov. 13, 2018 | **Updated 1:46 a.m. ET Nov. 14, 2018**

WASHINGTON – Still fighting political flaps that arose during his weekend in France, President Donald Trump said Tuesday that the Secret Service nixed a visit to a cemetery in the rain because the motorcade would have shut down Paris.

"When the helicopter couldn't fly to the first cemetery in France because of almost zero visibility, I suggested driving," Trump tweeted. "Secret Service said NO, too far from airport & big Paris shutdown."

Figure OA 14.4: News Headlines from 11/08/2018.
This figure captures a sample of headlines from Consortium publishers associated with 3 top topics from the week ending 11/08/2018. Panel (a) is a headline highlighting a submission deadline for a competitive South By Southwest VC pitch competition. Panel (b) highlights the release of an addition to the new release of Call of Duty. Panel (c) is a current events article about the US Secret Service.

Trump took heat back home after the White House announced it had canceled a trip to Aisne-Marne American Cemetery and Memorial near Paris "due to scheduling and logistical

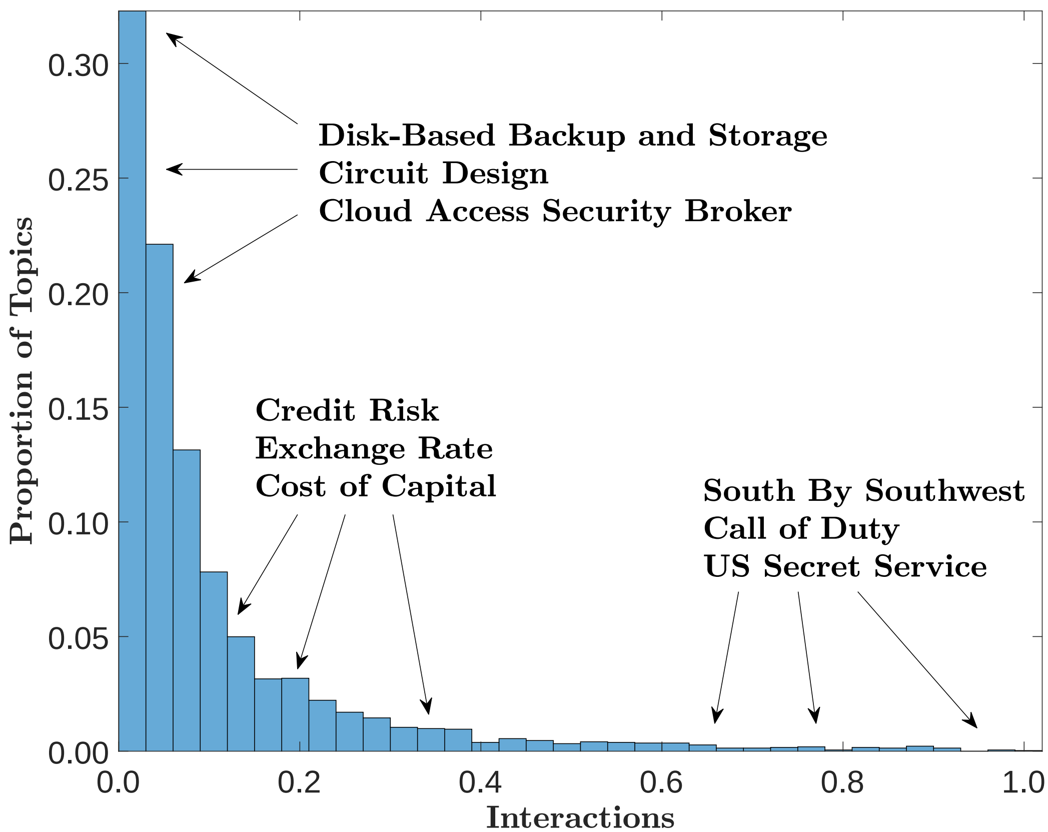


Figure OA.4.5: **Histogram.**
This figure is the histogram of topic interactions for all firms in NAICS 334 (Computer and Electronics Manufacturing) during the week ending 11/17/2018. We normalize the number of interactions with each topics by the number of interactions with the most popular topics in the given week (see equation (34). This measure of normalized interactions ranges from zero, for any topics with no interactions, to one, for the single topic with the most interactions.

OA.5 Risk Exposures and Pricing Errors

Section 4.2 links CS-RA$_{i,t}$ to firm-level measures of implied cost of capital, ICC$_{i,t}$, showing that firms moving from the lowest to highest value of CS-RA$_{i,t}$ have an ICC$_{i,t}$ that is between 110 and 150 basis points higher. Variations in ICC$_{i,t}$, however, may be related to asset-pricing anomalies rather then firm exposures to risk. To address this concern, in Section 5.1 directly links macroeconomic-related reading of firms' employees to variations in firm exposures to aggregate uncertainty using the instrumented variable approach of Alfaro et al. (2024). We find that employees at riskier firms today read relatively more macroeconomic news tomorrow. This relationship even holds with firm and date fixed effects, suggesting that this is strongly related to differences between firms in the cross-section. We, however, also found that high CS-RA$_{i,t}$ firms tend to be larger, have lower book-to-market (i.e., carry a potentially negative risk premia on these factors), but higher CAPM βs, higher gross profitability and lower asset growth (i.e., carry a potentially positive risk premia on these factors).

As the characteristic sorts reported in Section 5.2 were conducted on a univariate basis, this section provides similar evidence after jointly account for multiple risk factors. Using Fama and French (1993)-style regressions, we find that CS-RA$_{i,t}$ does indeed convey incremental information regarding expected rates of returns beyond the standard factors. After computing CS-RA$_{i,t}$ based on equation (6), we first average this measure across an entire calendar-quarter. We then compute the daily returns of a value-weighted long-short portfolio that buys firms in the highest decile of average CS-RA$_{i,t}$ and sells firms in the lowest decile of average CS-RA$_{i,t}$. This portfolio is then rebalanced at the beginning of every quarter (January 1st, April 1st, July 1st, and October 1st).

The daily returns of this long-short portfolio are then regressed onto the Fama and French (2015) factors, as well as the Carhart (1997) momentum factor. Table OA.5.5a presents the results. In keeping with the characteristic of the portfolio sorts from Section 5.2, we find that portfolio returns are negatively related to size (SMB$_t$) and positively related to both the gross profitability (RMW$_t$) and investment (CMA$_t$) factors. The portfolio returns, however, show a negative relationship to market returns (MKTRF$_t$) and a positive relationship to value (HML$_t$); the former becomes insignificant after adding all factors. Furthermore, momentum shows little relationship to the CS-RA$_{i,t}$-sorted portfolio.

Next, we ask if CS-RA$_{i,t}$ conveys any incremental information about expected rates of returns. We find that it likely does. In particular, in column (3), when we include the market, size, and value factors as controls, we obtain a weakly statistically significant alpha of 4.7% per year. This shrinks significantly when adding the gross profitability and investment factors to the regression. Despite being insignificant, however, the α is still around 200 basis points, which is close to that seen with the regressions using the implied cost of capital (ICC$_{i,t}$) in Section 4.2. Unfortunately, the weak statistical power using the Fama and French (1993) methodology may be related to our data's relatively short time-series, which only runs from May 2016 through June 2022.

Rather than measure cross-sectional differences in exposure using equation (6), we consider conducting these portfolio sorts using an alternative measure that weights topics depending on their importance within an industry. This measure, which we label as an *extensive margin* measure, computes industry-level *tf-iaf* weights, and multiplies these weights by an indicator variable that is equal to one if any firm i in industry j is paying attention to a given topic at time t. The advantage of this measure, which we refer to as CS-RA$^{ext}_{i,t}$, is that it is less likely to be influenced by any idiosyncratic reading behavior at the firm level.

We proceed as above in creating CS-RA$^{ext}_{i,t}$-sorted portfolios on a calendar-quarter basis, and regress the long-short portfolio returns onto standard factors. The results of these regressions are given in Table OA.5.5b. Overall, returns are much stronger and more stable. Including all factor controls, the CS-RA$^{ext}_{i,t}$ sorted portfolios generate an annualized alpha of 6.8%.

Table OA.5.5: Fama and French Regressions of CS-RA Sorted Portfolio
This table presents the risk exposures and pricing errors of a long-minus-short portfolio formed on the basis of each firm's relative attention to macroeconomic-related news, measured using equation (6). Specifically, we sort the cross section of firms at the end of each quarter t into 10 portfolios formed on the basis of CS-RA$_{i,t}$ averaged over the previous quarter. We hold each portfolio through the following quarter-end, at which point in time all portfolios are rebalanced. We compute the value-weighted return of a self-financing portfolio that buys the set of firms with the highest decile of CS-RA$_{i,t}$ and sells the set of firms with the lowest decile. We regress this value-weighted returns of these portfolios onto a constant (Column (1)), the excess market return (Column (2)), the Fama and French (1993) three-factor model (Column (3)), Fama and French (2015) five-factor model (Column (4)), and the Fama and French (2015) five-factor model that also features the momentum factor (Column (6)). The table reports exposure of the portfolio to each risk factor as well as the portfolio's pricing error (α). Panel A (Panel B) populates equation (6) using the extensive-margin measure of CS-RA$_{i,t}$ described in Section OA.5. The time-series regression are estimated using daily data from 2016 through 2022 and t-statistics, reported in brackets, are computed using Newey and West (1987) standard errors.

(a) Sorts Using Intensive Margin

	(1)	(2)	(3)	(4)	(5)	(6)
α	0.0379	0.0487	0.0473*	0.0339	0.0223	0.0223
	[1.23]	[1.62]	[1.85]	[1.45]	[1.01]	[1.01]
MKTRF_t		−0.0814***	−0.0486***	−0.0414***	−0.0113	−0.0113
		[−4.47]	[−4.26]	[−3.79]	[−1.20]	[−1.20]
SMB_t			−0.2951***	−0.2344***	−0.2084***	−0.2088***
			[−11.22]	[−8.86]	[−10.24]	[−9.87]
HML_t			0.2349***	0.1738***	0.0606***	0.0596***
			[13.91]	[10.36]	[3.67]	[3.12]
RMW_t				0.2571***	0.2659***	0.2654***
				[9.55]	[11.38]	[11.12]
CMA_t					0.3807***	0.3818***
					[11.76]	[11.41]
UMD_t						−0.0016
						[−0.12]
Observations	1,534	1,534	1,534	1,534	1,534	1,534
F-stat	.	20.0114	92.7200	93.1830	103.3557	86.3204

Table OA.5.5: Fama and French Regressions of CS-RA Sorted Portfolio—(Continued)

(b) Sorts Using Extensive Margin

	(1)	(2)	(3)	(4)	(5)	(6)
α	0.1000***	0.1022***	0.0974***	0.0749***	0.0679***	0.0681***
	[2.86]	[2.93]	[3.44]	[3.07]	[2.86]	[2.89]
MKTRF_t		−0.0161	0.0316*	0.0437***	0.0618***	0.0618***
		[−0.67]	[1.95]	[2.96]	[4.21]	[4.54]
SMB_t			−0.4370***	−0.3350***	−0.3194***	−0.3073***
			[−18.32]	[−14.97]	[−15.04]	[−14.68]
HML_t			0.1542***	0.0517***	−0.0164	0.0136
			[8.07]	[3.06]	[−0.88]	[0.66]
RMW_t				0.4319***	0.4372***	0.4501***
				[14.57]	[14.43]	[14.86]
CMA_t					0.2288***	0.1963***
					[6.83]	[5.86]
UMD_t						0.0483***
						[3.09]
Observations	1,534	1,534	1,534	1,534	1,534	1,534
F-stat	.	0.4459	113.4331	148.3137	131.8178	112.0885

Table OA.5.6: Firm Reading Versus Measures of Uncertainty

This table reports the results of estimating equation (9) using the same set of dependent variables as in Table 4 and 6, but after removing the COVID-related period between January and December 2020. Column (1) has asset growth, column (2) sales growth, column (3) employment growth, column (4) the hedging intensity measure of Campello et al. (2011) and column (5) the regulatory intensity measure of Kalmenovitz (2022) as the dependent variables. Variable construction is defined in the Internet Appendix OA.1. All specifications include industry-by-date fixed effects, and firm-level controls including CAPM β, the variables from Leary and Roberts (2014) and the Whited and Wu (2006) measure of financial constraints. The data underlying this regression span from 2016 through 2022, and all standard errors are clustered by firm.

	Real Activity			Mitigation Activity	
	Assets	Sales	Employment	Hedging	Compliance
	(1)	(2)	(3)	(4)	(5)
CS-RA$_{i,t-1}$	-0.0106^{**}	-0.0492^{***}	-0.0229^{**}	0.0262	1.6126^{**}
	[−2.34]	[−4.05]	[−2.47]	[1.22]	[2.01]
CS-RA$_{i,t-1}$ × ΔEPU$_{t-1}$	-0.0145^{***}	−0.0051	0.0031	0.0546^{***}	−0.1657
	[−3.64]	[−0.49]	[0.58]	[3.78]	[−0.39]
Date × Industry FE	+	+	+	+	+
CAPM-beta	+	+	+	+	+
Controls	+	+	+	+	+
Observations	33,647	33,274	6,880	7,165	20,343
R^2	0.0967	0.0940	0.1475	0.3428	0.7188