BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2024-37

Last updated: October 8, 2024

An Anatomy of Firms' Political Speech

by Pablo Ottonello,[1] Wenting Song[2] and Sebastian Sotelo[3]

[1] Department of Economics
University of Maryland and
National Bureau of Economic Research
ottonell@umd.edu

[2] Canadian Economic Analysis Department
Bank of Canada
wsong@bankofcanada.ca

[3] Department of Economics
University of Michigan and
National Bureau of Economic Research
ssotelo@umich.edu

DOI: https://doi.org/10.34989/swp-2024-37 | **ISSN** 1701-9397

Acknowledgements

We thank Stephanie Houle and Thomas Winberry for helpful comments. We thank Athiwat Thoopthong, Stella Xu, Frankie Andrews, Christina Blessing, Taylor Fredman, Olivia Goff, Braden Krajcirovic, Claudia Lujan Trujillo, Peter Micciche, Jenna Nagie, Aidan Owens, Angelina Parker, Nicolas Recio, Christopher Roebuck, and Joanna Zhang for excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada. The data for this paper was accessed between July 2022 and January 2023 through CIQ (2022), SEC (2022), and Twitter (2022).

Abstract

We study the distribution of political speech across U.S. firms. We develop a measure of political engagement based on firms' communications (earning calls, regulatory filings, and social media) by training a large language model to identify statements that contain political opinions. Using these data, we document five facts about firms' political engagement: (1) Political engagement is rare among firms; (2) Political engagement is concentrated among large firms; (3) Firms tend to specialize in specific topics and outlets; (4) Large firms tend to engage in a wider set of topics and outlets; (5) The 2020 surge in firms' political engagement was associated with an increase in the engagement of medium-sized firms and a change in the mix of political topics.

Topics: Firm dynamics; Market structure and pricing; Recent economic and financial developments
JEL codes: D22, D63, G41, L11, L20

Résumé

Nous nous penchons sur la distribution du discours politique parmi les entreprises américaines. Pour les besoins de l'exercice, nous concevons une mesure de l'engagement politique fondée sur leurs communications (téléconférences sur les résultats financiers, dépôts réglementaires et médias sociaux) en entraînant un grand modèle de langage à repérer les énoncés qui renferment des opinions politiques. Les données recueillies nous permettent de dégager cinq constats : 1) l'engagement politique est rare chez les entreprises; 2) ce sont surtout les grandes entreprises qui expriment des points de vue politiques; 3) les entreprises tendent à se positionner sur des sujets précis, dans des canaux précis; 4) l'éventail de sujets et de canaux est généralement plus large chez les grandes entreprises; 5) la forte poussée de l'engagement politique des entreprises en 2020 s'explique par une plus grande prise de position des moyennes entreprises et par un changement dans les enjeux politiques abordés.

Sujets : Dynamique des entreprises; Structure de marché et fixation des prix; Évolution économique et financière récente
Codes JEL : D22, D63, G41, L11, L20

1 Introduction

Firms have historically played an active role in the design of government policies in market economies. Traditionally, this role has been associated with efforts to influence regulations or industrial policies, often through lobbying and campaign contributions (e.g., Stigler, 1971; Grossman and Helpman, 1994; Bombardini and Trebbi, 2020, and references therein). Recently, however, firms appear to be increasingly using platforms such as earnings conference calls or social media to engage in political discussions on a broader range of topics, which extend beyond the regulations related to firms' businesses. At a time of high concentration in goods and labor markets (e.g., Philippon, 2019; De Loecker *et al.*, 2020; Manning, 2021; Kwon *et al.*, 2024), understanding how political engagement is distributed across firms becomes especially important.

Motivated by this concern, in this paper we study the distribution of political speech across U.S. firms. We begin by developing a measure of political engagement based on firms' communications. Our baseline measure focuses on earnings calls, during which managers of publicly listed firms hold webcasts or teleconferences to discuss financial results with investors, analysts, and other market participants. The nature of discussions in earnings calls often facilitates more candid and open exchanges (Hassan *et al.*, 2024), making them particularly suitable for studying political engagement. We complement this analysis with the more formal communication found in 10-K filings and the more informal communication found in social media posts on Twitter.

To construct our measure of political engagement, we proceed in two steps. In the first step, we identify statements that potentially contain political opinions. Our universe of political topics draws on a set of political issues that the Pew Research Center has been tracking since 1997, which we use to identify candidate political statements containing related keywords. In the second step, we train a large language model to identify which of these candidate statements actually contain political opinions. Our data show that the average frequency of political engagement throughout the sample is relatively low, with approximately 4 percent of firms making a political statement. Over time, we also observe a gradual increase in political engagement starting in 2016, followed by a sharp surge in 2020, consistent with other studies measuring firms' political engagement (Barari, 2024; Cassidy and Kempf, 2024).

Having constructed our new data set, we turn to our main goal of characterizing the distribution of political engagement across firms, through 5 new facts. The first fact is that political engagement is rare: More than half of the firms do not issue any political statements during the 15 years of our sample, while less than 1 percent of firms engage in political speech in more than 20 percent of their earnings calls. Even among firms that engage politically, the persistence of engagement is low.

The second fact is that political engagement is concentrated among large firms. We first show that the distribution of political engagement is highly concentrated. 10 percent of the firms account for approximately 40 percent of the total instances of political engagement of firms over this period—a higher degree of concentration than observed in the distribution of sales. We then document that political engagement is more prevalent among large firms (whether measured by log assets, sales, or employment). In particular, a one standard deviation increase in firm size is associated with about a 0.8 percentage point increase in the probability of political participation (or 20 percent of the mean of political participation).

Third, firms tend to specialize in specific topics and use specific outlets. In terms of topics, roughly 70 percent of firms engage in a single topic, and only 11 percent of firms engage in three or more topics. In terms of outlets, the majority of firms that engage politically (68 percent) do so using only one outlet (earning calls, 10-Ks, or Twitter). Only 5 percent of firms use all three outlets.

Fourth, large firms tend to engage in a broader set of topics and outlets. Among firms that engage politically, the smallest 10 percent of firms engage on average in 1.4 topics, while the largest 10 percent of firms engage in 1.8 topics. The corresponding averages for outlets used are 1.3 and 1.7 outlets, respectively.

Fifth, the 2020 surge in political engagement was associated with an increase in the engagement of medium-sized firms and a change in the mix of political topics. We show that the top 10 percent of firms by size account for about 22 percent of all political engagement until 2020, but their share declines to about 17 percent afterwards. This decline is accounted for by the rise in participation of firms between the 10th and the 50th percentile of size. In terms of topics, the surge of political engagement is explained, in large part, by an increasing number of firms that express views about environment, race relations, health policy, and criminal justice.

Related literature. Our paper is related to four strands of literature. First, we contribute to the long-standing debate on the role of firms in society. A traditional view, often attributed to Friedman (Friedman, 1962) holds that the role of firms is to maximize profits for its shareholders. Especially of late, this view has been challenged by those who maintain that firms should embrace a broader role, integrating social and political issues into their objectives (e.g., Broccardo *et al.*, 2022; Hart and Zingales, 2022) and aligning corporate actions with environmental, social and governance (ESG) goals (see Gillan *et al.*, 2021, for a review on ESG research in corporate finance), a view supported by recent empirical evidence suggesting that political polarization affects financial decisions (see Kempf and Tsoutsoura, 2024, for a review). Our contribution to this discussion is to highlight that large firms, regardless of the reason, also engage with the public through free-form communications in earnings calls, a form of communication hitherto unexplored in academic research in this context.

Second, we also relate to the emerging literature that analyzes corporate political speech. Barari (2024) documents that social media postings from the most recognized consumer brands in the U.S. feature an increasing resemblance to Democrat speech, which aligns with the political preferences of firms' key stakeholders. Conway and Boxell (2024) show that consumers increase their purchases from firms that express social stances consistent with their values. Closer to our work, and independently, Cassidy and Kempf (2024) develop a measure of firms' political engagement to study the partisan speech observed in tweets among S&P 500 firms, showing a sharp increase in Democrat-resembling speech starting in 2019. We contribute to this literature by focusing on the distribution of firms' political engagement, in terms of frequency, topics, and outlets. By analyzing the universe of publicly traded firms with available communications, we highlight the large degree of concentration of speech across firms and the role played by firms that are large in their own markets.

Third, we contribute to the literature studying the role of the distribution of firms in domestic and international markets. Part of this literature focuses on granularity and the importance of large firms. Gabaix (2011) argues that firm-level shocks translate into aggregate fluctuations. Di Giovanni and Levchenko (2012) show that trade openness, by making large firms grow larger, increases the importance of granular shocks. Gaubert and Itskhoki (2021) show that idiosyncratic firm shocks can shape aggregate comparative advantage. Our contribution to this literature is to characterize a new dimension of concentration among firms—political engagement—while emphasizing how political engagement relates to firm size.

Fourth, we contribute to a large and expanding literature that uses textual data to provide new measurements of firms' behavior and characteristics, including financial conditions (Loughran and McDonald, 2011), monetary policy communication (Hansen *et al.*, 2018), political risks (Hassan *et al.*, 2019), among many others (see Gentzkow *et al.*, 2019; Ash and Hansen, 2023, for comprehensive reviews). Our contribution is to develop a novel method to measure political engagement in different outlets of firm communication.

2 Data and Measurement

2.1 Data

We measure the prevalence of U.S. firms' political speech across three outlets of communication: earnings conference calls, regulatory filings, and social media posts. Our main analysis uses transcripts of earnings conference calls (hereafter "earnings calls"), in which managers of publicly listed firms hold webcasts or teleconferences to discuss financial results with investors, analysts, and other market participants. These calls typically occur following the release of regulatory disclosures (such as annual 10-K filings or quarterly 10-Q filings) and consist of two segments. The first segment contains a presentation by firm managers, during which they discuss the latest financials and disclose relevant information at their discretion. It is followed by a question-and-answer session, during which investors and analysts have the opportunity to pose questions for firm managers to address. Even though earnings calls are not mandatory, they have become common practice adopted by most firms after the implementation of Regulation FD (Fair Disclosure) by the U.S. Securities and Exchange Commission (SEC) in 2000.[1] Our sample consists of 283,920 earnings calls for 13,472 unique firms between 2008 and 2022, obtained from Capital IQ Transcripts (CIQ, 2022).

The second form of firm communication we study is their regulatory filings. Under the Securities Exchange Act of 1934, publicly traded firms in the U.S. are required to file Form 10-K with the SEC annually to disclose audited financial statements and provide comprehensive overviews of firm business conditions. We collect all electronically available Form 10-Ks filed by publicly traded firms in the U.S. between 1996 and 2022.[2] While 10-K filings vary in length, they are organized into standardized sections. We focus on the sections with greatest variation in language and therefore least susceptible to boilerplate legal statements and financial reporting: Item 1, which provides an overview of the firm's business; Item 7, which requires firm managers to discuss the firm's financial condition and results of operations; and Items 1A and 7A, which require firm managers to disclose general risk factors and market risk, respectively. Our sample consists of 83,674 filings from 14,707 unique firms, obtained from EDGAR (SEC, 2022).

The third form of firm communication we consider in our analysis is posts made by publicly traded firms on the social media platform Twitter (now X), which facilitates the dissemination of short messages, or "tweets," limited to 280 characters. We identify 3,110 publicly traded firms in the U.S. that have Twitter accounts and collect the complete set of tweets posted by these firms between 2014 and 2022 (Twitter, 2022).[3]

The three forms of firm communication are complementary in the analysis, as they differ markedly in formality, content, and targeted audience. 10-K filings are regulatory documents designed to disclose comprehensive information to shareholders and regulators, featuring for-

[1] According to the 2014 National Investor Relations Institute survey, 97 percent of publicly traded firms in the U.S. hold earnings calls (see Hassan *et al.*, 2019, for a more detailed discussion). Firms tend to restrict access to earnings calls to invited participants. However, the transcripts and audio recordings of earnings calls are stored and made publicly available on firms' websites, compliant with Regulation FD, which requires that all material nonpublic information disclosed to certain individuals or entities must also be disclosed publicly.

[2] The Electronic Data Gathering, Analysis, and Retrieval (EDGAR) system was developed by the SEC in 1993. Firms were phased into electronic filing between 1993 and 1996. We begin our sample in 1996, when our sample covers the universe of filers (see Song and Stern, 2020, for a discussion).

[3] For firms with multiple Twitter accounts, we use the firm's corporate account and exclude accounts related to customer service or internal communication. We start the sample in 2014, when Twitter gained popularity, and end the sample in 2022, before its ownership change.

mal and legal language. Earnings calls typically accompany the release of 10-K and 10-Q filings. While firm managers prepare scripted presentations for these calls, the subsequent question-and-answer sessions with investors and analysts elicit more spontaneous interactions. Social media posts, particularly on platforms like Twitter, allow firms to engage informally with investors, customers, and employees. Although these posts offer greater flexibility in content and timing, the short format constrains the depth and breadth of information that can be conveyed.

2.2 Methodology

We develop a method for identifying political engagement and apply it across all three forms of firm communication. Our methodology consists of two steps. In the first step, we identify statements from firm communication that potentially contain political engagement. In the second step, we classify whether each candidate statement actually contains engagement, using a large language model fine-tuned for this purpose.

2.2.1 Step 1: Identifying candidate statements on political engagement

The three types of communication in our data feature distinct vocabularies: written and formal in SEC filings, verbal and interactive in earnings calls, and casual and colloquial in tweets. Our methodology automatically detects the vocabulary firms use to discuss political issues, allowing us to identify political discussions across the three corpora. This approach closely relates to the methodology developed by Bloom *et al.* (2021), who use word2vec to discover technology-related vocabulary and study the diffusion of disruptive technology.[4]

To minimize discretionary inputs from researchers, we use the political issues from the American Trends Panel by the Pew Research Center, which has surveyed the political priorities of Americans since 1997. Table 1 lists the set of political issues that appear in the survey, including recurring issues that have appeared for at least 10 years, as well as topical issues added to the survey as they come into the spotlight. We remove economy-related issues to focus on purely political topics. For each political issue, we specify a core set of keywords in

[4] A proof-of-concept of this step was conducted using NL Analytics, a text analytics tool.

Table A.1, i.e., the "seed words." Since the list of seed words does rely on human inputs, we restrict it to a minimal set that is most relevant to each political topic.

TABLE 1
PEW POLITICAL ISSUES

Recurring issues			Topical issues	
Crime	Immigration	~~Budget deficit~~	Abortion	LGBTQ
Drug policy	Military	~~Economy~~	Criminal justice	Political system
Education	Race relations	~~Global trade~~	Free speech	Poor and needy
Environment	Social security	~~Jobs~~	Gun policy	Religion
Health policy	Terrorism			

The next step automatically detects different vocabularies firms use for political engagement across the three forms of communication. To do so, we train an embedding vector algorithm (word2vec, developed by Mikolov *et al.*, 2013) separately for each firm communication corpus. The algorithm trains a neural network to represent each word as a vector (i.e., embedding). It uses word ordering in addition to word frequencies to represent the joint distribution between words and therefore informs how words are related in domain-specific contexts. Our training sample for each corpus includes all 3-sentence snippets containing the seed words.[5] After the model is trained, we extract 20 words that are most related to the seed word by cosine similarity (e.g., "George Floyd" is identified by the trained word2vec model as being highly related to "systemic racism"). This procedure detects vocabulary that firm managers commonly use for political discussions in each form of communication, expanding our set of political keywords and minimizing false negatives.

Lastly, we make two adjustments to the expanded keyword set to minimize false positives. First, we use only bigrams and trigrams to ensure that the keywords are specific to the intended context. Second, we perform an exhaustive audit of the expanded keyword set to remove keywords unrelated to politics or those with ambiguous meanings (e.g., "carbon dioxide" is identified as related to climate change but is often used in contexts unrelated to it).

[5]We perform basic preprocessing to remove cases, punctuation, and stop words, but keep numbers because they remain informative for political discussions (e.g., "CO2 emissions"). We use the skip-gram implementation of word2vec and include unigrams, bigrams, and trigrams in the vocabulary.

2.2.2 Step 2: Using BERT to classify firm political engagement

The classification challenge. After identifying political discussions in firm communication, we classify whether a statement contains political engagement. Most statements identified through a keyword search do not involve political engagement but rather focus on firm business and political risks. For instance, both statements below contain the keyword "renewable energy" under the political topic of the environment, but only the second statement relates to political engagement:

> *Statement 1*: As I said at the beginning, sustained profitability is our overriding objective. We have built a platform that provides a broad array of energy efficiency and renewable energy solutions for a diversified base of clients to deliver sustainable profitability. The power of this platform is now becoming evident in our financials.
>
> *Statement 2*: Some heavily funded investors from California were driving this to try to get what they think is important for this state, and that's the constitutional amendment for renewable energy. So it tends to be—we like to refer to it as hedge funds in California who want to capture our Constitution and turn it into renewable energy for their benefit. So they haven't destroyed the California economy enough?

The key challenge in identifying political engagement is that determining whether a statement contains political engagement depends on the context in which keywords are used, rather than on the occurrence of keywords alone. To capture political contexts, we turn to transformer-based BERT, or Bidirectional Encoder Representations from Transformers. BERT is a large language model developed by Google and trained on thousands of books and Wikipedia pages (Devlin *et al.*, 2018). Due to the large amount of training data and its attention mechanism, it only requires a small fine-tuning sample to achieve high accuracy in downstream tasks, such as classification in our application.[6]

[6]Specifically, we use the uncased and base version of BERT, which contains 12 layers and 110 million parameters. We compare the performance of base BERT with two models that are pre-trained on specific corpora: SEC-BERT, pre-trained on SEC filings (Loukas *et al.*, 2022), and TwHIN-BERT, pre-trained on tweets (Zhang *et al.*, 2023). We obtain the model weights for all versions of BERT from Hugging Face.

Annotating the training sample. To form the training sample for fine-tuning BERT, we draw 3,500 statements from those identified through a keyword search, ensuring representation across industries and political issues. We employ a team of research assistants to annotate whether a statement contains political opinions. For each statement, we provide the annotator with the industry of the firm along with the following criteria for classifying whether the statement contains political opinions: "In this statement, a firm is expressing a statement about current political or social events." Each statement is annotated by two members of the research team. In cases of disagreement, a third member reads the statement to break the tie.

Out of 3,268 statements for which we received valid annotations, 130 were classified as political opinions. To form the training sample, we balanced the representation of statements, randomly drawing from those that do not contain a political opinion to match the number of statements that do. The rare nature of political speech in earnings calls poses a problem of small training samples, discussed in Abowd *et al.* (2021) in the context of linking survey and administrative data. We draw the training sample by political topics to ensure sufficient training in each topic, and verify that the training sample is representative in terms of firm size.[7]

Training BERT. Having constructed the training sample, we now use it to fine-tune BERT for identifying political engagement. For the fine-tuning, we need to specify hyperparameters that determine the rate of learning and how the training sample is read in. To avoid overfitting, we follow the procedure in Hansen *et al.* (2023) for hyperparameter selection. First, we set aside 15 percent of the human-labeled sample as the holdout test sample, which is not seen by the model during the hyperparameter-selection step and will be used to evaluate the performance of BERT compared with other language models.

Using the remaining 85 percent of the sample as the training sample, we perform a grid search over combinations of the hyperparameters: learning rates $\beta_{\text{lr}} \in \{2 \times 10^{-5}, 3 \times 10^{-5}, 5 \times 10^{-5}\}$, epochs $\beta_{\text{e}} \in \{10, 15, 20\}$, and batch sizes $\beta_{\text{bs}} \in \{16, 32\}$. We use a 5-fold cross-validation and select the hyperparameters that yield the highest average F1 score

[7]The average size (and standard deviation) of firms with political speech in the training sample is 16,991 (75,644) million dollars in total assets, and that of firms with nonpolitical speech is 17,206 (38,833).

across the 5 training splits. The resulting hyperparameters are $\beta_{\text{lr}} = 5 \times 10^{-5}$, $\beta_{\text{e}} = 10$, and $\beta_{\text{bs}} = 32$.

Model evaluation. We use the resulting model to classify statements in the holdout test sample. To assess the accuracy of the model, we compare it with a wide range of alternative language models:

1. *Dictionary*: We use the dictionary of political keywords constructed in Section 2.2.1. By construction, all statements identified through the keyword search in Section 2.2.1 are classified as positive under the dictionary model.

2. *SEC-BERT*: Rather than using BERT-BASE pre-trained on generic English language texts, we use SEC-BERT pre-trained on SEC filings (Loukas *et al.*, 2022). We fine-tune SEC-BERT with the same training sample used in baseline BERT. This allows us to assess the importance of pre-training the model to the specialized language of financial documents.

3. *All zeros*: Instead of using our human-labeled training sample, we fine-tune BERT with a training sample where all statements are classified as negative, corresponding to the median outcome. This allows us to assess the importance of the training sample.

4. *GPT-4*: GPT-4 is a generative large language model developed by OpenAI, trained on a large corpus of internet and digitized text (Achiam *et al.*, 2023). Under a zero-shot learning setting, we prompt GPT-4 with "Classify whether the snippet from a company's earnings call is a political or social statement. Return only 'yes' or 'no.' If the text is a political statement, then return 'yes.' If the text is not a political statement, then return 'no.'"

Table A.2 reports the performance of each model. The test sample consists of 377 statements, 7 percent of which contain political opinions. Row 1 reports that the baseline BERT model has an accuracy of 86 percent and F1 score of 0.89. Row 2 shows that a simple dictionary-based method performs poorly in classifying political engagement. Row 3 shows that SEC-BERT, a version of BERT pre-trained on the specialized language of SEC filings does not lead to a performance improvement and slightly underperforms the baseline BERT. Row 4 demonstrates the importance of the human-labeled training sample for fine tuning BERT. The accuracy of correctly classifying a true positive increases by 83 percentage points with our training sample compared to that with an all-zero training sample. Finally, Row 4 compares BERT with GPT-4, a larger model with magnitudes more parameters. BERT substantially outperforms GPT-4 in identifying true political speech, with an accuracy of 89 percent compared to that of 17 percent for GPT-4; in contrast, GPT-4 outperforms BERT in identifying true nonpolitical speech, with an accuracy of 99 percent compared to that of 85 percent for BERT. Even though GPT-4 provides an overall performance gain, we use BERT as our baseline model because the open-source transformers infrastructure ensures transparency and reproducibility (Dell, 2024). Nevertheless, we construct an alternative measure of political engagement using GPT-4 as robustness and find consistent patterns.

Using BERT to classify political engagement. The final model we use to classify political engagement is re-estimated on the entire human-labeled sample, which includes both the training and test data. We use this model to classify the candidate statements that the keyword search from Section 2.2.1 identifies as potentially political.[8]

The output from BERT is the probability distribution over each class label (i.e., political or nonpolitical). The higher the probability, the more confident the model is that a statement contains political opinions. To ensure the accuracy of the measure, we classify a statement as political engagement if the probability is above a 0.95 threshold. If the main topic discussed in the statement coincides with a firm's business description in Compustat, we classify the statement to be nonpolitical engagement, in order to differentiate our measure of political engagement from measures of political risks.

[8]We fine-tune two separate BERT models: one fine-tuned on the earnings-call training sample and used to classify statements from earnings calls and 10-K filings, and another fine-tuned on the tweet training sample and used to classify tweets. We fine-tune BERT on tweets following the same procedure described in Section 2.2.2: We draw a random sample of 1,500 tweets from candidate political tweets, annotate the training sample, and use it to fine-tune BERT. 515 tweets from the training sample are labeled as political speech, and the hyperparameters used for fine-tuning are $\beta_{\mathrm{lr}} = 5e-5$, $\beta_{\mathrm{e}} = 15$, and $\beta_{\mathrm{bs}} = 16$.

2.3 Aggregate patterns

Before characterizing the distribution of political engagement across firms, which is the main goal of our paper, we first use our new data set to document three aggregate patterns, which we reference later. Panel (A) of Figure 1 reports the average frequency of political engagement, measured by the share of firms that made political statements in their earnings calls between 2008 and 2022. Before 2016, the average frequency was relatively stable, hovering around 3 percent. Starting in 2016, the frequency of engagement gradually increased. In the summer of 2020—coinciding with the George Floyd protests and the COVID-19 pandemic and lockdowns—the frequency of engagement increased sharply, reaching an average of 9 percent. While it fell quickly thereafter, participation remained higher than during the earlier part of the sample, stabilizing at about 5 percent.

Panels (B) and (C) of Figure 1 document the large variation in the industries to which participating firms belong and in the topics that firms engage. Regarding industries, those with the highest average frequency of engagement are agriculture, health care, construction, and administrative support, while those with the lowest average frequency are arts and entertainment, accommodation and food, and other services. Regarding topics, those with the highest average frequency of engagement are "environment," "military," "social security," and "poor and needy."

3 Political Engagement across the Firm Distribution

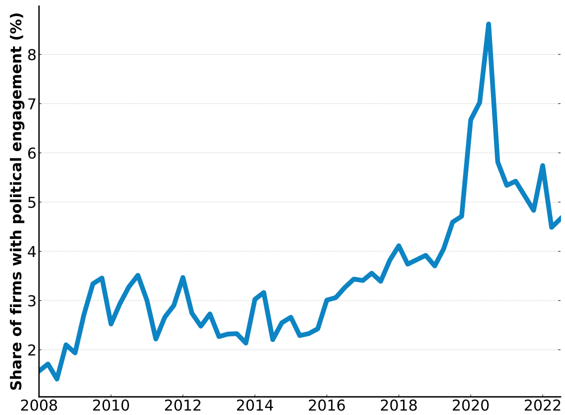
In this section, we characterize the distribution of political engagement across firms throughout our entire sample. Section 3.1 focuses on the frequency of engagement, while Section 3.2 studies the topics and outlets of engagement.

3.1 The frequency of political engagement

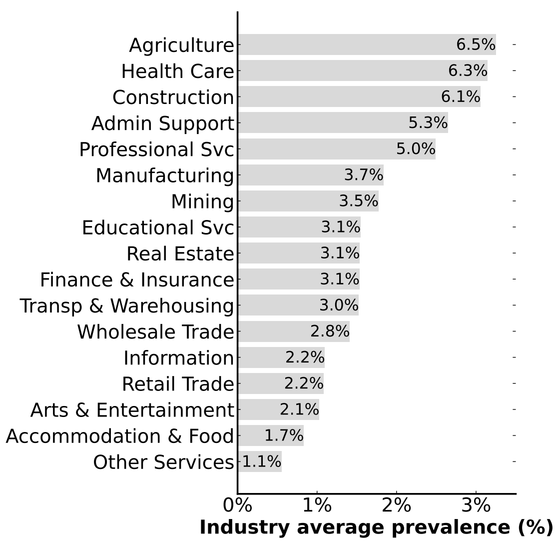
We begin by documenting two facts about the distribution of political engagement frequency: it is rare and concentrated among large firms. We describe each of these in the following sections.

Figure 1: The Average Frequency of Political Engagement: Evolution, Industries, and Topics

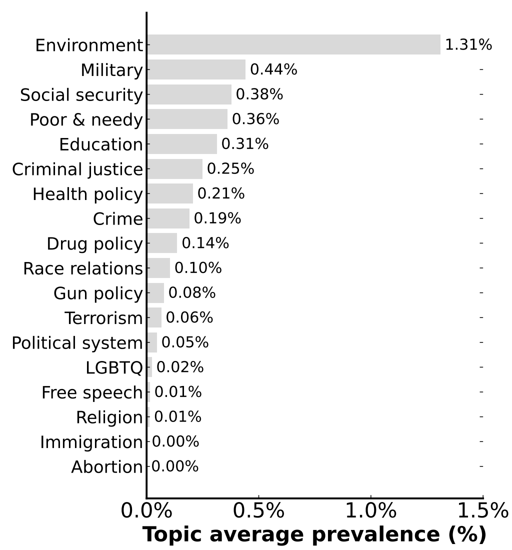
(A) Evolution



(B) By Industry



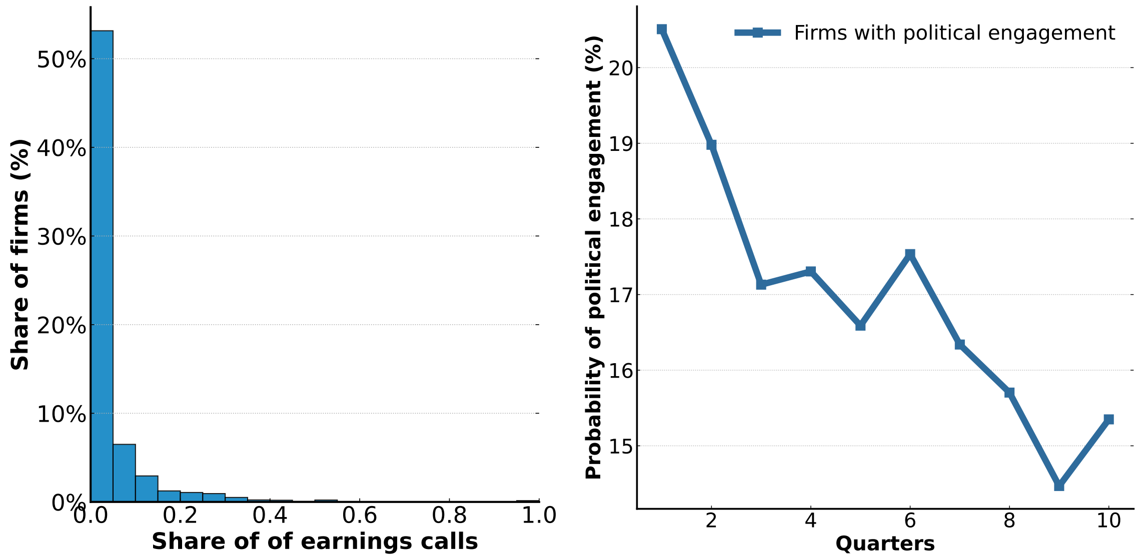
(C) By Topic



Notes: Panel (A) reports the share of firms with political engagement over the sample period 2008–2022. Panels (B) and (C) report the share of firms with political engagement in each industry and discussing each political topic, respectively.

Figure 2: The Distribution of Engagement Frequency

(A) Unconditional Distribution (B) Hazard Functions



Notes: Panel (A) is a histogram of the average fraction of earnings calls that contain political engagement over the sample period of 2008–2022. Panel (B) reports the probability of subsequent earnings calls containing political engagement for firms that have engaged politically with no missing earnings calls in the sample.

Fact 1: Political engagement is rare among firms.

Panel (A) of Figure 2 reports the distribution of the frequency of engagement. For each firm we compute the share of earnings calls in which the firm engages politically and then display the distribution of those shares. A key feature that stands out is that political engagement is rare. More than half of the firms do not issue any political statements during the 15 years of our sample, while less than 1 percent of firms engage in political speech in more than 20 percent of their earnings calls.

Complementing the unconditional distribution, Panel (B) of Figure 2 depicts the hazard function of political engagement, after having issued a political statement. The figure shows that the persistence of political engagement is low. Only 20 percent of firms that have engaged in a given quarter engage again in the following quarter. The share falls steadily, reaching 15 percent 10 quarters after having engaged.

Fact 2: Political engagement is concentrated among large firms.

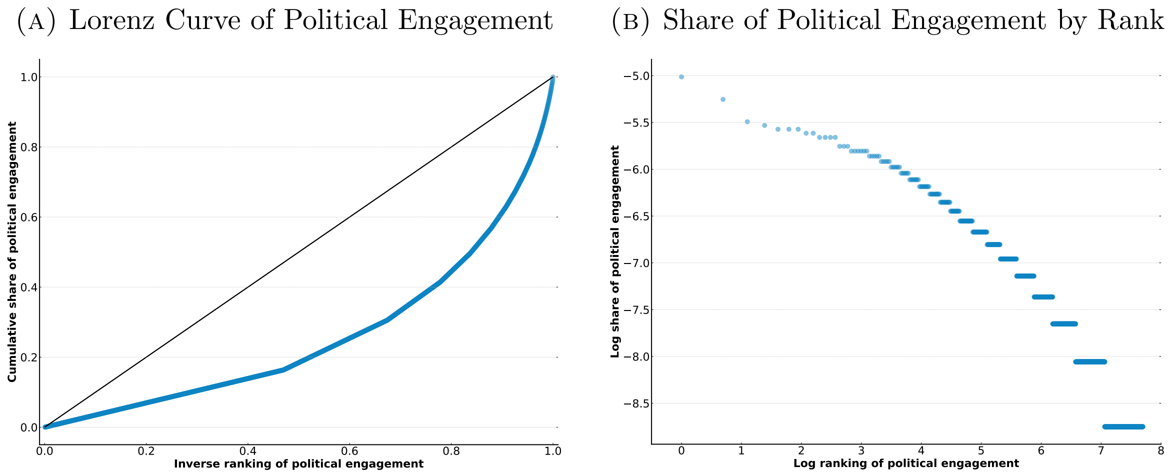
A counterpart to the infrequency of political engagement is its concentration. To study the concentration of political engagement, we define the total amount of participation as the number of quarter-firms that engage politically across all periods and firms. We then compute each firm's engagement share as the number of its earnings calls with political engagement out of this total. Panel (A) of Figure 3 shows the Lorenz curve corresponding to these firm-level shares and indicates that the most engaged firms account for a disproportionate share of total engagement. For instance, the top 10 percent of firms, ranked by engagement shares, account for 40 percent of total engagement. Panel (B) of Figure 3 reports the scatter plot of the rank-size relation. Following the firm granularity literature, we use this relation to study the top of the distribution. Although the distribution is not Pareto, it does exhibit a fat right tail: The slope for the top 1 percent of firms is -0.2.[9] As another metric of concentration, consider the collective share of mentions of the top 50 most engaged firms, which is 15 percent, while that of the top 100 most engaged firms is 24 percent.[10] We conclude that political engagement is highly concentrated among firms, and to an extent comparable, at least, to concentration in output markets.

Given the high concentration of political engagement, it is natural to ask what types of firms participate the most. We now document that political engagement is concentrated among large firms. Figure 4 presents binned scatter plots depicting the relationship between political engagement and firm size, measured by log real assets. Panel (A) presents the relationship using the raw data, pooling across firms and quarters. It indicates a positive association between political engagement and firm size. Panel (B) demeans these variables at the 4-digit NAICS sector level (equivalent to using a sector fixed effect in a regression context) and confirms that the relation between political engagement and size also holds within sectors. This relationship is quantitatively significant, showing that the share of firms that engage politically ranges from less than 3.5 percent for firms with the lowest

[9]The coefficients are -0.32 and -0.4 when we run the regressions using the top 5 and top 10 percent of firms. These coefficients imply substantially fatter tails than, for example, those in the sales distributions studied by Eaton *et al.* (2011)

[10]Gabaix (2011) meanwhile, shows that sales of the top 50 largest firms in the US are about 25 percent of GDP.

Figure 3: The Concentration of Political Engagement



Notes: In this figure, we rank a firm by its political engagement share, measured as the number of its earnings calls with political engagement as a share of total firm-quarter pairs containing political engagement among all firms in the sample period 2008–2022. Panel (A) plots the cumulative share of political engagement against firms inversely ranked by political engagement share. We invert the ranking in this panel to be consistent with the convention of Lorenz curves, so that the leftmost points correspond to least politically engaged firms. Panel (B) reports the scatter plot of a firm's log political engagement share against its log ranking of political engagement for all sample firms.

levels of real assets to 5.5 percent for the largest firms.

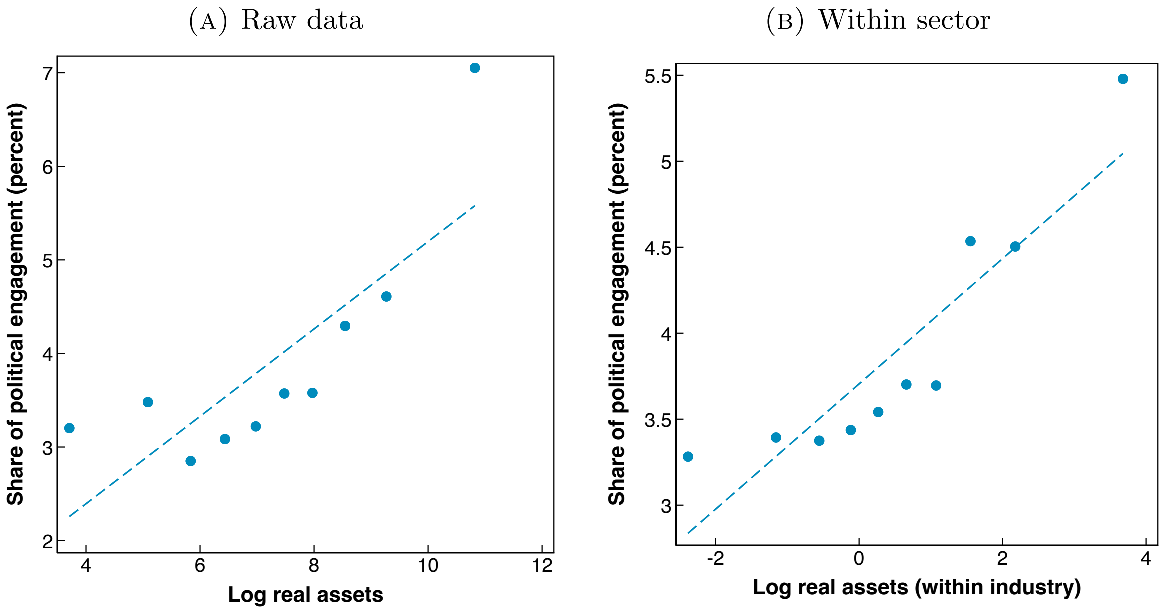
To document the systematic association between political engagement and firm size, we estimate variants of the following regression:

$$\text{engagement}_{it} = \alpha_s + \alpha_t + \beta \log \text{size}_{it} + \gamma X_{it} + \epsilon_{it}, \tag{1}$$

where engagement_{it} is a dummy variable that takes the value of 1 if firm i engages in period t and 0 otherwise; size_{it} is a measure of size, either real assets, real sales, or employment; X_{it} is a vector of firm-level controls; and α_s and α_t denote sector and time-fixed effects. We two-way cluster standard errors by firm and quarter. In the vector of firm-level controls, we include variables typically used in the corporate finance literature: firm age, leverage, and real sales growth. In addition we control for total firm lobbying (Kim, 2018, at the annual level), to control for a firm's general tendency to engage in politics. Appendix B provides a detailed definition of all variables used in the empirical analysis.

Table 2 reports the results from estimating equation (2) and indicates a strong and

Figure 4: Political Engagement and Firm Size



Notes: This figure reports binned scatter plots of the share of firms that engage politically against firm size. Each dot represent a decile of firm size, measured by log real assets. Panel (A) reports the relationship for all firms and quarters. Panel (B) reports the relationship between the share of firms that engage politically against firm size relative to industry average, measured by the residuals after regressing log real assets on 4-digit NAICS industry fixed effects.

robust association between political engagement and firm size. In particular, across all model specifications, a one standard deviation increase in firm size is associated with a 0.8–0.9 p.p. increase in the probability of political participation (or 20 percent of the mean of political participation), which is statistically significant at the 1 percent level. Among the other variables used as controls in (2), firm age, sales growth, and lobbying do not exhibit a statistically significant relationship with political participation, while leverage shows a negative relationship with political participation that is statistically significant at the 10 percent level.

3.2 The topics and outlets of political engagement

We now study the distribution of topics and outlets of political engagement. We document that firms tend to specialize in specific topics and outlets and that this specialization is less prevalent among large firms, who tend to engage in a wider range of topics and outlets. We

Table 2

Political engagement, firm size, and other firm-level variables

	Share of political engagement (percent)					
	(1)	(2)	(3)	(4)	(5)	(6)
Log real assets	0.841***	0.787***				
	(0.163)	(0.172)				
Log real sales			0.767***	0.759***		
			(0.159)	(0.167)		
Log employment					0.858***	0.802***
					(0.173)	(0.183)
Log age		-0.069		-0.066		-0.130
		(0.109)		(0.109)		(0.113)
Leverage		-0.226*		-0.216*		-0.216*
		(0.117)		(0.117)		(0.123)
Real sales growth		-0.000		-0.000		0.000
		(0.000)		(0.000)		(0.000)
Lobbying		0.103		0.119		0.125
		(0.086)		(0.089)		(0.091)
Observations	162080	144027	159655	143634	153578	136913
R^2	0.036	0.038	0.036	0.038	0.036	0.038
Industry FE	yes	yes	yes	yes	yes	yes
Quarter FE	yes	yes	yes	yes	yes	yes
Double-clustered SE	yes	yes	yes	yes	yes	yes

Notes: This table reports results from estimating variants of

$$\text{engagement}_{it} = \alpha_s + \alpha_t + \beta \log \text{size}_{it} + \gamma X_{it} + \epsilon_{it}, \tag{2}$$

where engagement_{it} is a dummy variable that takes the value of 1 if firm i engages in quarter t and zero otherwise; size_{it} is either real assets, real sales, or employment; X_{it} is a vector of firm-level controls, including firm age, leverage, real sales growth, and lobbying spending; and α_s and α_t denote sector and time fixed effects. All firm variables are standardized across the sample. Standard errors are clustered by firm and quarter. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

describe each of these facts in more detail next.

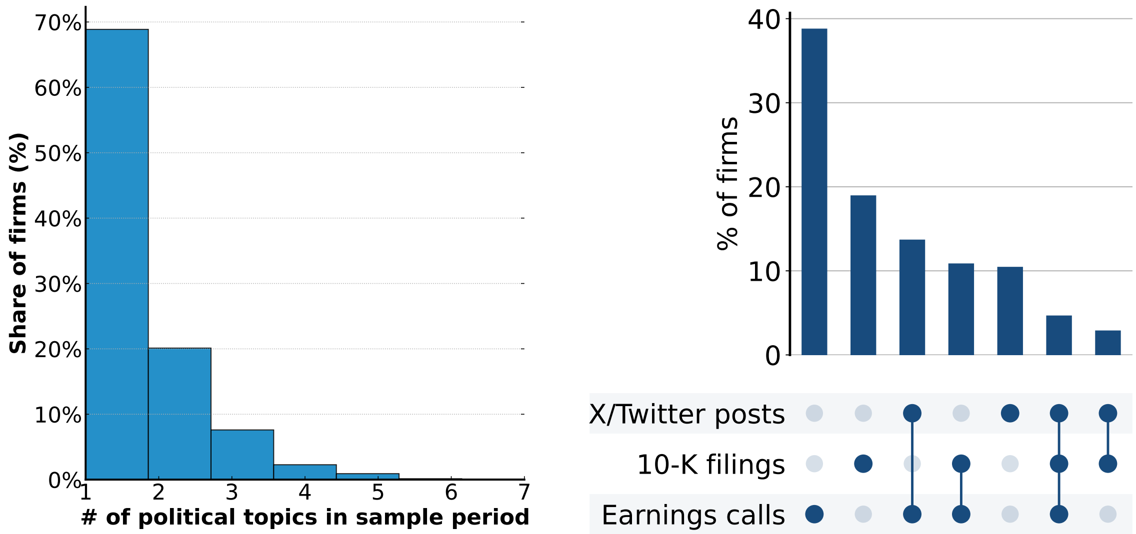
Fact 3: Firms tend to specialize in specific topics and outlets.

Panel (A) of Figure 5 reports the number of topics that firms engage in during our sample period. Roughly 70 percent of firms engage in a single topic, and only 11 percent of firms engage in three or more topics. In this sense, firms' political engagement is characterized by a specialization in terms of topics: When firms choose to engage, they tend to do so

Figure 5: Specialization in Topics and Outlets

(A) Topics

(B) Outlets

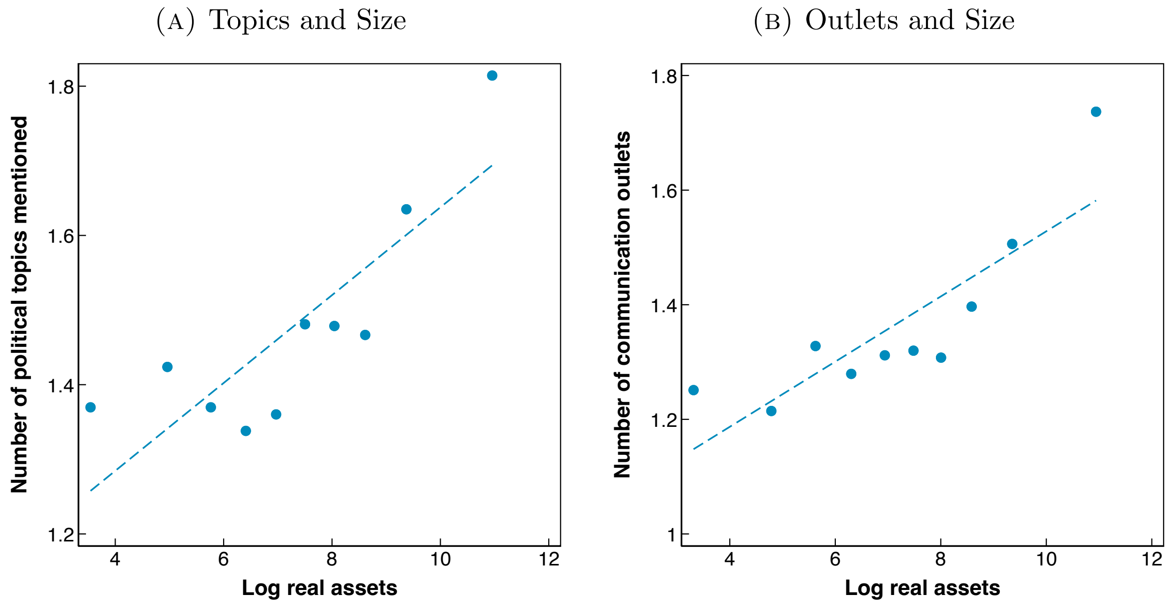


Notes: Panel (A) is a histogram of the total number of political topics that firms discuss. Panel (B) is a histogram of the combinations of communication outlets that firms use, ranked from the most used to the least used: earnings calls only, 10-Ks only, earnings calls and Twitter, earnings calls and 10-Ks, Twitter only, all three outlets, and 10-Ks and Twitter.

with a narrow set of topics. To complement this finding, Panel (A) of Appendix Figure A.1 reports the distribution of the average number of topics mentioned per earnings call featuring political engagement. The vast majority of these documents (more than 85 percent) involve only one topic. Only 6 percent of these documents involve two or more topics.

Panel (B) of Figure 5 shows that firms also specialize in the set of outlets they use for political engagement. In this analysis, we focus on the period from 2014 to 2022, during which, as discussed in Section 2, we have data available on three outlets of political engagement: earnings calls, 10-Ks, and Twitter. As indicated in the first three columns of Panel (B), the majority of firms that engage politically (68 percent) do so using only one outlet. Only 5 percent of firms use all three outlets. Complementing this finding, Panel (B) of Appendix Figure A.1 shows that this specialization in outlets is also present when we focus on firm-years as the unit of observation. Appendix Table A.3 shows that this specialization among outlets occurs despite the fact that the probability of engaging politically in earnings calls is associated with a higher probability of engaging in other outlets.

FIGURE 6: Topics, Outlets, and Firm Size



Notes: Panel (A) reports the binned scatter plot of the number of political topics mentioned by firms that have engaged at least once in the sample period against firm size. Panel (B) reports the binned scatter plot of the number of communication outlets used by firms that have engaged at least once in the sample period against firm size. In both panels, each dot represents a decile of firm size, measured by log real assets.

Fact 4: Large firms tend to engage in a wider set of topics and outlets.

Figure 6 documents how specialization across topics and outlets varies across the firm size distribution. Panel (A) shows that large firms tend to engage in more topics than small firms. We divide firms into size deciles, measured by log real assets, and report the average number of political topics mentioned by firms in each size decile. Among firms that engage politically, the smallest 10 percent of firms discuss an average of 1.4 topics, while the largest 10 percent of firms discuss an average of 1.8 topics. To further study the relationship between the topics of political engagement and firm size, we estimate a multinomial logit model in which we allow for firm size to have different effects on the likelihood of choosing each topic. Appendix Figure A.2 presents the results, showing that large firms tend to engage in the same topics as the average firm.

Panel (B) of Figure 6 shows that large firms tend to use more outlets for political engagement than small firms do. Among firms that engage politically, the smallest 10 percent of firms use, on average, 1.3 outlets, while the largest 10 percent of firms use 1.7 outlets.

To complement this analysis, Appendix Table A.3 estimates a linear probability model to understand the joint distribution of engagement across outlets. The results indicate that firms that engage in earnings calls are more likely to also engage in 10-Ks and tweets and that large firms are more likely to engage in both 10-Ks and tweets.

4 Accounting for the Surge of Political Engagement

So far, our analysis has concentrated on characterizing the distribution of political engagement across firms throughout our entire sample. We now focus on studying how the distribution of engagement changed during the surge of political engagement documented in Section 2.3, which occurred in the summer of 2020. We summarize our findings as follows:

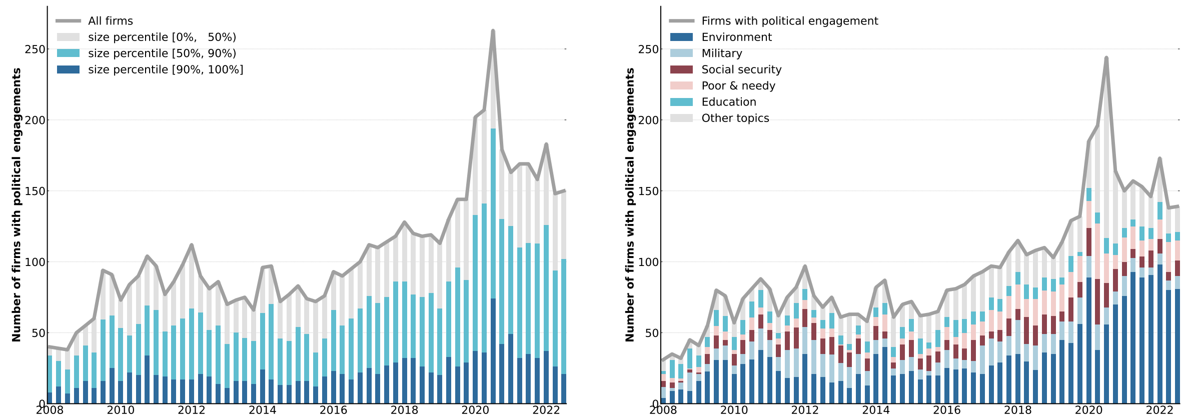
Fact 5: The 2020 surge in political engagement was associated with an increase in the engagement of medium-sized firms and a change in the mix of political topics.

Panel (A) of Figure 7 illustrates the role of firm size in the surge of political engagement. For this, we compute the share of the total number of firms engaging (solid line) that is accounted for by three groups of the firm-size distribution within quarters: firms (1) below the median, (ii) between the 50th and the 90th percentile, and (iii) above the 90th percentile. The figure shows that there has been a shift in the importance of smaller firms in accounting for total overall political engagement, especially in the wake of summer 2020. While the share of total engagement by the largest 10 percent of firms hovered around 22 percent until 2020, it declined to 17 percent by 2022. In contrast, the next 40 percent of the largest firms' share of political engagement has increased—from 40 percent in 2015, to 48 percent in 2020, and finally to 50 percent in 2022. Note that firms in all groups engaged more over time, but the engagement of middle-sized firms grew more, as shown in Appendix Table A.4.

Panel (B) shows the topics in which firms engage over time, plotting separately the top five topics (throughout our entire sample) and grouping the rest in a category called "other topics." Two key patterns emerge. First, over the whole sample, the composition of topics is relatively stable. The top five topics—"environment," "military," "social security,"

Figure 7: Accounting for the surge of political engagement

(A) Firm Size　　(B) Topics



Notes: Panel (A) reports the number of firms in each size bin that engage politically in each quarter. Size bins are calculated using firm size in each quarter. Panel (B) subsets to firms that engage politically and reports the number of firms that engage in a given political topic in each quarter.

"poor and needy," and "education"—account for the lion's share of engagement, with a participation rate usually higher than 90 percent. Since 2016, however, there has been a gradual increase in the participation of other topics. This increase was especially salient in 2020. As shown in Appendix Table A.5, among these other topics, the three main topics behind this growth are "race relations," "health policy," and "criminal justice."

5 Conclusions

We have documented new patterns of political engagement among U.S. firms and, in particular, demonstrated that the outsize role that large firms play in goods and labor markets is mirrored in the fact that they account for a large share of political speech. Such speech may reflect profit-maximizing firms trying to reach more consumers or it may reflect firms adopting a broad, "stakeholder" view of their operations. We do not take a stand on what the explanation is. While the speech patterns we document here are in line with the U.S. legal tradition of allowing firms to express political views, our findings also suggest additional caution in evaluating the role that large firms play in society.

References

Abowd, J. M., Abramowitz, J., Levenstein, M. C., McCue, K., Patki, D., Raghunathan, T., Rodgers, A. M., Shapiro, M. D., Wasi, N. and Zinsser, D. (2021). Finding needles in haystacks: Multiple-imputation record linkage using machine learning.

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S. *et al.* (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774*.

Ash, E. and Hansen, S. (2023). Text algorithms in economics. *Annual Review of Economics*, **15** (1), 659–688.

Barari, S. (2024). Political speech from corporate america: Sparse, mostly for democrats, and somewhat representative. *Journal of Quantitative Description: Digital Media*, **4**.

Bloom, N., Hassan, T. A., Kalyani, A., Lerner, J. and Tahoun, A. (2021). *The diffusion of disruptive technologies*. Tech. rep., National Bureau of Economic Research.

Bombardini, M. and Trebbi, F. (2020). Empirical models of lobbying. *Annual Review of Economics*, **12** (1), 391–413.

Broccardo, E., Hart, O. and Zingales, L. (2022). Exit versus voice. *Journal of Political Economy*, **130** (12), 3101–3145.

Cassidy, W. and Kempf, E. (2024). Partisan corporate speech. *Manuscript*.

CIQ (2022). Capital IQ Transcripts [dataset]. Accessed from July 1, 2022 to August 10, 2022.

Conway, J. and Boxell, L. (2024). Consuming values. *Manuscript*.

De Loecker, J., Eeckhout, J. and Unger, G. (2020). The Rise of Market Power and the Macroeconomic Implications ["Econometric Tools for Analyzing Market Outcomes"]. *The Quarterly Journal of Economics*, **135** (2), 561–644.

Dell, M. (2024). *Deep Learning for Economists*. Tech. rep., National Bureau of Economic Research.

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

Di Giovanni, J. and Levchenko, A. A. (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, **120** (6), 1083–1132.

Eaton, J., Kortum, S. and Kramarz, F. (2011). An Anatomy of International Trade: Evidence From French Firms. *Econometrica*, **79** (5), 1453–1498.

Friedman, M. (1962). Capitalism and freedom. University of Chicago Press.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, **79** (3), 733–772.

Gaubert, C. and Itskhoki, O. (2021). Granular Comparative Advantage. *Journal of Political Economy*, **129** (3), 871–939.

Gentzkow, M., Kelly, B. and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, **57** (3), 535–574.

Gillan, S. L., Koch, A. and Starks, L. T. (2021). Firms and social responsibility: A review of esg and csr research in corporate finance. *Journal of Corporate Finance*, **66**, 101889.

Grossman, G. M. and Helpman, E. (1994). Protection for sale. *The American Economic Review*, pp. 833–850.

Hansen, S., Lambert, P. J., Bloom, N., Davis, S. J., Sadun, R. and Taska, B. (2023). *Remote work across jobs, companies, and space*. Tech. rep., National Bureau of Economic Research.

—, McMahon, M. and Prat, A. (2018). Transparency and deliberation within the fomc: A computational linguistics approach. *The Quarterly Journal of Economics*, **133** (2), 801–870.

Hart, O. D. and Zingales, L. (2022). *The new corporate governance*. Tech. rep., National Bureau of Economic Research.

Hassan, T. A., Hollander, S., Kalyani, A., Van Lent, L. and Schwedeler, A., Markus Tahoun (2024). Economic surveillance using corporate text. *Manuscript*.

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

Kahle, K. M. and Stulz, R. M. (2017). Is the us public corporation in trouble? *Journal of Economic Perspectives*, **31** (3), 67–88.

Kempf, E. and Tsoutsoura, M. (2024). *Political Polarization and Finance*. Tech. rep., National Bureau of Economic Research.

Kim, I. S. (2018). Lobbyview: Firm-level lobbying & congressional bills database [dataset]. Http://web.mit.edu/insong/www/pdf/lobbyview.

Kwon, S. Y., Ma, Y. and Zimmermann, K. (2024). 100 years of rising corporate concentration. *American Economic Review*, **114** (7), 2111–2140.

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

Loukas, L., Fergadiotis, M., Chalkidis, I., Spyropoulou, E., Malakasiotis, P., Androutsopoulos, I. and George, P. (2022). FiNER: Financial numeric entity recognition for XBRL tagging. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, Association for Computational Linguistics.

Manning, A. (2021). Monopsony in labor markets: A review. *ILR Review*, **74** (1), 3–26.

Mikolov, T., Chen, K., Corrado, G. and Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Ottonello, P. and Winberry, T. (2020). Financial heterogeneity and the investment channel of monetary policy. *Econometrica*, **88** (6), 2473–2502.

Philippon, T. (2019). *The great reversal: How America gave up on free markets*. Harvard University Press.

SEC (2022). Electronic Data Gathering, Analysis, and Retrieval [dataset]. Accessed and collected from July 1, 2022 to August 13, 2023.

Song, W. and Stern, S. (2020). Firm inattention and the efficacy of monetary policy: A text-based approach. *Manuscript*.

Stigler, G. J. (1971). The theory of economic regulation. *The Bell Journal of Economics and Management Science*, pp. 3–21.

Twitter (2022). Tweets posted by publicly traded firms [dataset]. Accessed and collected from July 1, 2022 to January 31, 2023.

Zhang, X., Malkov, Y., Florez, O., Park, S., McWilliams, B., Han, J. and El-Kishky, A. (2023). Twhin-bert: A socially-enriched pre-trained language model for multilingual tweet representations at twitter. In *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*, pp. 5597–5607.

Appendix

A Additional Tables and Figures

Table A.1
Seed Words for Political Issues

Political issue	Seed words
abortion	abortion, reproductive rights
crime	crime prevention, juvenile crime, death penalty, death row, capital punishment, gun violence, violent crime
criminal justice	police, criminal justice, black lives matter, capitol riot, defund police
drug policy	marijuana legalization, drug addiction, drug overdose, vaping, opioid epidemic
education	education, student loan, student debt, title ix
environment	climate change, global warming, extreme weather
free speech	free speech, cancel culture, first amendment rights, offensive speech, censorship, misinformation, fake news
gun policy	gun, rifle, gun policy, second amendment, open carry, assault rifle, gun violence, background checks
health policy	health policy, mental health
immigration	immigrant, refugee, immigration policy, immigration enforcement, deportation, daca
lgbtq	lgbtq, diversity equity inclusion, gender identity, transgender, trans rights, gender affirming, pronouns, nonbinary
military	national defense, war iraq, war ukraine, veteran, afghan troops
political system	supreme court, separation church state, gerrymandering, democracy, voting rights, electoral college
poor and needy	safety net, universal basic income, homeless, homelessness, poor needy, economic inequalities, income inequality, low income americans
race relations	discrimination, prejudice, systemic racism, national belonging, racism, who black, who white, whiteness, white fragility, white supremacy
religion	religious, religious liberty, religious groups, christian nation, bible
social security	medicare, social security
terrorism	terrorism, terror attacks, terrorist, cyber warfare

Table A.2
Performance of BERT and alternative models

	NLP Model	F1 Score			Accuracy		
		Overall	Recall	Precision	Overall	Nonpolitical	Political
1	Baseline BERT	0.89	0.96	0.86	0.86	0.85	0.89
2	Dictionary	0.00	0.00	0.05	0.05	0.00	1.00
3	SEC-BERT	0.88	0.95	0.84	0.84	0.85	0.83
4	All zeros	0.92	0.90	0.95	0.95	1.00	0.00
5	GPT-4	0.94	0.93	0.95	0.95	0.99	0.17

Notes: This table reports the performance of language models in classifying whether a statement contains political engagement. Let TP, FP, TN, FN, and n denote true positives, false positives, true negatives, false negatives, and the total number of statements from the test sample based on model classification. Recall is computed as $\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$; precision is computed as $\mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$; the overall F1 score is computed as $2/(1/\mathrm{Precision}+1/\mathrm{Recall})$; and accuracy is computed as $(\mathrm{TP}+\mathrm{TN})/n$.

Table A.3
Distribution of engagement across outlets

	Engagement in 10-Ks		Engagement on Twitter	
	(1)	(2)	(3)	(4)
Engagement in earnings calls	0.018**	0.016**	0.025***	0.013***
	(0.007)	(0.007)	(0.005)	(0.005)
Size (log real assets)		0.002**		0.014***
		(0.001)		(0.001)
Age		-0.000***		0.001***
		(0.000)		(0.000)
Real sales growth		-0.000***		-0.000
		(0.000)		(0.000)
Return on equity		0.000***		-0.000*
		(0.000)		(0.000)
Observations	34142	28710	149562	125295
R^2	0.066	0.066	0.032	0.066
Double-clustered SE	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes
Quarter FE	yes	yes	yes	yes

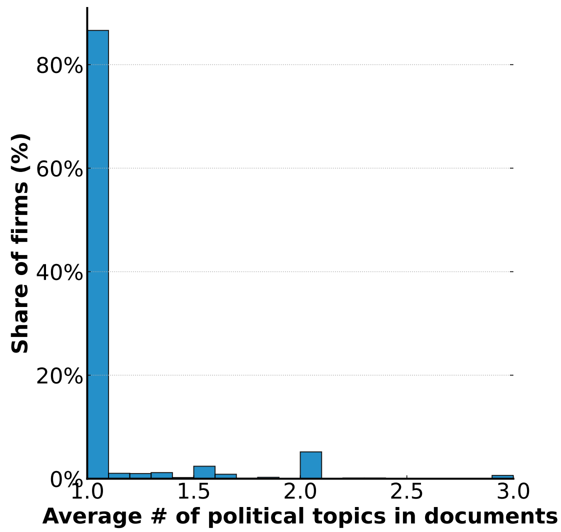
Notes: This table reports estimates from the linear probability model

$$\text{engagement}_{it}^{k} = \delta_s + \delta_t + \beta \cdot \text{engagement}_{it}^{\text{earnings}} + \Gamma' Z_{it} + \varepsilon_{it}, \text{ for } k \in \{\text{10-Ks, tweets}\},$$

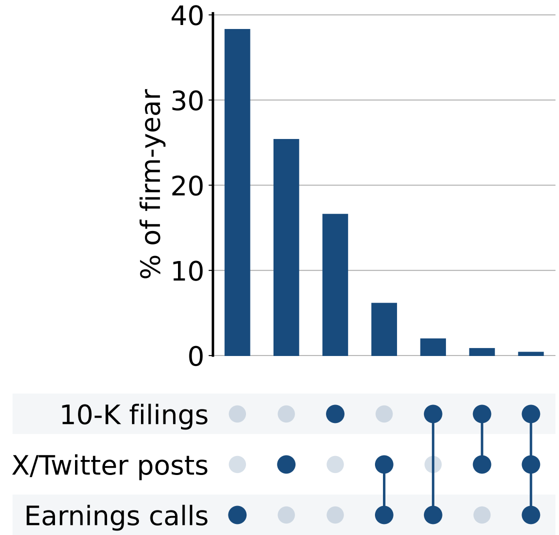
where $\text{engagement}_{it}^{k}$ is a binary variable that takes the value of 1 if firm i engages in political speech in period t through outlet $k \in \{\text{earnings calls, 10-Ks, tweets}\}$, $\{\delta_s, \delta_t\}$ are sector (4-digit NAICS) and quarter fixed effects, and Z_{it} is a vector of firm controls including size, age, real sales growth, and return on equity. Columns 1 and 2 are based on annual data from 2008 to 2022, and Columns 3 and 4 are based on quarterly data from 2014 to 2022. Standard errors are double clustered by firm and time. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

Figure A.1: Alternative Measures of the Distribution of Topics and Outlets in Political Engagement

(A) Average number of topics per firm document with political engagement



(B) Outlets used for political engagement per firm-year



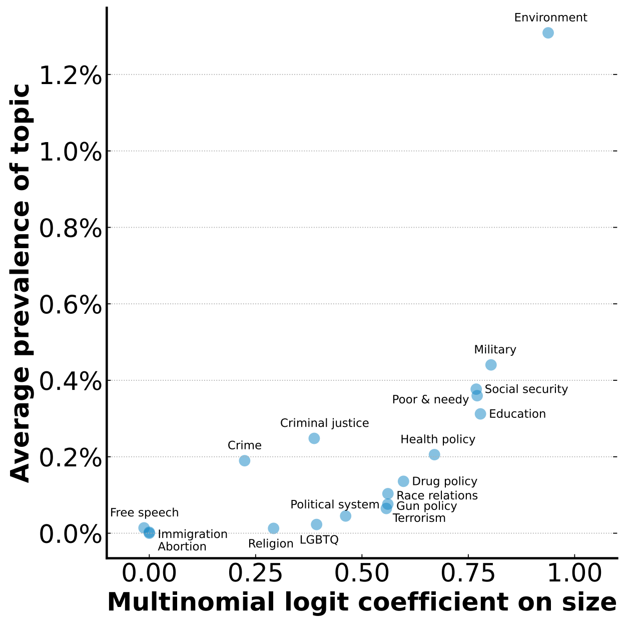
Notes: Panel (A) is a histogram of the average number of political topics that firms discuss over the sample period that is available for communication outlets (2014–2022). Panel (B) is a histogram of the combinations of communication outlets that firms use per document, ranked from the most used to the least used: earnings calls only, Twitter only, 10-Ks only, earnings calls and Twitter, earnings calls and 10-Ks, 10-Ks and Twitter, and all three outlets.

Table A.4
Contribution to political engagement growth 1997–99 to 2020–22, by size

Firm size percentile		Avg engagement (%)	Growth (p.p.)
Small	[0%, 50%)	1.31	0.37
		(11.36)	
Medium	[50%, 90%)	1.75	1.05
		(13.12)	
Large	[90%, 100%]	0.84	0.32
		(9.10)	

Notes: The first column reports the average share of firms that engage politically (as percent) for each size bin, with the standard deviation of firm engagement reported in parenthesis. The second column reports the political engagement growth for each size bin (in percentage points), defined as the difference between the average share of firms that engage politically in 2020–2022 and this share in 1997–1999.

Figure A.2: Size and the Probability of Engaging in Each Topic

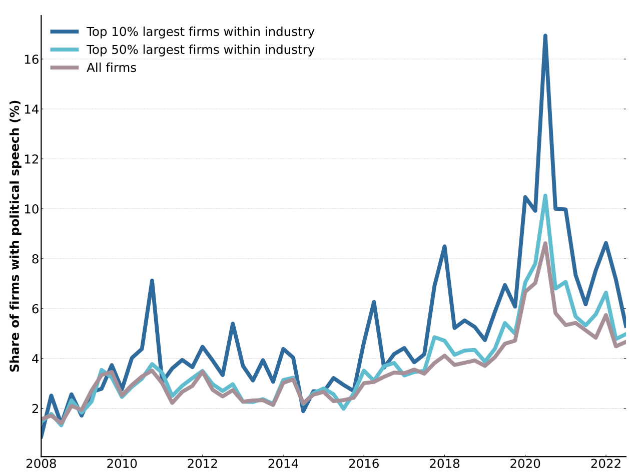


Notes: This figure reports the scatter plot of the average shares of firms that engage in a given political topic against the coefficients β for the topic estimated from the multinomial logit regression:

$$\log \frac{\Pr(\text{topic} = j)}{\Pr(\text{topic} = \text{immigration})} = \beta_j \text{size}_{it} + \varepsilon_{it}$$

for firm i in quarter q and topics j specified in Table A.5, excluding "immigration" as the reference topic and "abortion" with zero engagement.

Figure A.3: Increase of Political Engagement by Firm Size



Notes: This figure reports the share of firms with political engagement for all firms, firms whose size relative to industry average (4-digit NAICS) is in the top 50 percentile in each quarter, and firms whose size relative to industry average is in the top 10 percentile in each quarter.

Table A.5
Contribution to political engagement growth 1997–99 to 2020–22, by topic

Topic	Avg engagement (%)	Growth (p.p.)
1. Environment	1.79	1.46
	(13.25)	
2. Race relations	0.09	0.48
	(3.08)	
3. Health policy	0.19	0.17
	(4.34)	
4. Criminal justice	0.22	0.15
	(4.73)	
5. LGBTQ	0.02	0.07
	(1.42)	
6. Social security	0.36	0.05
	(6.01)	
7. Poor and needy	0.33	0.02
	(5.71)	
8. Immigration	0.00	0.00
	(0.56)	
9. Abortion	0.00	0.00
	(0.00)	
10. Education	0.27	-0.01
	(5.20)	
11. Religion	0.01	-0.01
	(1.09)	
12. Gun policy	0.07	-0.02
	(2.70)	
13. Crime	0.17	-0.02
	(4.14)	
14. Terrorism	0.07	-0.04
	(2.67)	
15. Free speech	0.01	-0.04
	(1.16)	
16. Political system	0.05	-0.07
	(2.33)	
17. Drug policy	0.12	-0.15
	(3.48)	
18. Military	0.39	-0.23
	(6.22)	

Notes: The first column reports the average share of firms that engage in a political topic (as percent), with the standard deviation of topic-specific firm engagement reported in parenthesis. The second column reports the growth in topic engagement (in percentage points), defined as the difference between the average share of firms that engage in a topic in 2020–2022 and in 1997–1999. Topics are ranked by the growth in topic engagement.

B Data Construction

This appendix provides details on the firm financial variables used in the empirical analysis, based on quarterly Compustat data. The definition of the variables follows that in Kahle and Stulz (2017) and Ottonello and Winberry (2020).

Variables

1. *Size*: the log of total real assets (`atq`), deflated using the BLS implicit price deflator.
2. *Real sales*: sales (`saleq`) deflated using the BLS implicit price deflator.
3. *Real sales growth*: log differences in real sales.
4. *Employment*: number of employees (`emp` from Compustat Annual).
5. *Age*: number of years since CRSP listing.
6. *Leverage*: the ratio of total debt (sum of `dlcq` and `dlttq`) to total assets (`atq`).
7. *Return on equity*: the ratio of income before extraordinary items (`ibq`) to market capitalization (`cshoq` times `prccq`).