Iterative Hand-Coding and Computational Text Analysis: Evidence from Public Comments in US Federal Agency Rulemaking

June 17, 2024

Abstract

This paper describes methods for integrating human coding of texts with computational text analysis and preprocessing to increase the inferential power of hand-coding by several orders of magnitude. I argue that human coding and computational text analysis are more powerful when combined in an iterative workflow. Text analysis tools can strategically select texts for human coders—texts representing larger samples and outlier texts of high inferential value. Preprocessing can speed up hand-coding by extracting features like names and key sentences. Humans and computers can iteratively tag entities using regex tables and group texts by key features (e.g., identify lobbying coalitions by common policy demands) Applying simple search and text-reuse methods to public comments on all U.S. federal agency rules, a sample of 13,000 hand-coded comments yields 41 million as-good-as-hand-coded comments regarding both the organizations that mobilized them and the extent to which policy changed in the direction they sought. This large sample of as-good-as-hand-coded documents enables new analyses of the relationships between lobbying coalitions, social movements, and policy change. For example, I find that public pressure to address climate change and environmental justice movements had large effects on policy documents, but a small number of national advocacy organizations dominate lobbying coalitions. When tribal governments or local groups lobby without the support of national advocacy groups, policymakers typically ignore them.

Keywords hand-coding, text analysis, data science, dynamic data, regex table

1 Introduction

The workflow I describe is designed for cases where we observe a large number of texts aimed at influencing policy. That is, it will be useful when researchers aim to answer the question: who demands what, how, and when? Characterizing policy demands is a first step in assessing the impact of political efforts. To assess the impact of both what people and groups demanded and who got what.

Because political advocacy is organized, multiple people and organizations may advance the same demands and supportive information. Researchers can leverage this fact to reduce the number of documents we read by targeting the most informative documents. If we know that 100 documents make the same claim, we don't need to read it 100 times. While I focus on computational methods that enable more efficient qualitative analysis, not outsourcing qualitative judgment about the text, I also offer preliminary results from comparing the results of human coding and Open AI's GPT in Appendix Section B.

Because a large number of texts almost certianly indicates one or more organized campaigns, three features are key: First: which group, organization, or campaign is each text associated with? Second, what does it ask policymakers to do? Finally, how did policymakers respond these demands? The workflow I describe helps answer these questions. I pay special attention to the first: linking individual texts to organizations and broader campaigns—because once accomplished, this allows us to use information collected about the policy demands and lobbying success in one text to make inferences about others.

1.1 Rulemaking Data

Ninety percent of U.S. law is now written by agencies rather than Congress (West and Raso, 2013). Legally binding agency rules give specific meaning and force to legislation and presidential agendas. Thus, the real impact of most new statutes and executive orders is largely unknown until the ink is dry on the agency rules and regulations that implement them.

As federal policymaking shifted toward the executive branch, so have interest group lobbying and public pressure campaigns. While most federal agency policies receive little public attention, activists occasionally expand the scope of conflict by targeting agency policymaking with letter-writing campaigns, petitions, protests, and mobilizing people to attend hearings, all classic examples of "civic engagement" (Verba and Nie, 1987). Yet, the relationship between the scale of public engagement and policy change remains untested. Indeed, we have much to understand about the causes and effects of public pressure campaigns before we are in a position to ask if they are a mechanism for groups to influence policy. Most critically, we must understand who mobilizes public pressure campaigns and why.

To examine the relationship between public pressure campaigns and lobbying success, I collected a corpus of over 58 million public comments on proposed agency rules. While most prior studies of public pressure campaigns targeting federal agencies focused on the Environmental Protection Agency, my sample of agency rules spans 60 agencies, allowing new insights into the scale and influence of public pressure across agencies.[1] From 2005 to 2020, these 60 agencies posted 42,426 rulemaking dockets to regulations.gov for public comments. Only 816 of these rulemaking dockets were targeted by one or more public pressure campaigns, but this small share of rules garnered 99.07 percent (57,837,674) of all comments. Using text analysis tools to strategically select 10,894 comments for hand coding in an iterative process ultimately yields a new dataset of 41,342,776 as-good-as-hand-coded texts.

---

[1]The EPA provides more extensive metadata on public comments than most agencies, enabling studies that do not require working with text data.

1.2   Methods

I leverage an iterative process of computational text analysis and hand-coding to attribute comments to the organizations, campaigns, and broader coalitions that mobilized them. I first identify comments that share text. I find that a 10-word (10-gram) phrase repeated across more than a few comments is either text copied from the proposed rule or language provided by an organized campaign. Thus, for each comment text, I first remove all 10-word phrases that appear in the proposed rule (including the preamble and call for comments). Then, I identify all comments that share ten-word phrases with 99 or more other comments. Finally, I collapse these form letter comments into one representative document for hand-coding. Text reuse thus allows me to code one sample document that is valid for all commenters using the same form letter.

Before and during hand-coding, I attempt to identify the organization(s) that submitted or mobilized each comment by extracting the names of organizations and elected officials from the text of each comment. For comments that do not include the name of an organization or elected official, human coders used an internet search on portions of the comment's text. This often identified the organization that provided the model letter text. As additional organizations were identified by hand, they were added to an entity extraction method based on regex tables, and model text linked to them was added to the clustering method. This was especially helpful for astroturf groups that provide model public comments while obscuring their identity. I thus attribute each comment to both the organization behind it (its author if submitted by the organization or the mobilizing organization for form letters) and the lobbying coalition to which that organization belonged.

I also identify comments submitted by elected officials, with special attention to members of the U.S. Congress. Members of Congress often learn about issues from—and are spurred to act by—public pressure campaigns. Legislators often submit comments from their constituents, either on their own or attached to their own comments. For example, several members of Congress attached mass mail or petitions from their constituents to their comments on the Consumer Financial Protection Bureau's controversial Payday Loan rule. Public pressure campaigns often seek to influence policy by informing elected officials of their constituents' demands. Many campaigns collect the zip code of letter signers so they can forward constituent comments to their representatives. Some form letters include a line for signers to "CC" (carbon copy) their member of Congress (see, for example, EPA-HQ-OAR-2005-0161-2624). When members of Congress comment on agency rules, they are often aware of public pressure campaigns. Many are spurred to engage in bureaucratic policymaking pressure campaigns and their constituents who participate in them.

To study the influence of organizations and coalitions, I collapse these data into one observation per organization or coalition per the proposed rule for analysis. I identify lobbying coalitions both by hand and by

textual similarity. Co-signed comments and form-letter comments are always assigned to the same coalition. I then identify the main substantive comment submitted by each organization's staff or lawyers, which are usually much longer than supporting comments like form letters.

Comments selected for hand-coding are coded for their level of lobbying success. Hand-coding includes recording the type of organization, the lobbying coalition to which each comment belongs, the type of coalition (primarily public or private interests), their policy demands, and the extent to which the change between the draft and final rule aligned with their demands. The level of alignment between policy demands and policy outcomes is a standard approach for assessing lobbying success with these kinds of observational data.

My approach to measuring lobbying success starts with policy demands raised in comments. The dimensions of conflict on which I judge lobbying success are those issues identified by commenters. Unlike other studies, the issues I use to assess lobbying success do not come from first reading the policy or any a priori concept of what each policy fight is about. Instead, I read the change between the draft and final rule with an eye for alignment with commenters' requests (including requests that specific parts of the draft policy do not change.) Using commenter requests to identify the dimensions of the conflict has advantages and disadvantages. Compared to other potential measures of success, it is more likely to focus on things that commenters care about and miss policy issues that other non-commenting segments of the public might care about.

1.3   Findings

These new data reveal who participates in public pressure campaigns and why. I find that pressure campaigns are almost always a conflict expansion tactic used by less-resourced "grassroots" groups; "astroturf" campaigns are surprisingly rare. However, the resources and capacities required to launch a campaign cause a few larger policy advocacy organizations to dominate. Over 80 percent of public comments were mobilized by just 100 organizations, most of which lobby in the same public interest coalitions. As a result, the public attention that pressure campaigns generate is concentrated on the small portion of policies on which these organizations focus. I also find no evidence of negativity bias in public comments. Instead, most commenters supported draft policies during the Obama administration but opposed those of the Trump administration, reflecting the partisan biases of mobilizing groups.

Comparing the success of campaigns across agencies, I find that the federal agencies that are most often the targets of public pressure are also the most responsive to public pressure. My hand-coded measure of whether commenters got the policy outcome they sought allows me to assess whether public pressure campaigns increase lobbying success. Finally, I assess potential mechanisms by which mass public engagement may affect policy. I focus on congressional oversight as a mediator in policy influence, i.e., the extent to which public pressure

campaigns affect policy indirectly through their effects on legislators' oversight behaviors. I find that members of Congress are more likely to engage in rulemaking when advocacy groups mobilize public pressure and that lobbying coalitions are more successful when they mobilize more legislators. However, pressure campaigns are related to policy outcomes under very narrow conditions. Lobbying coalitions are more successful when they mobilize more members of Congress, but legislators disproportionately align with private interest (e.g., business-led) coalitions, not the public interest coalitions that run most public pressure campaigns.

These new data also allow me to assess the discursive effects of the climate and environmental justice movements. I find that agencies are more likely to add language addressing climate change or environmental justice in their final rules when public comments raise concerns about climate change or environmental justice. However, the representation of both movements in federal policymaking is limited to a small number of national advocacy organizations. When less-well-resourced groups raise concerns about climate change or environmental justice, their concerns are almost always ignored.

2   Challenges and Opportunities in Studying Political Texts

Studying political texts found in the wild presents several methodological challenges. Some of these challenges are shared by other kinds of texts, such as open-ended responses to questions in surveys and lab experiments, but this paper focuses on texts generated by real-world politics.

Unlike experimental contexts where the outcome of interest is the difference in words or phrases used across treatment groups, differences in observational text data arise in systematic ways from the political process we aim to study. Open-ended survey responses often reflect rhetoric from media or political campaigns. Texts generated by political campaigns are even more directly linked. For example, letter-writing campaigns targeting government officials often provide talking points or letter templates; the processes generating political action simultaneously generate the text that government officials and political scientists observe. Likewise, social media posts and protest signs employ shared slogans and hashtags. Words and phrases represent ideas and issue frames by which campaigns aim to affect politics and policy. Similarities and differences among texts can thus be used to identify organized efforts.

At the same time, because texts are generated in rich political contexts, much is often left unsaid. Slogans, hashtags, vague policy demands, and even petitions often have much more politically relevant meaning in light of the policy fight for which they were generated. Individual participants may not even be aware of the implicit policy proposals advanced by their words. For example, depending on the political context, people demanding that the government "act on climate" may mean regulating power plants, tax credits for electric cars, or using the Defense Production Act to require firms to manufacture heat pumps. Individual participants in a pressure

campaign might not even know what specific policy changes they are helping to advance. This does not mean that the organization behind the campaign does not have specific policy demands they hope to achieve by mobilizing the mass public around a slogan. It is thus essential to study the policy impact of texts like tweets and public comments in the context of the broader campaign to which they belong.

Any study of political texts in the wild must be informed by the key observation that politics is organized. Generally, people need to be mobilized to engage in politics.

3   Hand-coding Dynamic Data

Before reviewing specific tools and methods, I briefly address the broader data management systems that make iterative hand-coding and computational text analysis feasible in a team setting. If working alone, a researcher may take an iterative approach with little additional infrastructure. One can use computational tools to select and preprocess a document to hand-code, code it, update the selection and preprocessing algorithms, and generate a second document. However, the logistics of a team setting require that computational work and hand-coding happen more or less simultaneously. It is inefficient to distribute one document at a time to coders and make them wait until all coders are done coding their first document to update and re-run the automated tools.

Here, I focus on a workflow using Google Sheets and Drive, but a similar workflow could be achieved with other cloud servers and software. A key feature of Google Sheets is that edits are visible to all users simultaneously. There is no need to send updated documents; all team members always have access to the most current version of the data. A second key feature of Google Sheets is that they can be edited with the googlesheets4 R package. An R script can pull in updates to the data, use hand-coded data to update algorithms (more on this below), generate a computationally improved version of the data, and immediately update the sheet that the hand-coder sees with the improved data. Third, spreadsheets conveniently organize a large amount of metadata generated by computational text tools that may be useful for human coders. Metadata can include hyperlinks. Hyperlinks are easy to generate automatically and can helpfully point to multiple documents on a server. For example, one column of hyperlinks may point to the locations of the original PDF documents on Google Drive (or any other server), and a second column may point to files with the plain text extracted.

For the task of coding comments on draft agency rules, I generated initial spreadsheets for each of the 200 rules in my sample. These initial sheets had between 1 and 850 comments. Where a rule received a small number of comments, the spreadsheet contained all comments on the rule. For rules with thousands or millions of comments, I selected key comments for hand coding as described in Section 6.

Figure 1: Example Coded Comments in a Google Sheet

org_name	position	comment_type	coalition	coalition_type	org_type	ask	ask1	ask2
national Association of	5	org	LIUNA	Private	corp; county	Counties support revising the NEPA in order to strengthen the involvement of	Counties urge CEQ to ensure	
Bloomington Indiana	1	org	Partnership Project	Public	gov;local	Blooington objects to several aspects of the rule, and oppose the weakening of	Cumulative outcomes	Your propo
Environmental Defense Fund	1	org	Partnership Project	Public	ngo; environmental	Environmental Defense Fund has only 1 ask	Environmental Defense Fund	
Energy and Wildlife Action	5	org	Energy and Wildlife Action	Private	corp group; energy coalition	EWAC has reviewed the Proposed Regulations and generally supports the	EWAC recommends	While suppo
Campo Band of Mission Indians	1	org	Association on American Indian	Private	gov;tribe	The tribe is requesting for more Public hearings	Executive Order 13175	The E reque
New York State Office of the	1	org	Partnership Project	Public	gov; state	Oppose the rule and urges to withdraw the rule	First, we object to the	Secoi objec
American Council of	5	org	American Council of Engineering	Private	corp group	ACEC welcomes those CEQ proposed changes to NEPA implementing	However, ACEC	ACE(recon
Franklin Regional	1	org	Partnership Project	Public	gov;local	FRCOG is against the revisions proposed for changing the NEPA.	In addition, CEQ proposes	CEQ propo
Kansas Water Authority	1	org	Partnership Project	Public	gov;local	KWA is requestion an extension of the comment period	In order to adequately	
Northwest Indian	1	org	Association on American Indian	Private	tribe association;	NWIFC is comprised of the 20 treaty Indian tribes in western Washington, that	Inserting a new threshold	Simil propo
Erie County Department of	1	org	Erie County Department of	Public	gov;county	While we recognize the need to improve certain aspects of the NEPA environmental	Lead Agencies must be given	Erie (does
NFIB	5	org	LIUNA	Private	corp gorup	NFIB genrally supports the CEQ rule but has requests for modifications	Modify proposed 40	confo propo

A team of 15 coders worked in separate spreadsheets. For training and in order to assess inter-coder reliability, we initially coded the same comments on the same rules. Once calibrated, coders worked on separate rules. When a coder began working on a new rule, they saved a version of the spreadsheet with their name at the end. As information about organizations, coalitions, and policy demands was recorded, the google drive API allowed it to be pulled in and analyzed in real-time.

In addition to dynamically improving the uncoded data, as more documents were coded, real-time analysis enabled a series of systematic checks for irregularities in coders' work. The script pulling in data from Google Drive recorded coder names as a variable so that coding decisions could be traced to the coder. For example, when new information about comments was coded on a Google Sheet, scripts measured the variance in the commenter positions, policy demands, and lobbying success (see Section 7). This helped to assess the coherence of hand-coded coalitions. Where coders grouped organizations into the same coalition but coded them as having very different policy positions, these coding decisions were flagged for review, allowing review and feedback in real time.[2] For example, Figure 2 shows a suspicious coalition lobbying on the IRS internet gamboling rules that were flagged because one of its members was coded as having been mostly successful while the others were coded as not having their demands met. All coalition members used similar language, requesting that the IRS clarify that certain practices were exempt from the regulation. However, while all organizations discussed dog racing, one organization asked that the agency state that "pari-mutuel wagering" was except, which the IRS did,

---

[2]For an example of a dashboard for reviewing coder decisions in real-time, including code to pull and merge data from Google Sheets see https://judgelord.github.io/correspondence/FERC/DOE_FERC-letter-coding.html

while the others requested that the IRS state that dog racing was exempt, which the IRS did not. Despite the texts being similar, the key phrase–*what* it was that commenters sought exemption for—differed. The commenter requesting an exemption for pari-mutuel wagering was then put into a different lobbying coalition.

Figure 2: Incorrectly Labeled Coalition Identified by Automated Check

▸ J ◂	▸ P ◂	▸ R ◂	▸ V
org_name	coalition_co	ask	success
Sport View Television Corporation	Greyhound racing industry	the Board and Treasury must specifically state that pari-mutuel wagering, which is state sanctioned and regulated, is legal and therefore outside the scope of this regulation and should be permitted.	1
Greene Group, Inc.	Greyhound racing industry	The Congressional Record clearly states that if the use of the Internet in connection with dog racing is approved by state regulatory agencies and does not violate any Federal law, then it is allowed	-2
American Greyhound Racing, Inc.	Greyhound racing industry	The Congressional Record clearly states that if the use of the Internet in connection with dog racing is approved by state regulatory agencies and does not violate any Federal law, then it is allowed	-2

Additionally, the new information provided by coders allowed constant updates of algorithms that auto-coded organizations and coalitions, as described in Section 4. Updated auto-coding automatically re-populated the Google Sheets, reducing the remaining work for coders. Similarly, information from ongoing coding improved the selection of informative comments, often reducing the number of remaining comments requiring coding, as described in Section 6.

4   Itereatve Processing

Researchers typically see the computational tasks discussed in this section as "preprocessing" steps. In this section, I briefly describe several preprocessing steps but focus mainly on steps that can often be done more iteratively.

4.1   Preprocessing: digitizing, cleaning, and summarizing

In many cases, using text data requires that we first extract it from pdf. In addition to enabling quantitative text analysis, wrangling text data into machine-readable plain text can help with certain coding tasks. Instead of typing out key features that coders are tasked with recording (e.g., named entities or key quotes), coders can copy and paste text from plain text documents. For example, coding the policy demands of commenters meant

recording verbatim quotes from comment letters. Without providing plain text versions of the documents, this would require re-typing quotes wherever here original files were not machine-readable, making quotes vulnerable to human error.

Where documents are longer than a paragraph, hand-coding can often be sped up by providing a computer-generated summary of the document. Relevant summary information depends on the aim of hand-coding. For coding public comments, I provided two types of summary information, both using actual sentences in the document.

First, I provided a three-sentence summary of comments using the `textrank` software to identify the three sentences intended to be representative of the document.[3] Indeed, federal agencies and their consultants use this same software when creating their own summaries of comments. Summarizing can be useful in several ways. Because lobbying coalitions coordinate on talking points, it is often possible to code policy positions and coalition membership from a few sentences. Thus, for studying the impact of comments and pressure campaigns on agency rules, summaries of comments can provide codes with information about the commenter's position, the coalition they belong to, and potentially some of their most important policy demands.

Second, I selected sentences that contained phrases of substantive import. Specifically, for my studies of climate and environmental justice advocacy, I extracted any sentences that contained "climate change" or "environmental justice," respectively, and put them into spreadsheet columns. This aided coders in paying special attention to policy demands related to these issues. It was also necessary for coding tasks specific to these issues. For example, where policy demands related to environmental justice were present, the coders recorded additional information, including whether the environmental justice demands were about local or national issues and whether they were the organization's primary concern.

Finally, summaries of policy documents were also helpful. Agency rules average over eighty pages and are full of dense technical jargon. In addition to the rule's abstract, a two-sentence summary **of each section** of the rule can help coders get oriented to the content of each section of the rule and speed up finding the appropriate section to examine in order to assess whether the requested changes to the rule were made.[4]

4.2   Iterative entity tagging with dynamic regex tables

Through an iterative process of computational text analysis and hand-coding, comments are linked to the organizations, campaigns, and broader coalitions that mobilize them. I am thus able to attribute each comment to both the organization behind it (its author if submitted by the organization or the mobilizing organization for

---

[3]For more on textrank, code for applying it to comments, and comparison to hand-selected summaries, see: https://judgelord.github.io/rulemaking/textrank_comments

[4]For more on using `textrank` to summarize policy documents, and R code to do so for agency rules, see https://judgelord.github.io/rulemaking/textrank_summary

Table 1: Regex Table Created from Center For Responsive Politics Data

parentName	pattern
3M Co	3M Co|3M Cogent|3M Health Information Systems|Ceradyne|Cogent Systems|Hybrivet Systems
Teamsters Union	Brotherhood of Locomotive Engineers & Trainmen|Brotherhood of Maint of Way Employ Div|New Engl

form letters) and the coalition to which that organization belonged. I also identify comments submitted by elected officials, with special attention to members of the U.S. Congress. This process involves using regular expressions to search comment texts and metadata for known entities. Known entities are entities for which we already have required information (either from prior coding or from an external source), such as the organization's type (non-profit, company, government, etc.). The organization behind each comment is identified and either linked or added to a growing list of organizations known to comment. This corpus of known organizations is then included in the next search.

With this approach, I identify the organizations responsible for over 40 million comments, including all organizations responsible for mobilizing 100 or more comments with repeated text–either identical or partially unique texts containing shared language.

NLP tools make entity extraction simple. However, political entities go by many aliases, often frustrating the linking of records. Because various datasets of entities are not created with regular expression matching in mind, few provide aliases or variants of names. When datasets include substantial metadata, probabilistic record linking tools may help link imperfect records (e.g., Enamorado et al., 2017). However, researchers must often match records on name strings that may differ across datasets. Below, I outline two approaches to this task that can be done iteratively.

4.2.1   Consolidating entity name variants with regex tables

A regex table is one way to link multiple strings to a single entity. A regex table is a lookup table where one of the two columns is a regular expression. Regex tables are used by Google Tag Manager and other tools where a lookup table aims to capture a many-to-one relationship (e.g., many names for the same entity). For example, the Center for Responsive Politics tracks entities that donate to political campaigns or file lobbying disclosures. These data include the parent company or organization of each entity. For example, union locals are members of a national union, and subsidiary companies are owned by a parent company. A user could generate a regex table from these data by pasting all "child" organizations for each parent together with a regex "or" separator. Table 1 shows a regex table created from the Center for Responsive Politics lobbying data for the Teamsters Union (five aliases) and the company 3M (eight aliases).

Table 2: Regex Table from the legislators R Packages, Legislators Named Elizabeth in the 117th Congress

bioname	pattern
FLETCHER, Elizabeth	elizabeth fletcher|e fletcher|na fletcher|(^|senator |representative)fletcher.{1,4}tx|fletcher, eliz
WARREN, Elizabeth	elizabeth warren|elizabeth herring|elizabeth ann warren|e warren|elizabeth a warren|elizabe
ESTY, Elizabeth	elizabeth esty|e esty|elizabeth h esty|na esty|(^|senator |representative)esty|esty, elizabeth|est
DOLE, Elizabeth Hanford	elizabeth dole|elizabeth hanford dole|e dole|elizabeth h dole|na dole|(^|senator |representativ
FURSE, Elizabeth	elizabeth furse|e furse|na furse|(^|senator |representative)furse|furse, elizabeth|furse elizabet
PATTERSON, Elizabeth J.	elizabeth patterson|elizabeth j patterson|e patterson|na patterson|(^|senator |representative)
HOLTZMAN, Elizabeth	elizabeth holtzman|e holtzman|na holtzman|(^|senator |representative)holtzman|holtzman,
ANDREWS, Elizabeth Bullock	elizabeth andrews|elizabeth bullock andrews|e andrews|elizabeth b andrews|(^|senator |repr
GASQUE, Elizabeth Hawley	elizabeth gasque|elizabeth hawley gasque|e gasque|elizabeth h gasque|(^|senator |representati

Table 3: Regex Table of Typos from the legislators R Packages, Legislators Named Elizabeth

typos	correct
elezabeth	elizabeth
elizabeth(| [a-z]*)etsy($| |,|;)|etsy, elizabeth	elizabeth esty
elizabeth(| [a-z]*)varren($| |,|;)|varren, elizabeth	elizabeth warren

The `legislators` R package uses regex tables to identify U.S. legislators in text using various common permutations of legislator names (e.g., "Elizabeth Warren, Senator Warren, or Warren, Elizabeth). As users provided nicknames, this package expands the regex pattern (e.g.,"Liz Warren"). Table 2 shows a regex table for members of the 117th Congress named Elizabeth. The `legislators` package contains a second regex table of common typos, for example, those introduced by ORC (e.g., "Senator Varren") that can be used to replace strings as a preprocessing step that increases the number of legislators successfully matched by the main regex table. Table 3 shows a regex table for common typos or OCR-introduced errors for legislators named Elizabeth. Note that the misspelling "Elezabeth" is replaced by "Elizabeth" any time it appears in the provided text, potentially increasing matches for both legislators named Elizabeth.

4.2.2 Initial regex tables can be developed inductively and deductively.

Inductively building regex tables inductively involves constructing from scratch tables of known entities found in the data. Deductively building regex tables involves using existing databases of organizations (e.g., organizations in the Center for Responsive Politics data) as search patterns. There are advantages and disadvantages to both approaches. Researchers will want to do both when confronting an unknown population of potential entities. When the full population of entities is known (e.g., searching a database of congressional letters for members of Congress), a deductive approach is sufficient.

Inductively, I started with the metadata field for the comment author's organization and then used entities from the text when this field was empty or useless (as it was for large sections of these data).

Where metadata existed for commenting organizations, I first cleaned up strings by removing extra white space prefixes (e.g., ".* on behalf of") and suffixes (e.g., "LLC"). I then counted the frequency of all the strings appearing in this field. Many of the most frequent strings were aliases for the same organizations. These multiple aliases were the start of the inductive regex table. Using the inductive regex table to replace all strings with known organizations, I began the work of iteratively consolidating the list of most frequent commenters from the top (most frequent) down until every string that appeared more than 100 times was attached to a known organization or coded as useless (e.g., "anonymous").

I repeat this inductive process for comments that lacked useful metadata but now with entities extracted from the text. Like many political texts, public comments often mention many entities that are not the authors of the comment, especially the agency to whom the comment is addressed. Without manually inspecting the text of the comment, the secondary regex table of extracted entities (only used when the metadata was useless) was mainly reliable for short comments and form letters that did not contain citations and references to entities that were not (but a priori, could have been) a comment author. However, when provided to coders, this list of extracted entities helpfully provided a short list of potential authors. Human coders with an understanding of the politics of the particular policy process are often able to quickly pick out the correct, only plausible entity from the list of extracted entities. Thus, while only reliable for auto-coding comment authors in certain cases, extracted entities can also be seen as a useful form of summary metadata to speed up hand-coding.

4.2.3   Iterative and-coded and adding to regex tables

Hand-coders frequently generate information that can improve regex tables like the ones above. Indeed, parent organizations are generally identified qualitatively by Center for Responsive Politics staff. The nicknames and typos included in the `legislators` package were identified by dozens of research assistants who hand-coded texts known to contain the names of members of Congress.

For the sample of rulemaking comments that were hand-coded, coders were provided any original metadata and the "best guess" organization from the inductive and deductive regex tables described above. Coders identified organizations where no entity was identified and occasionally corrected the method. For example, the initial inductive method coded "Earthjustice on behalf of Alabama River Association" as "Alabama River Association." Hand-coders corrected this to Earthjustice, the main mobilizing organization responsible for the comment. "Earthjustice on behalf of Alabama River Association" was then added to the regex table such that any future occurrences would be codded as "Earthjustice."

Likewise, starting from the premise that nearly all participation in rulemaking results from organized lobbying efforts, hand-coders searched for organizations behind comments. For comments that do not include

the name of an organization or elected official, coders used an internet search on portions of the comment's text. This often identified the organization that provided the model letter text.

Hand coders identified many organizations and variations of organization names that were not in the top 100 most frequent commenters or one of the existing databases. As additional organizations were identified by hand, they were added to the regex tables. As new organizations are added and new search patterns are added, regex tables grow. To the extent that the new organizations and patterns match yet-uncoded observations, the share of documents matching a known organization also grows. For documents selected for hand-coding, this means one less step. Once a coder has identified an alias, all observations they encounter in the future will have the correct organization tagged (i.e., the correct organization will already be filled in the spreadsheet).

5   Identifying Coalitions with Iterative N-gram Clustering and Hand-coding

Because politics is organized combat (Hacker and Pierson, 2010), the proper unit of analysis for assessing the impact of lobbying and pressure campaigns is rarely the individual or individual organization. Instead, the proper unit of analysis is a lobbying coalition—a group of organizations advocating for the same policy changes in their comments on a draft rule. Advocacy organizations work together on campaigns. For example, Save our Environment submitted both sophisticated comments and collected signatures from hundreds of thousands of people on several rulemaking dockets. Save our Environment is a small non-profit with a simple WordPress website almost entirely dedicated to mobilizing public comments. It is run by The Partnership Project, a coalition of 20 of the largest environmental advocacy organizations in the United States, including the Sierra Club, Natural Resources Defense Council, Greenpeace, and the World Wildlife Fund, with the aim of "harnessing the power of the internet to increase public awareness and activism on today's most important environmental issues" (Saveourenvironment.org, 2021). Several Partnership Project members, including the Sierra Club, EarthJustice, and NRDC, also submitted technical comments and mobilized hundreds of thousands of their own supporters to comment separately on the same rules. These lobbying and mobilizing activities are not independent campaigns. These organizations and the people they mobilize are a coalition.

To mobilize broader support, advocacy organizations often engage smaller organizations, which, in turn, mobilize their members and supporters, often with logistical support and funding from the larger national organization. For example, for the same campaign where the Gulf Restoration Network mobilized hundreds of restaurants that serve sustainable seafood, one of their larger coalition partners, the Pew Charitable Trusts, mobilized thousands of individuals, including members of the New York Underwater Photography Society. These smaller organizations did not identify themselves as part of Pew's campaign, but their letters used almost identical language.

Identifying which people and organizations belong to which coalition is thus a crucial first task for any study of public pressure campaigns. I use several strategies to identify whether a pressure campaign mobilizes a given comment. I first use textual similarity to identify clusters of similar comments, reflecting formal and informal coalitions. Comments with identical text indicate a coordinated campaign.

To link individual comments and public pressure campaigns to the more sophisticated lobbying efforts that they support (if any), I identify the lobbying coalition(s) (if any) to which each comment belongs. Some individual commenters and organizations are unaffiliated with a broader lobbying coalition, but, as I show below, most people and organizations lobby in broader coalitions.

Importantly, even campaigns that achieve very low public response rates appear in these data. Because campaigns aim to collect many thousands of comments, it is implausible that even the most unpopular position would achieve no supportive responses. For example, Potter (2017) found Poultry Producers averaging only 319 comments per campaign. While this is far from the Sierra Club's average of 17,325 comments per campaign, it is also far from zero. (These numbers are from Potter's sample of EPA rules; the Sierra Club's average is even larger in my sample; see Table 4.)

Probabilistic clustering methods—especially unsupervised methods like k-means clustering and latent Dirichlet allocated topic models—have received a great deal of attention from methodologists, but there is less empirical work using these methods. In part, this may be a result of comment features of data that make unsupervised text models difficult. First, there are often an unknown number of dimensions of conflict in a given policy process. Additionally, there may be major dimensions of variation that do not map onto political conflict—for example, sophisticated organizations on both sides may use polite and technical language, resulting in similar word frequency distributions. In contrast, members of the mass public on both sides may be less polite. (2) some desired clusters (e.g., coalitions) often contain vastly more documents and have more diversity of word use within their coalition than between some members of a coalition and a small number of comments on the smaller side.

To identify coalitions in rulemaking, I start with an approach that relies on text reuse. Text reuse is a much more robust indicator of a connection between documents than, for example, word frequencies. Long strings of words appearing in the same order are unlikely to appear by chance.

5.1   Collapsing form letters with text reuse

I first identify comments that share text. I find that a 10-word ("10-gram") phrase repeated across more than a few comments is either text copied from the proposed policy or language inspired or provided by an organized campaign. Thus, for each comment text, I first remove all 10-word phrases that appear in the proposed rule

(including the preamble and call for comments). Then, I identify all comments that share ten-word phrases with 99 or more other comments. Finally, I collapse these form letter comments into one representative document for hand-coding. Text reuse thus allows me to code one sample document that is valid for all commenters using the same form letter.

For each comment on a rulemaking docket, I identify the percent of words it shares with other comments using a 10-gram moving window function, looping over each possible pair of texts to identify matches.[5] When actors sign onto the same comment, it is clear that they are lobbying together. However, various businesses, advocacy groups, and citizens often comment separately, even when they are aligned. Text reuse (using the same ten-word phrases) captures this alignment. When individuals use identical wording, I interpret that to mean they're endorsing the same policy position as part of a lobbying coalition.

1) shows the percent of shared text for a sample of 50 comments on the Consumer Financial Protection Bureau's 2016 Rule regulating Payday Loans. Comments are arranged by the document identifier assigned by regulations.gov on both axes. The black on the diagonal indicates that each document has a perfect overlap with itself. Black squares off the diagonal indicate additional pairs of identical documents. For example, 100 percent of the words from Comment 95976 are part of some tengram that also appears in 95977 because the exact same comment was uploaded twice. The cluster of grey tiles indicates a coalition of commenters using some identical text. Comments 91130 through 91156 are all partial or exact matches. All are part of a mass comment campaign by Access Financial. The percentage of identical text is lower than in many mass-comment campaigns because these are hand-written comments, but the n-gram method still picks up overlap in the OCRed text in the header and footer. Tengrams that appear in 100 or more comments indicate a mass comment campaign. Some agencies use similar "de-duping" software (Rinfret et al., 2021) and only provide a representative sample comment. In these cases, my linking method assumes that the example comment is representative, and I link these comments to others based on the text of the sample comment provided.

N-gram windows have advantages and disadvantages over other text-reuse methods. Compared to Smith-Waterman, n-gram matching is very fast.[6]

---

[5]More details and code implementing an n-gram moving window in R are available here: https://judgelord.github.io/rulemaking/payday_comments.html

[6]For more about n-gram window functions and comparisons with related partial matching methods such as the Smith-Waterman algorithm, see Casas et al. (2019) and Judge-Lord (2017).

Figure 3: Example: Identifying Coalitions by the Percent of Matching Text in a Sample of Public Comments
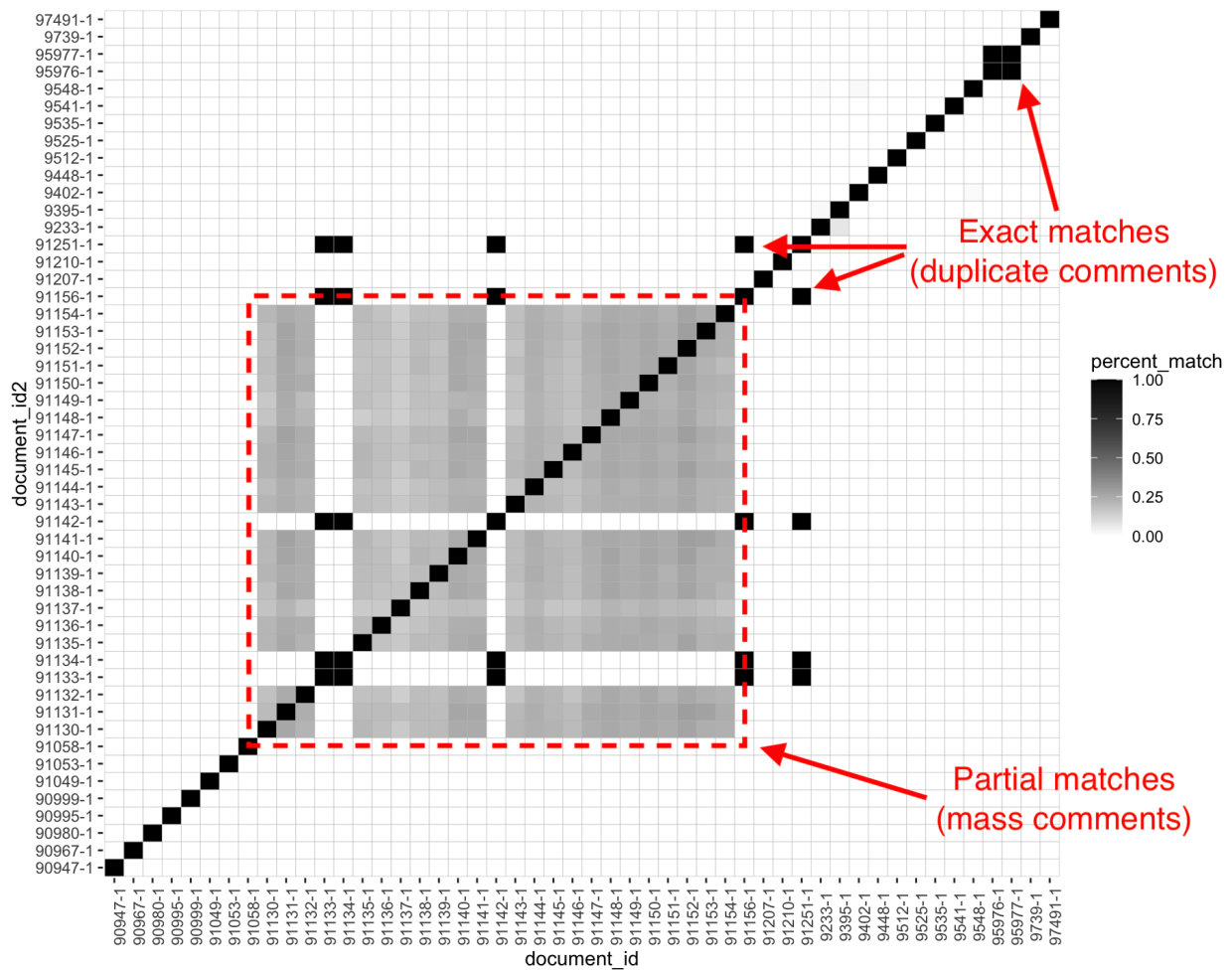
More importantly, n-gram matching is generally more inclusive of text that has been re-arranged.[7] The aim, in this case, is simply to identify texts that cross a minimum threshold of association that is unlikely to occur by chance. 10-gram text reuse is well-suited to this task.

5.2   Hand-coded coalitions and key demands

This approach differs from previous studies of mass comment campaigns in at least two ways. First, my methods allow me to identify coalitions consisting of multiple organizations. Previous studies measure mass comment campaigns at the organizational level. For example, Balla et al. (2020) analyzes "1,049 mass comment campaigns that occurred during 22 EPA rulemakings"— an average of nearly 50 "campaigns" per rule. By "campaign," Balla et al. (2020) mean an organization's campaign rather than a coalition's campaign. Especially on EPA rules, there are rarely more than two or three coalitions engaging in public pressure campaigns–one of the environmental advocacy groups and their allies, another of regulated industry groups and their allies. Using organizations as the unit of analysis means that observations are far from independent. An analysis that counts one coalition's campaign as 40 smaller "campaigns" with the same policy demands would count this one campaign as 40 observations.

In contrast, my methods allow me to measure levels of public pressure and lobbying success per organization *and* per coalition. Like previous studies, I identify the organizations responsible for mobilizing comments. Where other studies leverage the fact that the EPA gathers substantially similar comments, I can identify mass comment campaigns across dozens of federal agencies. Additionally, I further link common efforts by multiple organizations lobbying in a broader coalition. This allows for analysis with the lobbying coalition as the unit of analysis. By measuring comments per coalition, both through hand-coding and text reuse, I capture different levels of public pressure than we would see if we were only to measure the number of comments per organization.

The second major difference between my approach and previous research is that I do not compare policy-makers' responses to sophisticated comments to policymakers' responses to mass comments. Rather, I *attribute* mass comments to organizations and coalitions that also submit sophisticated technical comments. The set of comparisons one makes is critical to any study of responsiveness or policy influence. Researchers may reach different conclusions if they compare different things. Consider a study comparing how agencies respond to Sierra Club form letters to how they respond to the Sierra Club's sophisticated comments. Now consider a study that compares responsiveness to the Sierra Club's sophisticated comments between rules where they did

---

[7]Smith-Waterman aims to identify the best alignment between two documents. If two sections of text have been flipped within a document, Smith-Waterman will only detect one of them. If each sentence or 10gram were considered a "document," it is possible that a Smith-Waterman alignment threshold may come up with better results than exact matching.

and did not run a mass comment campaign. A study comparing the average influence of form-letter comments to the average influence of sophisticated comments is very different from a study that compares the influence of two sets of sophisticated comments with different *levels* of public pressure behind them. Where previous studies take the former approach, I take the latter.

5.3 Iteratively adding coalition members with text reuse

As hand coders added strings representing key policy demands, these bits of text provided an additional way to computationally identify coalitions. Because even sophisticated organizations collaborate on talking points and key policy demands, comments that shared these strings indicated a lobbying coalition. All other comments sharing these policy asks were thus automatically added to the coalition. This was especially helpful for astroturf groups that provide model public comments while obscuring their identities, especially the fact that they are centrally organized by the same lobbying firm.

6 Selecting Texts of High-inferential Value

Beyond speeding up the process of identifying organizations and coalitions, the mix of hand-coding and computational methods described above enables a sampling approach that can provide much greater leverage than a random sample. When confronted with more observations than can reasonably be hand-coded, researchers often opt for a random sample with the intent that their conclusions will be representative. A random sample may miss important but sparse observations in many populations, especially populations heavily skewed to one side on a key dimension. For example, public interest groups frequently mobilize thousands of public comments while the major regulated industry only submits one. A random sample of comments on the Trump Administration's Affordable Clean Energy Rule, for example, would almost be guaranteed to be 100% form-letter comments from environmentalists. This is deceptive in two ways. First, one might conclude that the environmentalists generally lacked sophisticated policy demands unless one drew one of the few comments from environmental lawyers. Second, unless one happens to draw the single comment from the American Petroleum Institute or a handful of other industry groups, one might erroneously conclude that regulated industry was absent. In fact, the regulated industry was much more successful in having their demands met than the environmental groups.

Instead of a purely random sample, I seek samples representing key dimensions of variation in two ways: representative texts and key outlier texts. I select several samples of comments for hand-coding.

One sample contains at least one comment from each group of 100 or more similar comments. As discussed above, I use text reuse to identify coalitions and select only one representative document for hand-coding. Because 99% of comments are form letters, collapsing from letters alone reduces the data by two orders of magnitude. When the same organization mobilizes around more than one form letter, we can further collapse

these form letter campaigns. This census of form-letter comments allows me to make valid observations about public pressure campaigns across agencies and over time.

Next, I take two samples from the population of rules and initially select all comments that either indicate they are from an organization or submit comments as an attached file.[8] This approach captures nearly all comments that do not form letters or bot-generated comments.ˆIndividual citizens rarely engage without being part of an organized campaign.] The first includes nearly all non-form letter comments on a random sample of 100 rules. A second sample includes nearly all non-form letter comments on another random sample of rules, weighted by the number of comments they received. These last two samples allow me to make inferences about lobbying coalitions that do and do not use public pressure campaigns.

Second, I select outlier texts of high inferential values. Having auto-coded many organizations, I use organization metadata (including metadata created by iterative coding) to ensure that at least one comment from each type of organization is coded. Commenter types include non-profits, businesses, elected officials, state governments, and tribal governments. Even if a regulated industry only submits one comment, this approach guarantees that it will be coded, thus enabling analysis of whose lobbying affects rules, even when the key texts might otherwise be lost in overwhelming volumes of text data. Further, I screen documents for rare characteristics that indicate potentially important outlier observations. These include comments that are particularly long and sophisticated, from a known list of often influential organizations (e.g., the U.S. Chamber of Commerce or members of Congress), or that contain many file attachments (often scientific reports backing up their data.) A tiny percentage of public comments meet these criteria, but nearly all most influential comments meet these criteria. Moreover, nearly all of the lead mobilizing groups—groups like Earthjustice and the Sierra Club—also submit long technical comments. Theoretically, the impact of any public pressure campaign should ultimately be observed in the impact of the technical comments of the mobilizers of public pressure.

Rarely are there more than a hundred sophisticated comments on a given rule, even rules that receive over a million comments. Thus, the combined effect of employing these methods is to reduce the number of comments by up to four orders of magnitude on the most challenging rules (from one million to one hundred).

As more comments on a rule are hand-coded, the number of comments selected for coding changes often decreases.

I then identify the main substantive comment submitted by each organization's staff or lawyers, which are usually much longer than supporting comments like form letters.

---

[8]nearly all minimally serious organizations submit comments as file attachments rather than typing in the text box provided by regulations.gov

7 Inferring Lobbying Success From the Success of Others in a Coalition

Comments selected for hand-coding are coded for their level of lobbying success. The extent to which the change between the draft and final rule aligned with their demands. The level of alignment between policy demands and policy outcomes is a standard approach for assessing lobbying success with these kinds of observational data (Yackee and Yackee, 2006).

The innovation in my approach is to use text reuse to automatically identify common policy demands across comments. Coders copy the text of the top three policy asks into three spreadsheet columns. I then search the text of all uncoded comments for these text strings. Where these text strings appear, I know that the commenter made the same policy demand, which necessarily had the same level of lobbying success and likely belongs to the same lobbying coalition. Unless the other comments containing the same ask are of particularly high inferential value, as-good-as hand-coded coalition membership and lobbying success means that the additional comments containing the same policy ask no longer need to be coded by hand.

7.1 Coding commenter demands

My approach to measuring lobbying success starts with policy demands raised in comments. The dimensions of conflict on which I judge lobbying success are those issues identified by commenters. Unlike other studies, the issues I use to assess lobbying success do not come from first reading the policy or any a priori concept of what each policy fight is about. Instead, I read the change between the draft and final rule with an eye for alignment with commenters' requests (including requests that specific parts of the draft policy do not change.)

Using commenter requests to identify the dimensions of the conflict has advantages and disadvantages. Compared to other potential measures of success, it is more likely to focus on things that commenters care about and miss policy issues that other non-commenting segments of the public might care about.

7.2 Coding policy positions

To assess whether organizations and their broader coalitions lobby unopposed or in opposition to other interests, I code the position of each organization on each proposed policy, given the direction of change from the current policy. Specifically, I trained research assistants to place comments on a spatial scale relative to the change between the status quo and proposed rules like the one shown in Figure 4. In Figure 4), x_1 is the current (status quo) policy, and x_2 is the new proposed policy on which commenters are commenting. Let p_i be commenter i's ideal policy.

In spatial models, whether an organization supports or opposes a proposed policy change generally depends on whether the policy is moving closer or further from its ideal policy. For example, if the ideal point of

Figure 4: Coding the Spatial Position of Comments on Proposed Policy Changes

Position of Commenter i (p_i) on Proposed Policy x_2, Given Current Policy x_1
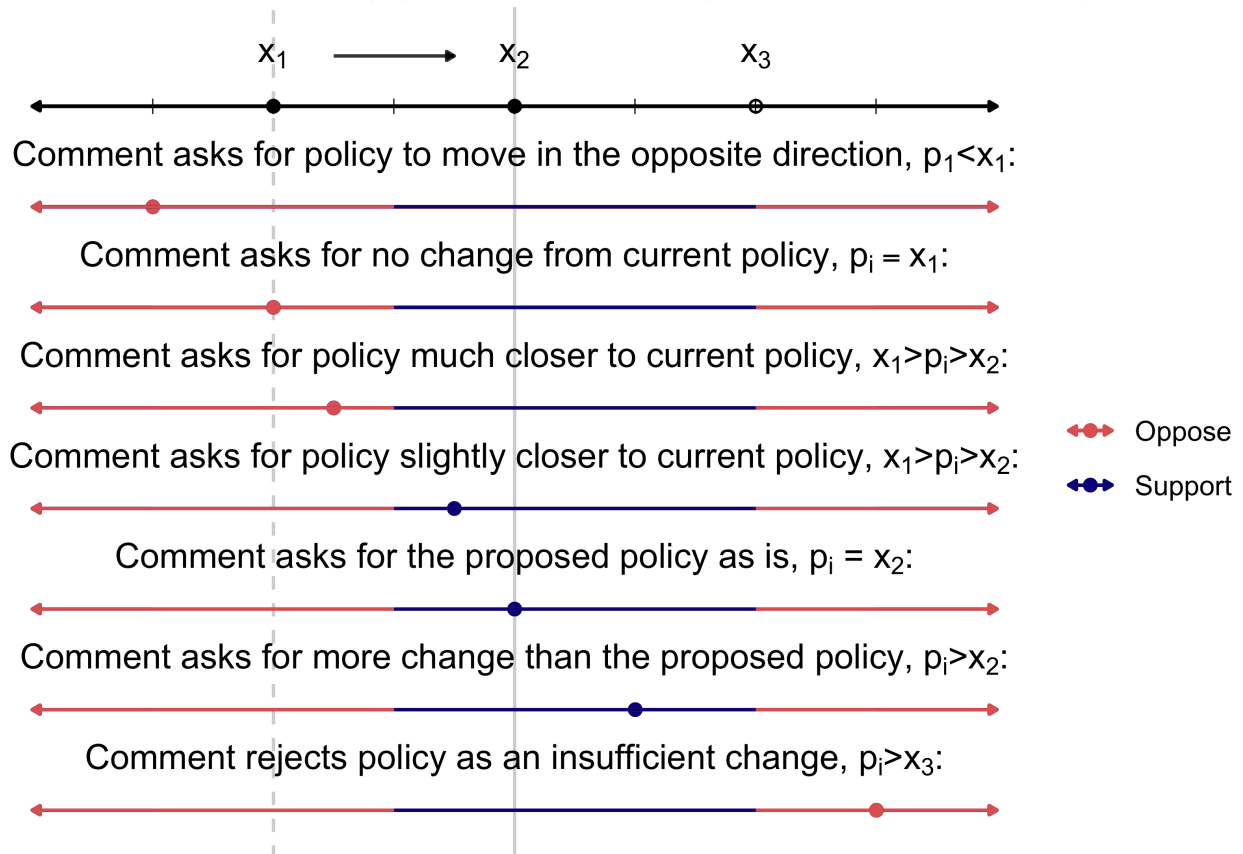
commenter 1 is the current policy (i.e., $p_1 = x_1$) or close to it, they will oppose any proposed change. Likewise, if the ideal point of commenter 2 is the new proposed policy ($p_2 = x_2$) or closer to it, they likely support the proposal.

While potentially incompatible with an assumption of single-peaked preferences assumed by most models, commenters do occasionally oppose a policy change for moving insufficiently in their preferred direction (e.g., describing the proposal as "too little" or "insufficient" to gain their support). For example, if a commenter prefers a more extreme change and will not accept anything less than a certain level of change ($p_i \geq x_3$), they may oppose x_2 as "insufficient." This is likely a result of the repeated game nature of policymaking, where commenters believe that rejecting a small change in their preferred direction (x_2) now is likely to result in a more extreme and preferred change (x_3) later.

If a commenter made statements like "We need stronger, not weaker regulations" or "These regulations are already bad for our business, we should not make them even more strict," they were coded as opposed to the proposed rule for moving in the wrong direction ($p_i < x_1$). If the commenter expressed a preference for the status quo over the proposed rule ($p_i = x_1$), they were also coded as opposing the proposed rule.

Conversely, when a comment included statements like "we applaud EPA's efforts to regulate, but would prefer less severe limits," this was coded as supporting the rule but asking for less change. If the commenter expressed unqualified support for the proposed rule ($p_i = x_2$) or requested even more policy change ($p_i > x_2$), they were almost always coded as supporting the rule.

Opposition to a proposed rule because it was insufficient to gain the organization's support was rare but did occur. For example, one commenter stated that "[w]hile the proposed rule may improve current protections to some degree, it is utterly inadequate…If the agency fails to revise the rule to incorporate such measures, then they should withdraw the proposed rule completely" (NOAA-NMFS-2020-0031-0668). Taking the commenter at their word, this was coded as opposition to the proposed rule, even though the commenter's spatial position is closer to the proposed rule than the current policy.

Having identified the coalition lobbying on each proposed rule and each organization's position, I assign each coalition's position as the position of the lead organization. For robustness, I also calculate the coalition's average position as the average position of its members. Coalition members usually have nearly identical positions, but occasionally, some take more extreme positions than others. For example, while all coalition members may have the same policy demands, some may ask for additional changes. I consider diverging interests to be one coalition only if the asks are entirely compatible with the position of organizations that did not ask for them. Conflicting policy demands indicate different coalitions.

8   Results: Patterns of Public Engagement in Rulemaking

8.1   Descriptive

8.1.1   Most Comments Result from Public Pressure Campaigns

Figure Figure 5) plots the number of comments received on regulations.gov each year from 2005 to 2020. Columns are shaded by whether I have classified each comment as part of a public pressure campaign (a mass comment campaign). Figure 5) shows that every year since 2007, the vast majority of comments on draft regulations posted to regulations.gov were the result of a public pressure campaign. All other comments (including comments from individuals acting alone and sophisticated comments from companies, governments, and other organizations) make up a small portion of all comments. Furthermore, the rise in the total number of comments from 2005 to 2013 is much steeper than the rise in the number of rules being published.

Figure 5: Public Comments, 2005-2020
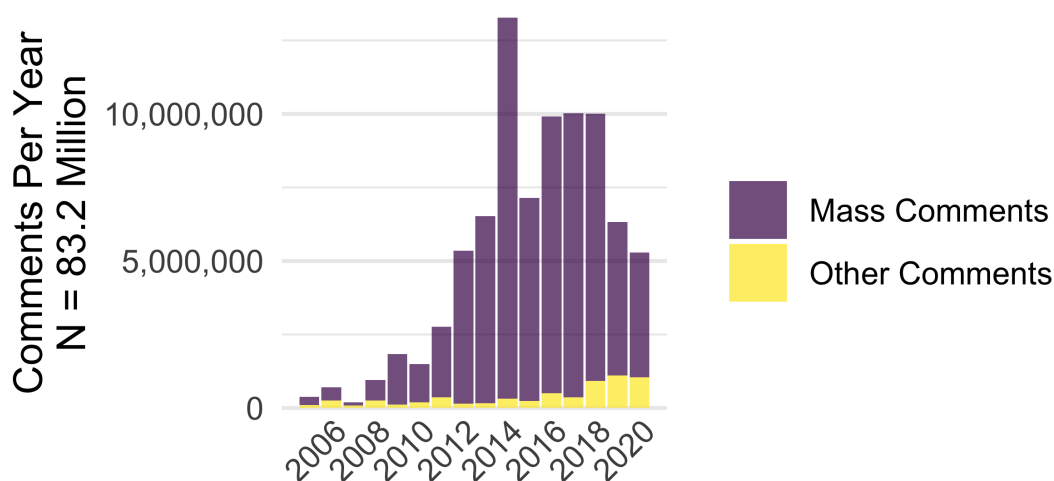
8.1.2   Most Comments and Campaigns are Mobilized by Public Interest Coalitions

Public pressure campaigns are almost exclusively organized by coalitions that include groups that engage in sophisticated technical lobbying. These coalitions include organizations that engage in sophisticated lobbying. Indeed, many of the most prolific organizers of public pressure campaigns are also among the organizations most frequently engaged in sophisticated lobbying. Public pressure is a compliment, not an alternative to sophisticated technical lobbying.

Table 4 shows the top organizers of comments posted to regulations.gov. The top ten organizations (NRDC, Sierra Club, CREDO, Environmental Defense Fund, Center For Biological Diversity, Earthjustice, World Wildlife Fund, National Wildlife Federation, Friends Of The Earth, Environment America) mobilized 42 percent of

comments on proposed rules posted to regulations.gov (25,947,612). All of these top ten organizations have lawyers on staff who engage in sophisticated lobbying, and all ten lobby together in the same coalitions. Nine are closely aligned environmental groups. Earthjustice began as the Sierra Club Legal Defense Fund. Eight of these nine organizations (all but the Center for Biological Diversity) are members of the Partnership Project, a 501c3 nonprofit founded by 20 leading environmental groups with the aim of " creating a sum of citizen participation and advocacy greater than they could generate acting apart" (Saveourenvironment.org, 2021). CREDO Action is a generic progressive group often mobilized to amplify progressive public interest campaigns. The top 100 organizations mobilized 71 percent of comments on proposed rules posted to regulations.gov (43,938,811). Each mobilized between 39,729 and 5,939,264 comments.

Table 4: Organizations Mobilizing the Most Public Comments 2005-2020

Organization	Rules Lobbied On	Pressure Campaigns	Percent (Campaigns /Rules)	Comments	Average per Campaign
NRDC	530	62	11.7%	5939264	95795
Sierra Club	591	110	18.6%	5111922	46472
CREDO	90	41	45.6%	3019150	73638
Environmental Defense Fund	111	31	27.9%	2849517	91920
Center For Biological Diversity	572	86	15.0%	2815509	32738
Earthjustice	235	59	25.1%	2080583	35264
World Wildlife Fund	69	9	13.0%	1133001	125889
National Wildlife Federation	141	36	25.5%	1113056	30918
Friends Of The Earth	127	28	22.0%	1051930	37569
Environment America	40	24	60.0%	833680	34737

The percent of rules on which each organization lobbies with a pressure campaign rather than without one (the "Percent" column in Table 4) shows only a few organizations using pressure campaigns the majority of the time they lobby. Most lobbying organizations that lobby in rulemaking rarely use pressure campaigns. The most extreme example is the American Petroleum Institute (API), which lobbied on hundreds of rules between 2005 and 2020, more than most of the other top mobilizing organizations. Yet it almost never uses

public pressure campaigns (at least in its own name). While API did sponsor a number of astroturf campaigns by front groups, API almost always lobbied without the aid of a pressure campaign. Almost all of these top mobilizing organizations usually rely on their legal and policy teams alone. The fact that so many of the top mobilizers are also highly sophisticated lobbying organizations like the Sierra Club and API lends support to my argument that public pressure campaigns are one tool that advocacy organizations may use in addition to more insider tactics.

The hand-coded sample includes 6,171 hand-coded documents representing over three million comments (including both sophisticated comments and the mass comments that support them).

Table 5: A Sample of Hand-coded Public Comments

Docket ID	Coalition	Comment type	Organization	Position	Success
DOT-OST-2018-0068	AAAE	Org	Finnair Plc	Supports Rule	2
FEMA-2016-0003	AASA	Org	Bergen County	Opposes Rule	2
FEMA-2016-0003	AASA	Org	AASA	Opposes Rule	2
OCC-2020-0026	ACLU	Org	ACLU	Opposes Rule	-2
OCC-2020-0026	ACLU	Org	Hoosier Action	Opposes Rule	-2
PHMSA-2012-0082	AFL-CIO	Org	AFL-CIO	Supports Rule	-2
PHMSA-2012-0082	AFL-CIO	Org	City Of Skokie	Supports Rule	-2
ED-2016-OESE-0032	AFT	Org	AFT	Supports Rule	1
USCIS-2006-0044	AILA	Org	Centro Romero	Opposes Rule	-1
USCIS-2006-0044	AILA	Org	World Relief	Opposes Rule	0
WHD-2019-0001	ANCOR	Org	Easterseals	Supports Rule	-2
WHD-2019-0001	ANCOR	Org	Leadingage	Supports Rule	-2
USCG-2010-0990	BOATU.S.	Org	Boatu.s.	Opposes Rule	-1
DOT-OST-2018-0068	CCD	Org	Autism Speaks	Opposes Rule	-2
CFPB-2019-0006	CFSA	Elected	Nolan Mettetal	Supports Rule	2
CFPB-2019-0006	CFSA	Elected	Mike Osburn	Supports Rule	2
IRS-2019-0039	CLC	Elected	Steve Bullock	Opposes Rule	-2
IRS-2019-0039	CLC	Org	Issue One	Opposes Rule	2
ETA-2020-0006	CUPA-HR	Org	TECHNET	Supports Rule	1
USCBP-2007-0064	EAA	Org	Sas Institute	Supports Rule	1

Table 5 shows a sample of hand-coded public comments. Docket ID is the identifier for each rulemaking. The Organization, Comment Type, and Coalition columns show how coders record the name and type of each organization or elected official, as well as the broader coalition to which they belong. The name assigned to each coalition is usually the name of one of the lead organizations.

The Position column in Table 5 is a collapsed version of the spatial position coding described in Section 7.2. To create a binary measure of support and opposition, I collapse the coding of each comment's spatial position into a dichotomous indicator of whether they ultimately support or oppose the rule. Finally, Lobbying Success— whether each comment got what it asked for in the change between a draft and final rule—is coded on a five-point scale from 2 to -2. "2" indicates that most of the commenter's requests were met. If the rule moved decidedly in the opposite direction as they would have liked it to move, this is coded as a "-2" (the opposite of total success). To measure these variables at the coalition level, I use the coding assigned to the lead organization or the average across coalition members. Because "lead" organizations are identified based on their leadership role in the coalition and the extent to which they represent the coalition's policy demands, the lead organization's coding is nearly the same as the average across coalition members in all cases.

Table 6: Organizations by Number of Rules on Which They Commented in the Hand-coded Data

Organization	Rules Lobbied On
U.S. Chamber of Commerce	27
Natural Resources Defense Council	10
OCEANA	10
Earthjustice	9
National Audubon Society	9
Sierra Club	8
American Petroleum Institute	7
Center For Biological Diversity	7
Edison Electric Institute	7
National Association Of Manufacturers	7

Table 6 shows the organizations that commented on the most rules in this sample: 522 organizations lobbied on more than one rule in the hand-coded data, some on as many as -Inf rulemaking dockets. Recall that this sample of rules is weighted toward rulemaking dockets that received more comments. Thus, the organizations lobbying on the most rules are not the same as those in the overall population. For example, recall from Table 5

that the American Petroleum Institute lobbied on nearly 400 rules, whereas the Pew Charitable Trusts lobbied on 120. Pew, however, used a public pressure campaign 5 percent of the time it lobbied, whereas the American Petroleum Institute used a public pressure campaign 0.3 percent of the time it lobbied. Thus, groups like Pew that more often use pressure campaigns are more likely to be lobbying on rules in this sample. While this sampling approach was necessary (a random sample of all rules would yield almost none with a pressure campaign), the statistical results should be interpreted as disproportionately reflecting variation in lobbying success in high-salience and contentious rulemakings.

Table 7 shows the number of hand-coded rules, documents, the coalitions and organizations to which those documents belong, and the total number of comments they represent for a sample of agencies. As expected with a random sample, the agencies with the most rules in this sample are also those with the most final rules posted to regulations.gov (as shown in Figure 7). The Environmental Protection Agency (EPA), Fish and Wildlife Service (FWS), National Oceanic and Atmospheric Agency (NOAA), Department of Transportation (DOT), and Internal Revenue Service (IRS) are all in the top ten agencies by the number of rulemaking dockets on regulations.gov. The Bureau of Safety and Environmental Enforcement, Consumer Financial Protection Bureau (CFPB), and Wage and Hour Division (WHD) of the Department of Labor are all above average and have a disproportionate number of rules with a large number of comments, making these agencies more likely to be selected into the weighted sample. Table 7 also illustrates how my method of collapsing documents with repeated text to one representative document allows me to reduce the number of documents requiring hand-coding by several orders of magnitude (compare the "Documents" and "Comments" columns).

Table 7: Hand-coded Data By Agency

Agency	Rules	Documents	Coalitions	Organizations	Comments
NOAA	7	323	27	162	514,761
FWS	6	881	23	508	1,677,778
BSEE	4	243	13	195	114,368
CFPB	3	1003	18	817	12,982
DOT	2	174	9	111	174
IRS	2	53	4	46	29,897
OCC	2	122	6	106	267
USCIS	2	129	6	71	129
WHD	2	264	8	214	44,613
CEQ	1	850	1	577	411,383

Table 7: Hand-coded Data By Agency

Agency	Rules	Documents	Coalitions	Organizations	Comments
DEA	1	15	3	15	15
DOI	1	41	2	36	41
ED	1	318	18	298	318
ETA	1	55	6	44	55
FEMA	1	380	9	349	3,139
ICEB	1	303	4	65	303
MSHA	1	52	3	43	52
NPS	1	100	3	68	100
OFCCP	1	188	4	31	7,173
OSHA	1	161	4	69	161
PHMSA	1	401	4	364	335,313
TREAS	1	60	8	52	60
USCBP	1	37	2	30	37
USCG	1	12	2	10	12

Figure Figure 6) shows a massive rise in the number of proposed rules targeted by public pressure campaigns (the bottom panel), greater than the overall increase in the number of proposed rules posted for comment on regulations.gov (the top panel). To some extent, the increase from 2005 to 2010 resulted from agencies using regulations.gov more systematically in the years after its launch in 2003. However, the ease of online organizing has also increased the frequency of public pressure campaigns. As mentioned earlier, less than 5 percent of proposed rules each year are targeted by a pressure campaign (note the necessary difference in the y-axis). However, this share is growing.

Figure Figure 7) shows the handful of agencies that publish the majority of proposed rules for public comment on regulations.gov (out of the 248 federal agencies that use regulations.gov). For the most part, these are also the agencies most often targeted by public pressure campaigns, but some agencies are relatively more or less likely to be targeted than others. For example, the U.S. Fish and Wildlife Service publishes a small share of rules overall but a large share of rules targeted by public pressure campaigns (many protecting threatened and endangered species habitat). In contrast, the U.S. Coast Guard and Federal Aviation Administration both

publish a large number of rules (mostly regulating transportation safety), but pressure campaigns rarely target these agencies.

Figure 6: Proposed Rules Open for Comment on Regulations.gov 2005-2020
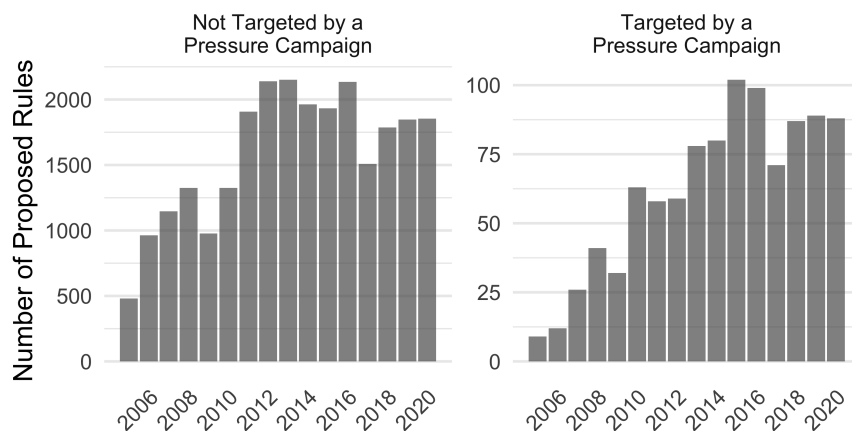

Figure 8 shows hand-coded support and opposition to proposed rules by different types of commenters and presidential administrations. Support and opposition coding come from the spatial position regarding the draft and final rule, as shown in Figure 4. Comments from a corporation ("Corp.") were overwhelmingly opposed to Obama-administration policies and more supportive of Trump-administration policies. Elected officials more often write in opposition than in support of a proposed rule across administrations. In contrast, individuals, organizations, and the mass comments these organizations mobilized overwhelmingly supported Obama-administration policies and opposed Bush- and Trump-administration. Mass and individual comments are especially polarized. This reflects the partisan asymmetry in mobilizing organizations; the individuals (unique comments) and mass comments (form letters) mobilized by progressive public interest groups' campaigns overwhelmingly supported Obama-era policies and opposed Trump-era policies.

Most of these comments belong to lobbying coalitions and are thus not independent observations. When Friends of the Earth and the Sierra Club lobby together on a rule, the success of each depends on the other. Thus, I group comments into coalitions. The hand-coded sample includes 256 "coalitions," 102 of which are single-organization "coalitions" (not coalitions), leaving 154 true coalitions of multiple organizations lobbying together.

Docket ID	Coalition	Position	Size	Businesses	Type	Mass
DOT-OST-2018-0068	A4A	Supports Rule	17	6	Public	0

Docket ID	Coalition	Position	Size	Businesses	Type	Mass
OCC-2020-0026	ACLU	Opposes Rule	75	0	Public	158
OFCCP-2014-0004	ACLU	Supports Rule	11	0	Public	7141
ICEB-2015-0002	AFL-CIO	Supports Rule	5	2	Public	0
USCIS-2010-0017	AFL-CIO	Opposes Rule	2	0	Public	0
ED-2016-OESE-0032	AFT	Supports Rule	3	0	Public	0
TREAS-DO-2007-0015	AHC	Supports Rule	1	0	Private	0
DEA-2018-0005	AMA	Supports Rule	3	3	Public	0
CFPB-2016-0025	CBC	Supports Rule	1	0	Public	0
ETA-2020-0006	CUPA-HR	Supports Rule	6	4	Public	0
WHD-2019-0003	EPI	Opposes Rule	40	1	Public	44352
DEA-2018-0005	HSCA	Opposes Rule	9	3	Public	0
NA	HSCA	Opposes Rule	1	0	Public	0
DOT-OST-2018-0068	IAADP	Supports Rule	12	1	Public	0
OCC-2011-0014	ISDA	Supports Rule	1	1	Private	0
ETA-2020-0006	ITSERVE	Opposes Rule	15	9	Private	0
OCC-2020-0026	MLA	Supports Rule	13	12	Private	0
DOT-OST-2018-0068	N/A	Opposes Rule	1	1	Private	0
USCIS-2010-0017	NCAPA	Supports Rule	14	5	Public	0
WHD-2019-0001	NCPA	Supports Rule	16	9	Private	0
FEMA-2016-0003	NCSL	Opposes Rule	2	0	Public	0
CFPB-2019-0006	NCUA	Supports Rule	5	3	Private	0
ETA-2020-0006	NFAP	Supports Rule	5	2	Public	0
FWS-R9-ES-2008-0093	NMBA	Supports Rule	2	2	Public	0
FWS-HQ-NWRS-2012-0086	NRDC	Supports Rule	13	0	Public	53258
USCIS-2006-0044	NYCLA	Supports Rule	3	0	Public	0
NOAA-NMFS-2011-0117	OCEANA	Supports Rule	5	0	Public	12882
BSEE-2012-0005	PEW	Supports Rule	4	0	Public	1526
BSEE-2013-0011	PEW	Supports Rule	15	0	Public	24911
FEMA-2016-0003	PEW	Supports Rule	217	201	Public	2765

Table 9

Coalition Type	Business led	Non-business led	Mass comments	No mass comments
Private	78	20	11	87
Public	16	140	35	123

Types of Lobbying Coalitions in the Hand-coded Sample

Docket ID	Coalition	Position	Size	Businesses	Type	Mass
NOAA-NMFS-2008-0096	PEW	Supports Rule	17	2	Public	124675
NOAA-NMFS-2012-0059	PEW	Opposes Rule	31	3	Public	102845
NOAA-NMFS-2013-0050	PEW	Opposes Rule	8	1	Public	75238
DEA-2018-0005	PHRMA	Supports Rule	3	2	Public	0
BSEE-2013-0011	SHELL	Opposes Rule	90	78	Private	87690
FEMA-2016-0003	SSIT	Opposes Rule	1	0	Public	0
IRS-2016-0015	SWISS RE	Supports Rule	1	1	Private	0
CFPB-2019-0006	TRUE	Opposes Rule	1	0	Public	0
TREAS-DO-2007-0015	UC GROUP	Opposes Rule	1	1	Private	0
ICEB-2015-0002	VERIZON	Supports Rule	3	1	Public	0
WHD-2019-0001	WFCA	Supports Rule	54	38	Private	0
CFPB-2019-0022	WSECU	Supports Rule	1	0	Public	0
TREAS-DO-2007-0015	XPRESSBET	Opposes Rule	1	1	Private	0

A Sample of Hand-coded Data Summarized by Coalition

Lobbying coalitions range in size from 2 to 217 organizations. **?@tbl-data-coalition-comments** shows a sample of coded data summarized at the coalition level. Even though the same organization may lead coalitions in multiple rulemakings, each rule's lobbying coalitions are different, so I consider them separate observations. For example, the American Civil Liberties Union (ACLU) led a coalition in 2014 with a small number of organizations and a medium-size pressure campaign in support of a rule requiring additional Equal Employment Opportunity reporting from the Department of Labor's Office of Federal Contract Compliance Programs (OFCCP). The ACLU also led a very different coalition in 2020 with a large number of organizations and a very small public pressure campaign against a rule rolling back regulations on banks published by the Office of the Comptroller of the Currency (OCC).
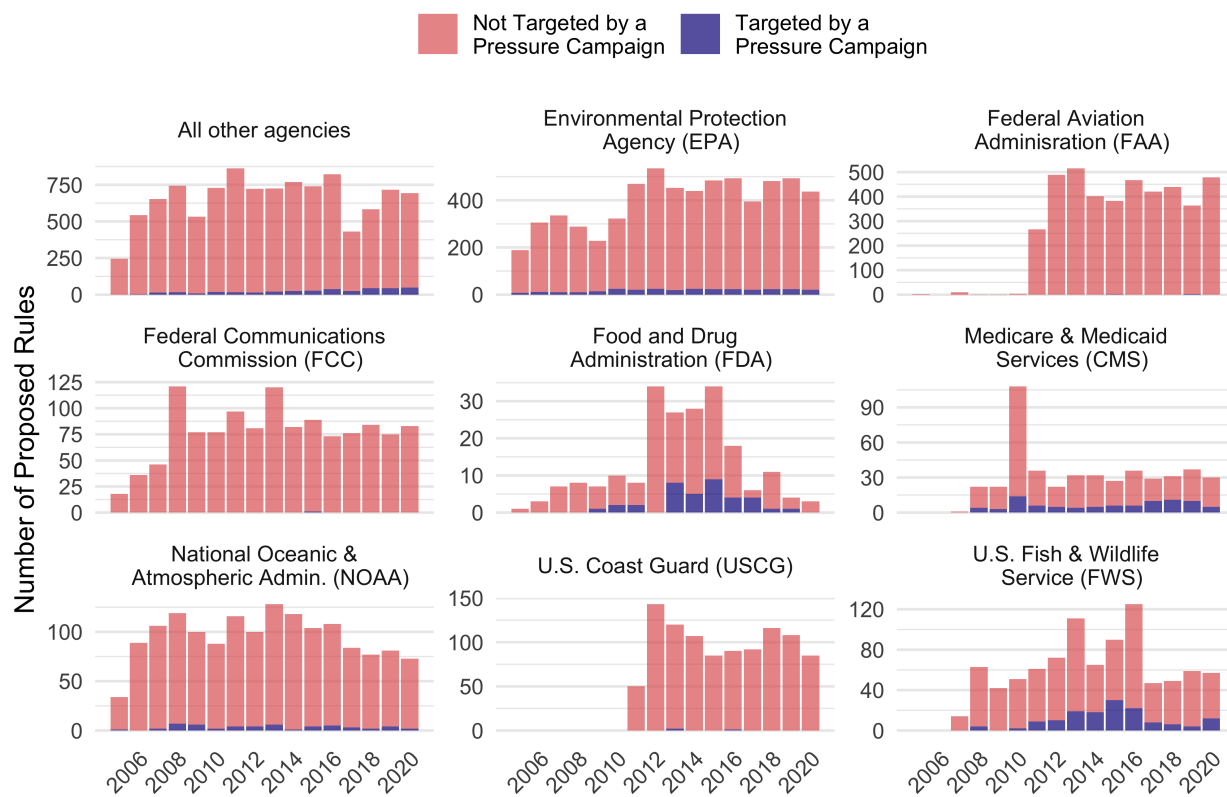
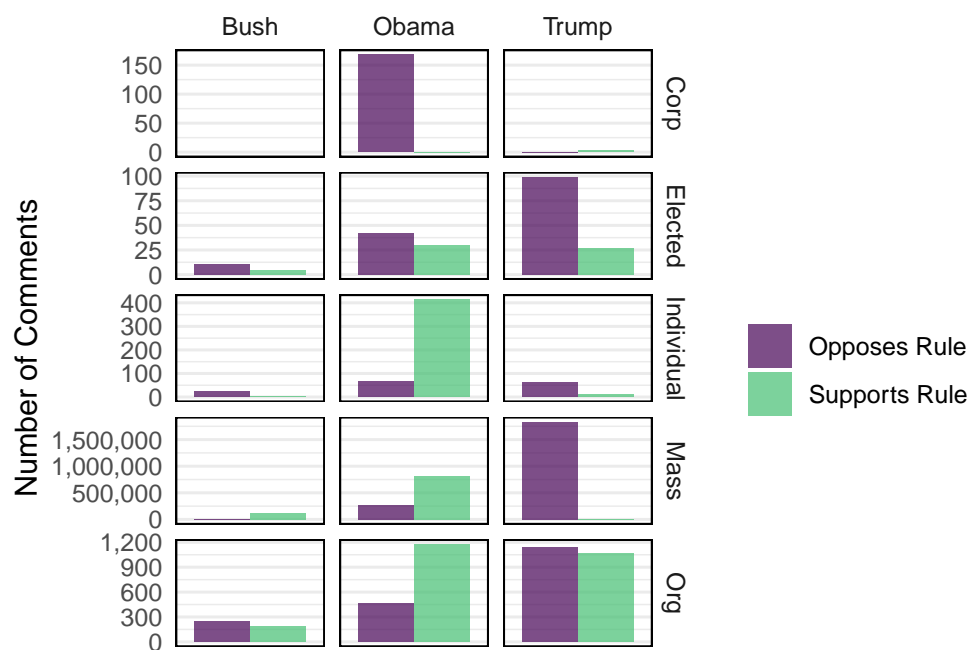**Figure 7:** Proposed Rules Open for Comment on Regulations.gov 2005-2020 by Agency

Figure 8: Hand-coded Comments By Type and Position on Proposed Rule

Table 9 shows the number of coalitions coded as "public interest" and "private interest" by whether the majority of organizations in the coalition are for-profit businesses and trade associations or non-businesses (governments and nonprofits): 37 percent are majority business coalitions. 62 percent are public-interest coalitions. As Table 9 shows, the hand-coded "public interest vs. private interest" distinction is highly correlated with the share of businesses in the coalition but not perfectly. These two measures diverge in cases where public interest coalitions mobilize a large number of business allies or where private interest coalitions mobilize a large number of non-business allies. Thus, while the share of businesses and trade associations is more objective, the public-private distinction is likely a better measure of coalition type.

Several coalitions may lobby on the same rule. One coalition's lobbying success is correlated with another coalition's lobbying success to the extent that they are asking for the same or contradicting policy changes. However, by grouping organizations into coalitions, I account for many of the causally-related policy requests (those organizations lobbying on an issue *because* another organization is lobbying on that issue).

8.1.3   Comments from Legislators Correlate with Public Pressure

One mechanism by which campaigns may influence policy is by mobilizing members of Congress. Thus, I count the number of legislators in each lobbying coalition. Figure 9 shows the number of comments from members of Congress received during rulemaking by a sample of federal agencies. There is massive variation in

the level of attention that members of Congress pay to different agencies and rules. The spikes in attention to each agency correspond with public pressure campaigns targeting rules from that agency. Oversight letters are frequently co-signed by multiple members from the Senate, House, or both chambers. Some of the rules on which members of Congress commented appear in the hand-coded sample. Table 10 shows the number of comments from the most common types of elected officials in the hand-coded data. Members of the U.S. House and Senate are the most common.

Figure 9: Number of Rulemaking Comments from Members of Congress per Year, 2005-2020 to Nine Federal Agencies
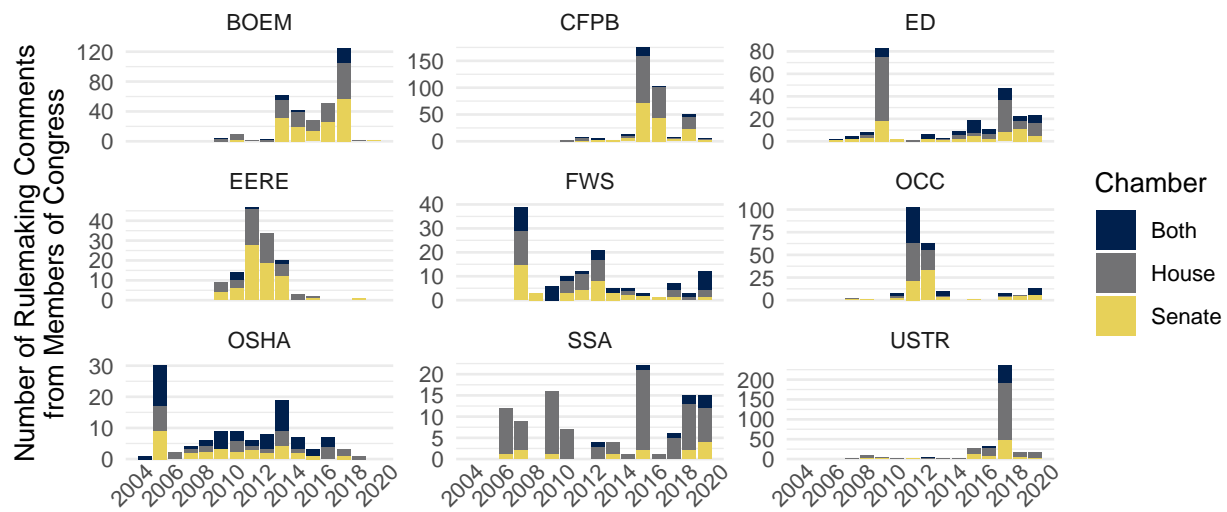
Comments received by the Bureau of Ocean Energy Management (BOEM),
Consumer Financial Protection Bureau (CFPB), Department of Education (ED),
Office of Energy Efficiency and Renewable Energy (EERE), Fish and Wildlife Service (FWS),
Office of the Comptroller of the Currency (OCC), Occupational Safety and Health Administration (OSHA),
Social Security Administration (SSA), and U.S. Trade Representative (USTR)

Table 10: Comments from Elected Officials in the Hand-coded Data

Elected Official Type	n
Senate	81
State Elected Official	56
House	53
Local Elected Official	14
Governor	13

8.1.4   The Dependent Variable: Lobbying Success

The dependent variable is the extent to which a lobbying coalition got the policy outcome it sought, which I measure in several ways.

First, on a sample of rules, I trained a team of research assistants to hand-code lobbying success for each organization or elected official, comparing the change between the draft and final rule to each organization's demands on a five-point scale from "mostly as requested" to "significantly different/opposite from requested direction" as described in Section 7. Additionally, for each comment, coders identify the main overall policy demand, the top three specific demands, and the corresponding parts of the draft and final rule texts. This does not capture rule changes on which no organization commented. While such changes may be significant, they are not helpful for measuring lobbying success.

Lobbying success on each specific demand was then coded for each organization and coalition. Both the overall score and average score across specific demands both fall on the interval from -2 ("significantly different") to 2 ("mostly as requested"). A team of undergraduate research assistants then applied the codebook to all comments likely to be from organizations or elected officials on a random sample of rules. Several rules were double-coded by the whole team.

In the models below, *coalition lobbying success* is the mean of hand-coded lobbying success on a five-point scale, {-2, -1, 0, 1, 2}.

The average hand-coded success per organizational comment is zero (N = 6,171). The average success for organizational comments associated with a mass comment campaign is -0.746 (N = 3,857).

### 8.1.5   Coalition size and coalition success

The number of supportive comments generated by a public pressure campaign (the main variable of interest) is a tally of all comments mobilized by each organization or coalition that ran a mass-comment campaign on a proposed rule. Because the marginal impact of additional comments likely diminishes, models typically include either the logged number of comments or a quadratic term to account for non-linear effects. If a coalition mobilizes more than 99 form-letter comments on a proposed rule, I code that coalition as having a mass comment campaign (*campaign* = 1). Where a coalition only submits technical comments from lawyers and does not mobilize public support, the binary measure, *campaign*, and the numeric measure, *mass comments*, are 0.

2) shows a scatterplot of the dependent variable (*lobbying success*) and main predictor (*mass comments*) for each coalition. Coalition lobbying success ranges from total success (2) to total loss (-2). The number of mass comments ranges from 0 to 701,917. The size of each point represents the size of each coalition (the

number of organizations and elected officials). The color indicates whether the coalition is led by private or public interest groups. For example, one extremely large private coalition of payday lenders mobilized over a million comments during the Obama administration. This coalition was moderately successful at reducing the stringency of the regulation but did not stop it from going through.

The view of the data in 2 does not show a clear relationship between public pressure and lobbying success. There were relatively more (and more successful) public interest campaigns in the Obama years. Likewise, there were more (and more successful) private interest campaigns in the Trump years. The largest campaigns are mostly public interest campaigns, and public interest campaigns are more frequent than private interest campaigns overall.

Figure 10: Lobbying Success by Number of Supportive Comments
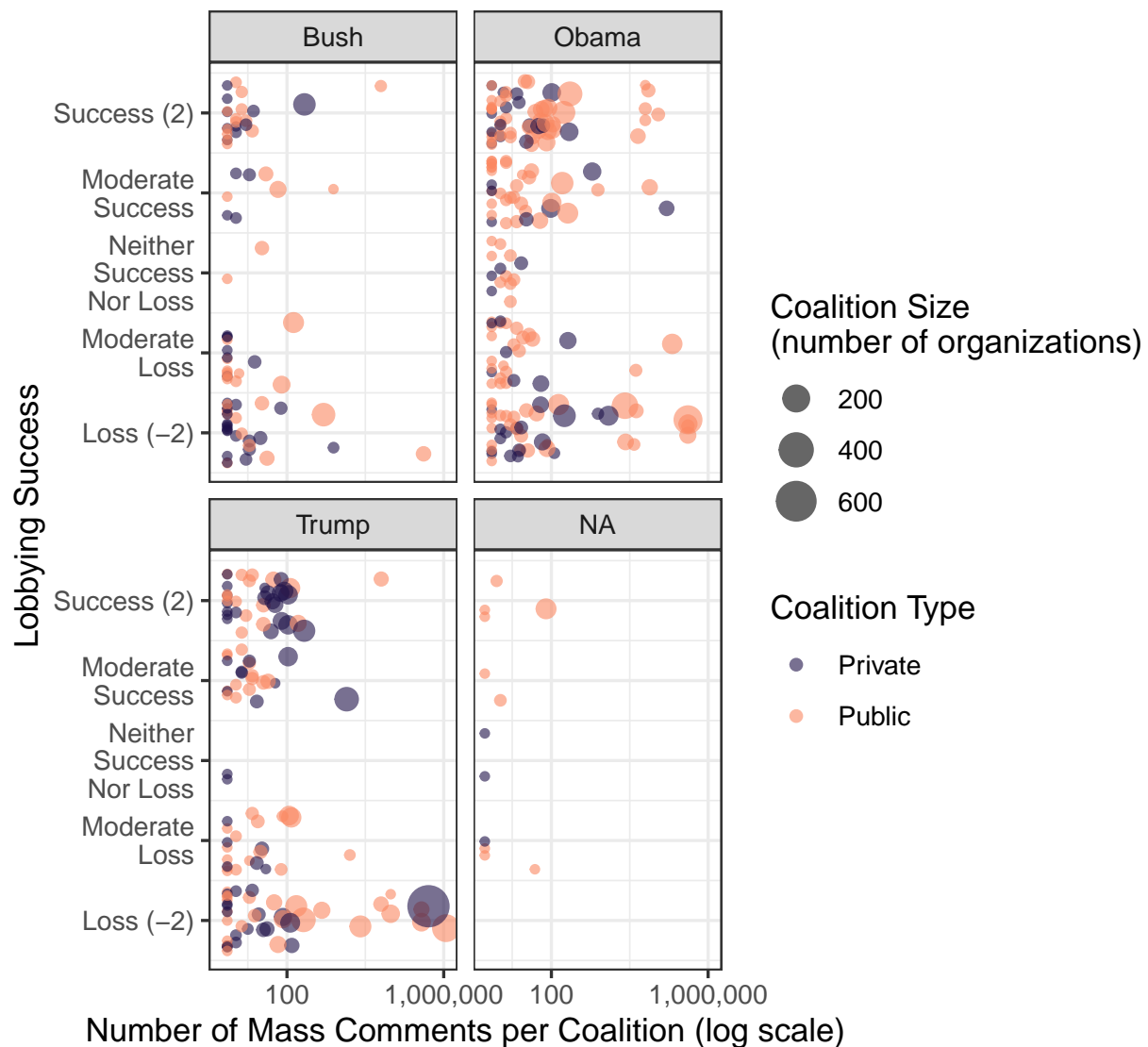
9 Conclusion

Because hand-coding is costly, researchers often turn to unsupervised or machine-learning methods to scale, cluster, or tag large text corpora. In some cases, however, the methods outlined here may yield better results since they do not rely on probabilistic inferences. Instead, exact matching with text reuse and other features go a long way toward extending the power of a small hand-coded sample.

Because politics is organized and organizing power is often concentrated among a few key actors, researchers studying a particular policy fight (or set of policy fights) may be best served by using powerful text analysis tools to identify key actors and those associated with them.

This paper has focused on the advantages of iterative workflow using hand-coding and exact matching tools. Parallel workflows using iterative human coding and machine learning tools likely offer similar advantages, and future work should compare the relative strengths of different approaches to "human-in-the-loop" workflows.

Extending political science intuitions about the organizational nature of politics, a natural next step for studying lobbying coalitions is to integrate network analysis tools with exact matching. Public pressure campaigns are coordinated by a relatively small number of organizations that repeatedly lobby both with and without pressure campaigns. The tools reviewed in this paper help link organizations' advocacy behavior over time, even as they use different monikers and join different coalitions. Lobbying coalitions range from two organizations to hundreds. These data are flush with opportunities for network analysis using organizations as nodes and coalitions as edges, and the methods outlined above enable such research. Simple measures like node centrality may tell us a great deal about the structure of advocacy coalitions and how they change over time and across policy areas.

Finally, a major advantage of the iterative methods and workflow is that they are fairly simple tools that can be adopted piecemeal by researchers to increase the inferential power of hand-coding. Some of these methods—like a regex table to detect U.S. legislators—are already available in easy-to-use R packages that may help researchers easily sift through large corpora to identify key observations (e.g., those talking about legislators) or may save time by pre-populating data sheets with information that can be quickly collected by machines. The allocation of work between humans and machines will vary with every research project, but iteratively combining the strengths of each will almost always be the most efficient and powerful approach.

9.1 References

Balla, S. J., Beck, A. R., Meehan, E., and Prasad, A. (2020). Lost in the flood?: Agency responsiveness to mass comment campaigns in administrative rulemaking. *Regulation & Governance*.

Casas, A., Denny, M. J., and Wilkerson, J. (2019). More Effective Than We Thought: Accounting for Legislative Hitchhikers Reveals a More Inclusive and Productive Lawmaking Process. *American Journal of Political Science*, 64:5–18.

Enamorado, T., Fifield, B., and Imai, K. (2017). fastlink: Fast probabilistic record linkage with missing data.

Hacker, J. and Pierson, P. (2010). *Winner Take All Politics*. Simon and Schuster.

Judge-Lord, D. (2017). Measuring Change and Influence in Budget Texts. American Political Science Association Annual Conference.

Potter, R. A. (2017). More than Spam? Lobbying the EPA Through Public Comment Campaigns. *Brookings Series on Regulatory Process and Perspective*.

Rinfret, S., Duffy, R., Cook, J., and Onge, S. S. (2021). Bots, fake comments, and e-rulemaking: the impact on federal regulations. *International Journal of Public Administration*, 0(0):1–9.

Saveourenvironment.org (2021). https://saveourenvironment.org/who-we-are/.

Verba, S. and Nie, N. H. (1987). *Participation in America: Political Democracy and Social Equality*. University of Chicago Press.

West, W. F. and Raso, C. (2013). Who Shapes the Rulemaking Agenda? Implications for Bureaucratic Responsiveness and Bureaucratic Control. *Journal of Public Administration Research and Theory*, 23(3):495–519.

Yackee, J. W. and Yackee, S. W. (2006). A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy. *Journal of Politics*, 68(1):128–139.

Appendix

A   Real-time quality control

Figures Figure A1 and Figure A2 illustrate some advantages of real-time analysis of hand-coded data. Figures Figure A1 show an example of how one might inspect overall patterns among coders (e.g., how many letters to the Federal Energy Regulatory Commission referenced a specific company). Figure A2 shows a more mundane but essential quality check—looking for cases where coders used incorrect special symbols. Project-specific checks like this can be assembled into a dashboard to track and inspect hand-coding progress.[9]

Figure A1: Checking for Disparities Among Coders in Real Time

```
# differences in pro and anti coding
FERC_coded %>%
  mutate(company_identified = !is.na(ProBusiness) | !is.na(Anti
Business)) %>%
  group_by(Freelancer, company_identified) %>%
  count() %>%
    spread(Freelancer, n) %>%
  knitr::kable()
```

company_identified	Gabriela Iocab	Jobelle	Lara Basiao	Maki Ianovic	Priyal Singh
FALSE	295	108	932	551	620
TRUE	399	192	427	441	374

B   Comparing Hand-coding to GPT-4

B.1   Method

As a first step in evaluating the potential for Large Language models (e.g., OpenAI's GPT) to aid these research tasks, I compare hand-coded comments to the output that GPT-4 returns when provided with the same inputs as undergraduate RAs: a proposed rule text, a final rule text, comment text, comment metadata, and a codebook with instructions of for recording they comment author, the organization they represent, their position on

---

[9]An example dashboard is available here: https://judgelord.github.io/correspondence/FERC/DOE_FERC-letter-coding.html

Figure A2: Checking for Incorrect Coding in Real Time

```
# Find problems with "/" instead of ";" sep
d %>% select(ID, ProBusiness, AntiBusiness, Freelancer) %>%
  filter(str_detect(AntiBusiness, "/") | str_detect(ProBusines
s, "/") ) %>%
  knitr::kable(caption = "Suspicious backslash use to look at")
```

Suspicious backslash use to look at

ID	ProBusiness	AntiBusiness	Freelancer
20180216-0016	NA	AmerenUE/AmerenMO	Gabriela Iocab
20171108-0007	FFP Missouri/Rye Development	NA	Gabriela Iocab
20151118-0010	NA	Transco/Williams	Gabriela Iocab
20150520-0074	NA	NIPSCO/USFWS	Lara Basiao
20140723-0052	Tranco / Williams	NA	Lara Basiao
20080201-0218	NA	Broadwalk Pipeline LLC / Gulf Crossing LLC; ABL Land Services	Priyal Singh

the proposed rule, their specific policy demands, and their success in achieving those policy changes in the final rule. It is likely that GPT will perform better with instructions optimized for LLM input (i.e., "prompt engineering").[10]

A Python script to construct GPT input ("prompts") from the same file structure described in Section 4.

To make the results easy to compare, GPT is prompted to return the result for each item in the codebook separated by commas, which are saved as a CSV file. These CSV files can then be stacked to make a spreadsheet identical to Figure 1. The result is two identically-formatted spereadsheets, with one row per comment and columns for each variable in the codebook.

Systematically comparing results requires a third spreadsheet that matches the other two, except that each cell now characterizes the similartiy or difference between the corresponding cells in the human coder and GPT's spreadsheet. For variables that are quantitative (e.g., the spatial `position` coding described in Section 7.2) and discrete-choice (e.g., the type of organization), comparing human coder and GPT output can be auto-populated. For example, if the human and GPT both classified a comment as from a business that preferred the status quo, the `org_type` and `position` columns for that comment are assigned a 1 for a perfect match. Otherwise, they get a 0. Comparing open-ended output such as the comments policy demands, requires a new codebook for human-coders to characterize the difference.

B.2   Results

B.2.1   Research tasks that GPT does well

GPT was surprisingly effective at identifying comment authors and the names of organizations mentioned in comments. This is important because metadata such as the comment author and the organization they represent is missing from many comments on regulations.gov. Some agencies systematically collect this information, and some do not. Regulations.gov generally allows submitters to input whatever they like, making metadata messy and requiring entity-tagging methods like those discussed in Section 4.2. GPT may help further automate these tasks. However, I have yet to test how well the names GPT extracts align with the organization's official names (e.g., from IRS records). It would be extremely expensive to give GPT hundreds of thousands of possible organization names and ask it to match them. Other exact or probabilistic matching methods are likely better suited to this task.

GPT's classification of commenters as businesses, nonprofits, government entities, elected officials, or other individuals generally matched human coders. GPT's classification of types of nonprofits (trade associations, unions, advocacy organizations, etc.) and types of government entities (cities, counties, states, etc.) generally

---

[10]Because the inputs are large, it is necessary to use GPT-4, which has more memory, rather than GPT-3.5 Turbo, which is much cheaper (~2 cents per comment rather than .2 cents per comment).

matched human coders as well. This type of classification seems like a good task for GPT as long as the prior stage of identifying the comment author is done correctly.

B.2.2   Research tasks that GPT does poorly but better than other automated methods

Like other machine learning methods, GPT tended to classify comments as supporting the rule if they used positive-sentiment language and opposed if they used negative-sentiment language. However, initial comparison appear to show GPT doing a better job at detecting when positive or negative sentiments were targeted at the status quo rather than the new proposed policy and thus correctly classifying comments as supporting or opposing change.

When asked to return policy demands in a comment, GPT generally returned plasuble demands that matched human coders significantly better than the sentences selected by the TextRannk algorithm.

B.2.3   Research tasks that GPT does poorly

GPT generally failed to capture changes to the draft and final rules. When asked to identify changes, GPT always returned a description of changes that mirrored those highlighted by the agency in the final rule preamble and did not detect changes that were not mentioned in the preamble.

Because GPT relied on the agency's description of changes between the draft and final rule, the results for commentate success were, at best, related to how the agency discussed the contentment's concerns. More often, GPT results for commenter success were not based on the change between the draft and final rule but just on the relationship between the words in the comment and the words in the final rule. GPT returned high values of success when commenters requests were mirrored in the final rule (even if that language already existed in the draft rule) and low values of success when the commenter used wording that differed from the wording of the final rule (even where the agency made changes in response to the commenter's demands).