Decoding Influence: A Text-Based Comparative Study of Business and Labor Interest Group Power in U.S. Labor Rule-making

Saniya Jain

Introduction

Regulations may carry the weight of law, but their final shape is often carved by those who know how to speak the agency's language: interest groups. The formulation of public policy in the United States is not solely the domain of elected officials and public servants but also that of interest groups, ranging from business associations to labor unions, that play a critical role in shaping policy outcomes, particularly through their engagement in the rulemaking process. Lobbying, coalition-building, and public comment submissions are among the key mechanisms these groups use to assert influence over regulations that affect their constituencies (W. D. Costain et al). However, the extent and nature of interest group influence on policy-making, especially in the bureaucratic rule-making process, remain deeply contested and understudied in empirical terms (Yackee and Yackee 128).

This study focuses on labor policy in the U.S., an arena that not only endures conflict between organized labor and business interests but also remains underexplored in policy influence studies, with most studies focusing on environment, energy, and financial regulations (Yackee and Yackee 129).

Recent studies have demonstrated that elite business interests dominate policy discourse, especially in conventional media outlets. In an analysis of 7,747 news stories across 80 policy issues, corporations and business groups were cited in 36.5% of all mentions, while labor unions accounted for just 13.2%, despite representing a large and historically influential segment of the workforce (Danielian et al.). This disproportionate media visibility raises critical questions: If labor groups are marginalized in public discourse, how else might they exert influence over the policymaking process? And are they successful in doing so?

Labor policy affects millions of workers, shaping workplace safety, wages, and collective bargaining rights. The US has always had a history of strong unions protecting workers from exploitation, unfair labor practices, lack of benefits, and occupational hazards in the face of contradictory beliefs and interests from businesses. While today, 11.1% of workers are unionized, union membership has been on a steady decline and most Americans believe this is damaging for the economy (Pew Research Center). And since a large measure of union strength comes from its ability to achieve policy objectives in the face of opposition, the unpopularity of unionization calls into question the power of labor unions and their advocacy strategies in policy arenas.

In this analysis, I compare business with labor because labor is also a large, organized interest group that frequently stands in opposition to business interests. Understanding the extent of labor's influence is essential to identifying what strategies and resources unions can leverage to safeguard the welfare of American workers.

This study examines how business and labor interest groups seek to influence labor regulations proposed by three key U.S. federal agencies tasked with protecting worker rights: the Equal Employment Opportunity Commission (EEOC), the Occupational Safety and Health Administration (OSHA), and the National Labor Relations Board (NLRB). These agencies create rules through a formal notice-and-comment process, where organizations and individuals are invited to submit feedback on proposed regulations. This provides a valuable opportunity to assess how different groups communicate their policy positions and whether those positions are reflected in final rule outcomes.

At the center of this inquiry is the question of how to measure “influence.” Influence is inherently difficult to observe directly, especially when it is diffuse, subtle, or structural. As Charles Lindblom argued, businesses hold structural power that makes them uniquely influential, even in the absence of overt lobbying or campaign contributions. Legislators and regulators may favor business interests not merely due to political alignment but because of an underlying assumption that business cooperation is necessary for economic stability and political legitimacy (Ambrosius et al.). This study asks: To what extent do business interest groups exert influence in regulatory settings that are meant to oversee them? Do they “win” more often than labor groups, and if so, under what conditions?

To investigate these questions, I employ a mixed-methods approach. First, I use quantitative text analysis, specifically the Wordfish algorithm in R, to estimate and compare the ideological positions of interest group comments and the language of the final rule. This allows me to systematically assess which groups' preferences were more closely aligned with the final policy, a proxy for influence. Then, I apply Qualitative Comparative Analysis (QCA) to six regulatory cases, identifying combinations of factors, such as rule salience, comment quality, and resource availability, that contribute to the success or failure of each group in shaping policy outcomes.

This thesis is structured into eight different sections, starting with a literature review discussing past methodologies and current research in influence studies, followed by a detailed breakdown of my research hypothesis, a holistic and critical analysis of the methodologies employed, and a discussion of data sources. The last few sections present the empirical findings of the analysis,

comparing the relative success of business and labor groups across cases, and discuss the implications of these findings for our understanding of structural power, regulatory capture, and democratic accountability. And finally, concludes by summarizing key contributions and offering directions for future research.

Literature review

Understanding how interest groups influence policymaking has long been a central concern in political science, particularly in contexts where policy outcomes emerge through bureaucratic, rather than legislative, channels. Scholars have developed multiple theoretical and empirical strategies to measure influence, broadly categorized into process-tracing, preference attainment, attributed influence, and quantitative text analysis.

One of the most challenging aspects of such a study is that influence is hard to define: power can mean control over outcomes, agenda-setting (keeping issues off the table), or shaping preferences, the "third face" of power, as Andres Dür would say (Dür 2008).

Measuring Interest Group Influence

Below, I expand on the diverse empirical methods researchers have employed to assess interest group influence:

1. Process Tracing and Qualitative Analysis

Scholars like Andreas Dür (2008) have emphasized mapping out the causal chain from lobbying activities to final policy outcomes using interviews and document analysis. This involves tracing. While rich in detail, this method is time-intensive and difficult to scale.

2. Preference Attainment Models

   This approach, advanced by Dür and Klüver, measures influence by estimating how closely policy outcomes align with interest group preferences. If a final rule mirrors a group's ideal point more closely than others', that group is considered to have exerted more influence (Klüver 2009; Dür 2008). This strategy forms the backbone of this study's methodological framework.

3. Attributed Influence Surveys

   These rely on self-reported or expert-perceived influence, as seen in the work of Dür and others. While useful for gauging perceptions, such surveys often reflect bias and cannot confirm actual policy impact.

4. Quantitative Text Analysis

   Innovations in automated text analysis tools like Wordfish and Wordscores have enabled researchers to quantify the alignment between public comments and final rules. Klüver (2009) uses these tools to assess EU policymaking, while Yackee & Yackee (2006) apply manual content coding to U.S. labor regulations. Both studies find that business groups often exert disproportionate influence, due in part to their capacity to provide detailed, technical, and often legally strategic comments during the public consultation phase.

Scholars like Dür and De Bièvre emphasize that interest groups can influence policy through various channels, including direct access to decision-makers, influencing who gets selected into

power, shaping public opinion through voice strategies, or exerting structural economic power. However, measuring this influence is notoriously difficult due to challenges in distinguishing true preferences from strategic bargaining positions, accounting for counter-lobbying, and capturing how public salience or preferences may shift during deliberation (Dür 2008).

In response to these challenges, they suggest the preference attainment method, which allows the measure of influence to be concretized as the alignment between interest group preferences and final policy outcomes. The idea is that the distance between an outcome and the ideal point of an interest group reflects the influence of this interest group.

Dür argues this approach is effective because it directly compares policy positions to outcomes, revealing influence even when lobbying activities are not visible (Dür 2008) .

Yackee and Yackee employ the preference attainment methodology in their rigorous study measuring actual interest group influence using systematic coding and modeling. They analyse 40 rules and 1,693 public comments, hand-coding each to control for quality of comments and analyzing which groups' interests were reflected in the final policy outcome (Yackee and Yackee). Their analysis finds that business interests disproportionately influence final regulatory outcomes and that agencies are more likely to alter rules in line with the preferences of business commenters, particularly when business submissions make up a large share of total comments. Interestingly, their findings suggest that comment quality alone does not account for this bias, instead, it is the volume and cohesion of business input that seem to sway agency decisions.

This finding builds on earlier studies, such as Golden (1998), who found that while business groups are the most frequent participants in the comment process, she did not detect clear influence patterns, likely due to the limitations of her small sample. The Yackee study improves

upon these earlier works by incorporating a larger sample size, employing robust coding methods, and controlling for variables like rule salience and complexity.

Heike Klüver (2009) advances the same preference attainment approach by employing quantitative text analysis to measure interest group preferences systematically. Using methods like Wordfish and WORDSCORES, Klüver analyzes the degree to which policy texts align with interest group submissions during EU policymaking consultations. This method is not without its challenges, it can be difficult to determine preferences, control for alternative factors explaining a coincidence between preferences and outcomes, or even disaggregate a political issue into more specific issues to be able to code the success of an interest group (Andreas Dür et al).

Much like Yackee and Yackee and Klüver, my thesis adopts a preference attainment approach, focusing on the notice-and-comment rulemaking process, a particularly observable and text-rich avenue of policy influence. By analyzing public comments submitted by interest groups and comparing them with changes in the final regulatory texts, I use quantitative text analysis to infer whose preferences are reflected in outcomes, offering a transparent and replicable way to assess influence.

In their study, Yackee and Yackee view business comment volume and cohesion along with comment quality as variables affecting the magnitude of influence. They argue that well-reasoned, evidence-rich comments may increase the likelihood of influence, especially when submitted by multiple, like-minded business actors. Furthermore, Andreas Dür (2008) synthesizes the literature by proposing four core determinants of influence: resources (e.g., money, expertise, legitimacy), political institutions (e.g., the EU's complex, multilevel structure),

issue characteristics (e.g., technicality, salience, distributive vs. regulatory), and interest group strategies (e.g., lobbying at national vs. EU level).

Reflecting this literature, in my QCA, I included comment quality and the presence of citations as proxies for the informativeness and perceived legitimacy of stakeholder input. I also incorporated variables for resource level (to capture organizational capacity), issue salience (to capture variation in rule visibility), and interest group type (labor vs. business) to account for strategic and structural positioning within the regulatory process.

These factors together help draw out patterns and combinations of conditions that form influence strategies.

Hypotheses

In this paper, I hypothesize that business interest groups are more likely to achieve their objectives in labor regulations in the United States than labor interest groups. Specifically, I argue that business associations and corporate actors exert more influence during the notice-and-comment rulemaking phase of bureaucratic policy-making on labor issues than organizations such as labor unions, which primarily advocate for workers' rights.

H1: Business interest groups are more likely to see their preferences reflected in the final text of labor regulations than labor interest groups.

This hypothesis builds on the findings of Yackee and Yackee (2006), who, through a content analysis of over 1,600 comments on federal rules, find that business preferences were more

consistently aligned with final regulatory outcomes, even after controlling for comment quality. Their results suggest a pro-business bias in U.S. bureaucratic rulemaking, one that is not explained by superior arguments or higher-quality engagement. This aligns with a broader body of research showing that agencies often cater more to business groups than to diffuse interests such as labor or public interest organizations.

H2: Groups with greater resources and organizational capacity are more likely to achieve preference attainment in final rules.

Following the rational choice models discussed in both Dür and De Bièvre (2007) and Yackee and Yackee (2006), resource-rich actors are better positioned to engage in sustained lobbying, provide extensive comments, and maintain relationships with regulators. Business groups typically out-resource labor unions in terms of funding, professional staff, and lobbying infrastructure, enabling more effective participation in rulemaking.

H3: The quality of a comment positively correlates with influence, but does not fully explain business dominance.

While the Yackees assessed comment quality through proxies like expertise, length, and inclusion of attachments, they found that business comments were not systematically of higher quality than those from other groups. This suggests that while quality may matter, structural factors likely give business interests an edge regardless of argument strength.

This is further supported by Charles Lindblom's theory of structural power that provides a foundational explanation for these patterns. Lindblom (1977) argued that business occupies a privileged position in policymaking because policymakers perceive business cooperation as essential to economic stability and their own electoral survival. Even absent direct influence, the threat of business disinvestment or market instability compels legislators and regulators to favor business interests. This built-in advantage results in a system where business preferences are disproportionately reflected in policy outcomes, even in regulatory domains such as labor, where one might expect stronger labor representation.

Methods

There are several methods to assess interest group influence. Many of them are based on understanding legislators' perceptions of interest groups, or examining how and when legislators credit interest groups in pushing the needle in their decision-making. Some methodologies also look at the frequency of interactions between legislators and lobbyists, without being able to determine the true nature of the exchange, to assess influence (W. D. Costain et al).

Many of these methodologies require extensive data collection, scraping public opinion platforms, news channels, and news outlets to determine the prevalence of interest groups in the political space. Often, they examine subjective data based on legislators' judgment and perception of interest groups, which may not provide a clear picture of final policy outcomes in a testable way (W. D. Costain et al).

I have used a mixed methodology aligned with the Preference Attainment strategy employed in measuring interest group influence. Utilizing this, I measure the alignment between interest group preferences and final policy outcomes, assuming that the distance between an outcome and the ideal point of an interest group reflects the influence of this interest group. It is also assumed that preferences are publicly stated, wholly captured in comments, and stable during the comment period.

In this study, I collected public comments submitted by labor unions, business associations, nonprofits, and other actors on six proposed labor regulations issued by the Occupational Safety and Health Administration (OSHA), the National Labor Relations Board (NLRB), and the Equal Employment Opportunity Commission (EEOC) between 2010 and 2023.
For each regulation, I analyzed the proposed rule (baseline position), the final rule (outcome), and a selection of representative comments, focusing on prominent labor and business groups, as well as a mix of smaller nonprofits and independent commenters. Prominent labor unions and business groups were chosen based on their membership volume and general perception as the lead advocates in their respective spaces. These include organizations like AFL-CIO, UFWC, SEIU, NFIB, USCC, and NAM, that remain prominent in their respective spaces as high-resource, politically active organizations.

Choice of Methods

To estimate the ideological distance between interest group preferences and regulatory outcomes, I use Wordfish, a statistical scaling algorithm for unsupervised quantitative text analysis. Wordfish estimates the position of each document on a latent policy dimension based solely on word frequencies, without requiring human-coded reference texts or pre-defined categories. It

allows me to locate each group's comment, the proposed rule, and the final rule on the same policy spectrum.

Comments collected in the form of PDFs were converted into text files and loaded into Wordfish as datapoints in a CSV file. Before running the Wordfish model, the text was preprocessed to remove links, boilerplate headings, special characters, common stop words like "the", "and" etc., and words were lemmatized, or reduced to their simplest form (eg. "walking" becomes "walk"). This reduces the noise from repetitive words that are part of sentence writing but don't necessarily communicate anything about the political positions of the comments.

The wordfish model then scaled each document along the latent policy dimension and generated estimated positions with assigned theta values between -1 and 1 along the x-axis for each comment, "clustering" them into ideological groups.
A group was coded as a "winner" if the final rule moved closer to the group's position than the proposed rule. In cases where the final and proposed rule positions were the same to the naked eye, differences in theta values were used to understand the slight change in political positions of final rules.

While Wordfish helps identify alignment between preferences and outcomes, it does not explain why some groups win and others do not. To identify patterns in which factors or combinations of factors are associated with successful influence, I employ Qualitative Comparative Analysis (QCA). QCA is a case-oriented, configurational method used to identify combinations of causal conditions associated with a binary outcome. Rather than estimating average marginal effects like regression models, QCA examines which sets of conditions are necessary or sufficient for an

outcome, here, whether a group "won." This method is ideal for medium-N studies, those too small for regression models but large enough to explore cross-case variation systematically.

Using a crisp-set QCA (csQCA) approach, I binarized variables for salience, resource level, comment quality (high => 4), and group type (further explained below). I then created a truth table with each group-regulation pair as a case and its associated conditions. Table 2 in the appendix shows the input table for QCA using the QCA library in RStudio. The truth table and QCA analysis identifies consistent configurations associated with success, contradictory configurations where conditions do not clearly predict outcomes and coverage and consistency scores that help understand how many of the successful cases can be explained by the model's output.

QCA is especially well-suited to this study because my sample (approx. 30–40 cases) is too small for meaningful regression analysis but large enough for pattern identification. Moreover, QCA allows for equifinality, acknowledging that different groups may "win" through different combinations of strategies and contextual factors.

Operationalization of Key Concepts and QCA Variables

To translate theoretical variables into analyzable conditions, I defined the following:

Win (1= won, 0 = lost): A group "wins" if the final rule is closer to their stated position (estimated via Wordfish) than the proposed rule. This is a binary outcome used in the QCA.

High Salience (1= yes, 0 = no): A regulation is classified as high salience if it received over 200 public comments. This threshold distinguishes routine rulemaking from controversial, broadly contested issues where the stakes are higher and participation more intense.

High Comment Quality (1=yes, 0=no): Following the approach used in Yackee & Yackee (2006), I created a 6-point quality index based on the characteristics of each comment:

1. Identifies the commenter as an expert (e.g., lawyer, PhD, technical specialist)
2. Is longer than one page
3. Follows a detailed report format (e.g., includes a table of contents or executive summary, is structured with headings and subheadings like a report)
4. Cites multiple external sources (e.g., journal articles, government data)
5. Includes attachments (e.g., studies, legal briefs)
6. Suggests more than one change to the proposed rule

Each item is scored as 1 (yes), 0 (no), or ? (unclear). The sum yields a quality score ranging from 0 to 6. Documents with a quality score of 4 or higher are coded as high quality (1).

High Resource Level (1= yes, 0= no): Groups were coded as high-resource (1) if they met any of the following: (a) national-level umbrella organization, (b) union or association with over 100,000 members, or (c) commonly cited lobbying heavyweight. Others (e.g., local nonprofits, advocacy networks) were coded 0.

Labor Organisation (1= yes, 0 = no): Binary variable for labor (e.g., AFL-CIO, SEIU) vs. business (e.g., NAM, Chamber of Commerce) interests. If a group's (eg.non-profit or think tank) ideology aligned with labor interest, they were coded (1).

Data/Sources

My dataset consists of six labor-related regulations issued between 2010 and 2023 by three major U.S. federal agencies involved in labor policy: the Occupational Safety and Health Administration (OSHA), the National Labor Relations Board (NLRB), and the Equal Employment Opportunity Commission (EEOC). These regulations were selected based on their stage in the rulemaking process; they had completed the public notice-and-comment period and had final rules published, making them suitable for analyzing how stakeholder input may have influenced the outcome. To capture variation in issue salience and regulatory complexity, I intentionally included a mix of highly contested regulations, with thousands of public comments, and more routine rules with fewer than 50. All public comments were sourced from the regulations.gov docket system, and the final rule texts were retrieved either from the same platform or from the Federal Register.

The limitations of my data are that comments by interest groups stored in dockets don't provide a holistic picture of all the lobbying efforts from various groups. For the sake of simplicity, I will have to overlook lobbying efforts in the form of campaign support, interactions between lobbying groups and legislators, and other forms of lobbying; however, these are indeed important factors contributing to interest group influence and will be better captured in mixed-methodology studies as noted by Andreas Dur.

Below are the six regulations that this study explores:

1. Standard for Determining Joint-Employer Status

   This proposed rule by the NLRB revises the standard for determining joint-employer status under the National Labor Relations Act. It aims to clarify when two entities share

responsibility over workers by grounding the definition in common-law agency principles, particularly focusing on control over essential terms and conditions of employment.

Number of Comments: 12.94K

2. Regulations To Implement the Pregnant Workers Fairness Act

   This proposed rule by the EEOC implements the Pregnant Workers Fairness Act, requiring employers to provide reasonable accommodations for employees or applicants with limitations related to pregnancy, childbirth, or related medical conditions, unless doing so would cause undue hardship to the employer's business operations.

   Number of Comments: 98.56K

3. Representation-Case Procedures: Election Bars; Proof of Majority Support in Construction Industry Collective-Bargaining Relationships

   This proposed rule by the NLRB seeks to rescind and replace 2020 amendments related to representation-case procedures, including rules on election bars during unfair labor practice charges and voluntary union recognition. It also addresses rules specific to the construction industry. The goal is to better protect employees' rights to choose union representation and to support fair and stable collective bargaining processes.

   Number of Comments: 104

4. Walking-Working Surfaces and Personal Protective Equipment (Fall Protection Systems)

This OSHA proposed rule updates standards for walking-working surfaces and fall protection systems in general industry to better prevent workplace falls. It modernizes outdated regulations under 29 CFR Part 1910 and introduces flexible, hazard-specific protections such as personal fall arrest systems. The rule also seeks alignment with construction industry standards and incorporates public feedback through hearings.

Number of Comments: 108

5. Notification of Employee Rights under the National Labor Relations Act

   This proposed rule by the NLRB would require employers to post notices informing employees of their rights under the National Labor Relations Act (NLRA). The goal is to increase awareness of NLRA protections, promote employee engagement, and encourage compliance by both employers and unions. The rule also specifies the format, content, and enforcement mechanisms related to the notice requirement.

   Number of Comments: 6.12K

6. Clarification of Employer's Continuing Obligation to Make and Maintain an Accurate Record of Each Recordable Injury and Illness

   This OSHA proposed rule clarifies employers' continuing obligation to accurately record each recordable injury and illness under existing workplace safety standards. The rule emphasizes that employers must maintain complete injury and illness records for the entire retention period, even if the original record was incomplete or not created on time. It reinforces OSHA's enforcement authority and aims to improve workplace transparency and accountability.

Number of Comments: 22

Analysis and Findings

This section presents the empirical results of my analysis, beginning with the Wordfish estimates used to identify the policy positions of interest groups, proposed rules, and final rules for each of the six labor regulations studied.

The graphs below display the results of a Wordfish analysis estimating the policy positions of various actors involved in the rulemaking process for each of the six regulations.[1] The y-axis lists the names of organizations or entities that submitted comments, along with the proposed rule (PROP_RULE) and the final rule (FINAL_RULE) texts. The x-axis represents the estimated ideological position (Theta) of each text, where higher values (to the right) indicate a position more favorable to regulatory expansion (often aligned with labor interests), and higher values reflect more deregulatory or business-aligned positions. The blue dots represent labor organizations, and labor-leaning organizations, the red dots represent business organizations, and organizations that share business interests (advocates of deregulation, eg, Economic Policy Institute), and the green dots represent Proposed and Final rules.

[1] A list of acronym full forms can be found in the appendix

The theta score is a latent variable estimated by Wordfish that places each document on a single ideological dimension derived from word usage patterns across all texts. This dimension is not pre-specified but emerges from the data, making Wordfish particularly useful for unsupervised analysis of stakeholder positioning in contested policy debates. I have, however, set a direction such that comments aligning with that of AFL-CIO remain on the right side of the graph and those aligning with NFIB remain on the left. The horizontal line extending from each point is the confidence interval, which captures the uncertainty around the position estimate. Narrow intervals indicate high certainty in the position estimate, while wider intervals suggest more ambiguity, often due to shorter or less content-rich texts.

The distance between the proposed rule and the final rule shows the degree of policy shift that occurred during the rulemaking process. By comparing this movement to the positions of commenting organizations, we can infer which group(s) the final rule moved closer to.

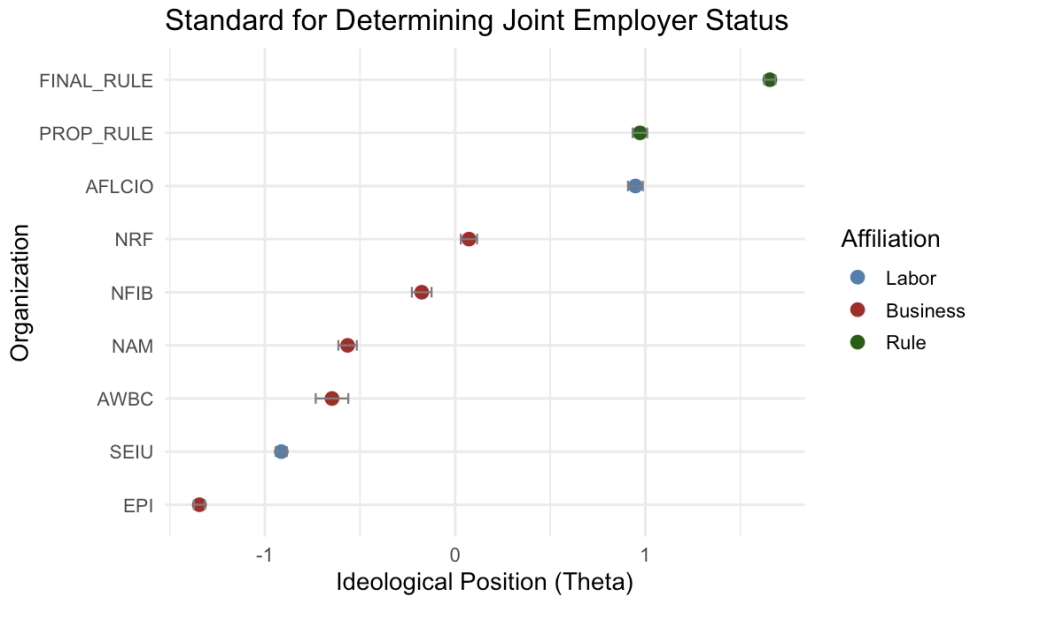

Figure 1. Standard for Determining Joint-Employer Status

In this graph (Fig. 1), we can see two clear clusters or sides forming, where the red dots (business organizations) are positioned on the left and the blue dot representing AFL-CIO is on the right. An interesting, unexpected result here is the position of SEIU, a major labor union, that is clustered with the red dots. This is odd because the comment of SEIU makes suggestions similar to AFL-CIO, but Wordfish aligns its ideological score with the business organizations. This is likely because of the frequency of words used in SEIU's comments, which may be similar to that of words used by organizations like NFIB and NAM in their comments. To understand this further, I produced a word-level visualization (Fig. 1.1), or a Wordfish feature scaling plot that helps analyze the "most influential" words that pull documents together.

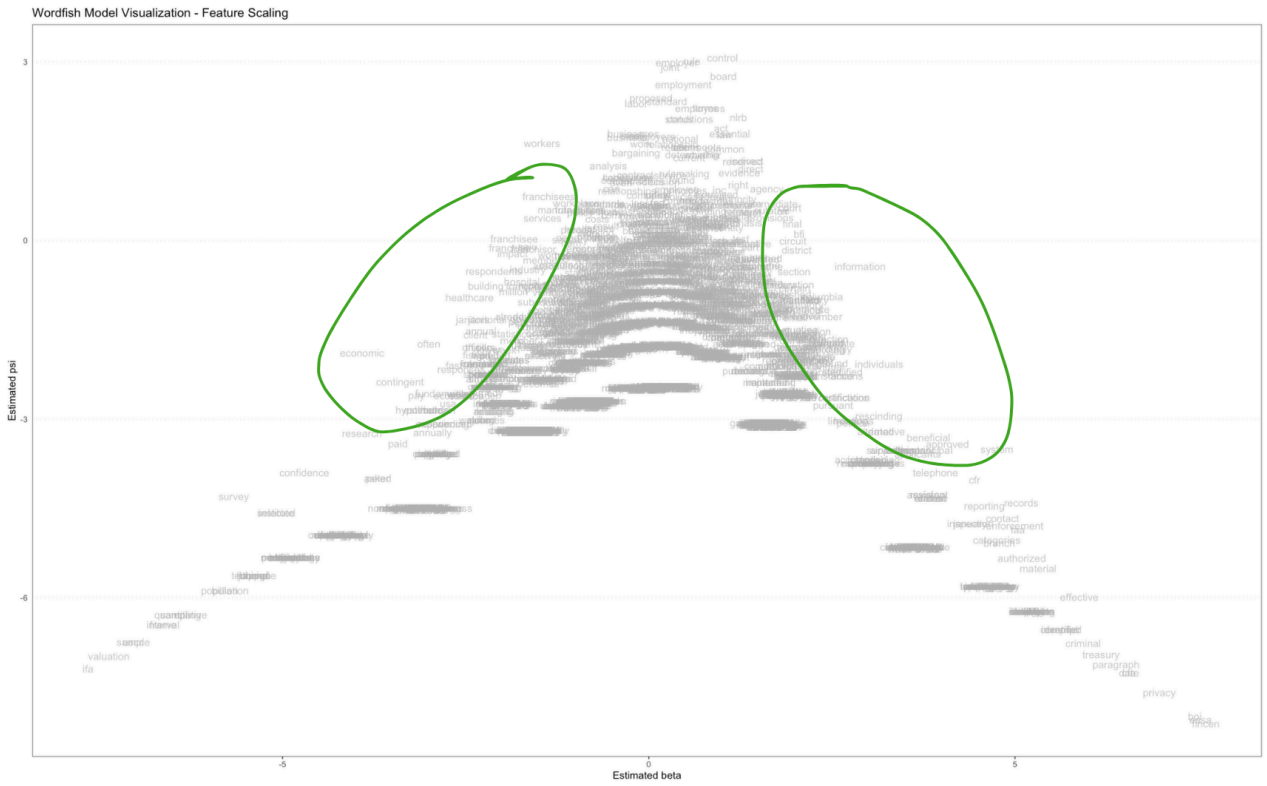

Figure 1.1 Word Level Visualization

All the words from the comments, proposed rule and final rule for this particular regulation are plotted along two axes. The horizontal axis (beta) represents how extreme the scaling of each feature (word) is: features far to the left and far to the right drive the placement of documents on the document-level left-right scale. Centrally located features have minimal impact. The vertical axis (psi) describes how common that feature is: more common features appear at the top of the model, while rarer ones appear at the bottom. The "most influential" words are usually the ones appearing on either side of the peak of the graph, circled in green. Here we see that words like "economic", "franchise", "healthcare", and "building" are common words used in business organizations' vocabulary, while labor groups use more words like "appeal", "information", and "appropriate". [2]

It is evident that wordfish is not an infallible method to identify political positions and ideologically cluster documents, but it is quite effective when working with large corpora of documents for empirical strategies.

Looking back at Fig. 1, the final rule appears to shift closer to the position of labor-affiliated organizations such as the AFL-CIO, suggesting that labor groups were more successful in shaping this particular regulation than business associations like the NFIB or NAM, whose positions were further away on the ideological spectrum.

[2] Image with 30 most influential words on each side is available in the appendix for a clearer look.

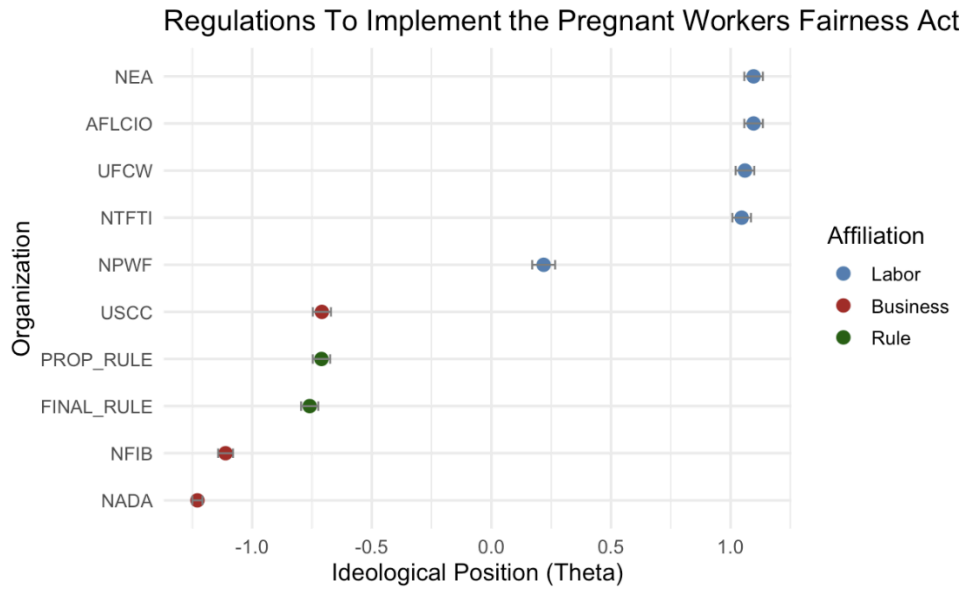

Figure 2. Regulations To Implement the Pregnant Workers Fairness Act

Similarly, from Fig. 2, we note the formation of two clear clusters with the blue dots, or labor organizations, to the right side and the red ones to the left. We note that the final rule moves slightly to the left, suggesting that the proposed rule was revised to include more words and ideas aligning with business organizations. For the sake of our study, even such slight movements will be considered a "win"; therefore, here it appears the business groups have won. Below is a zoomed-in feature scaling plot that shows the influential words in this set of documents:

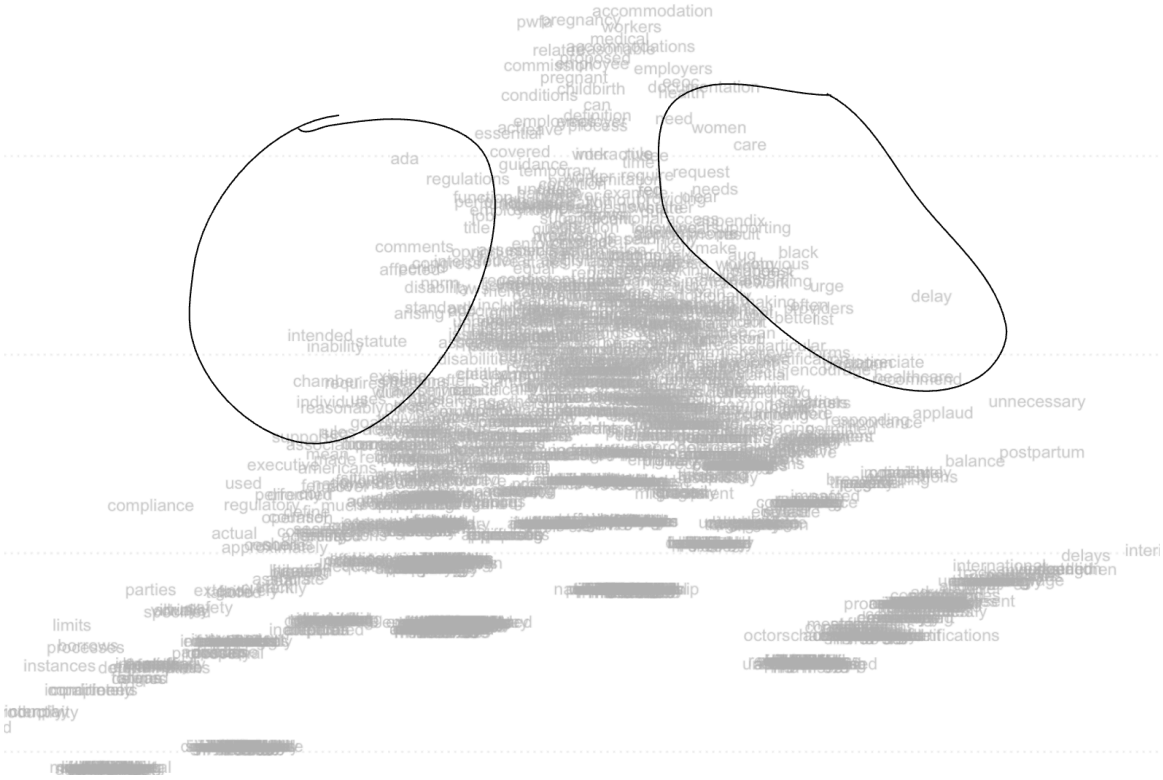

Figure 2.1 Wordfish Feature Scaling Plot for Pregnant Workers Act Documents

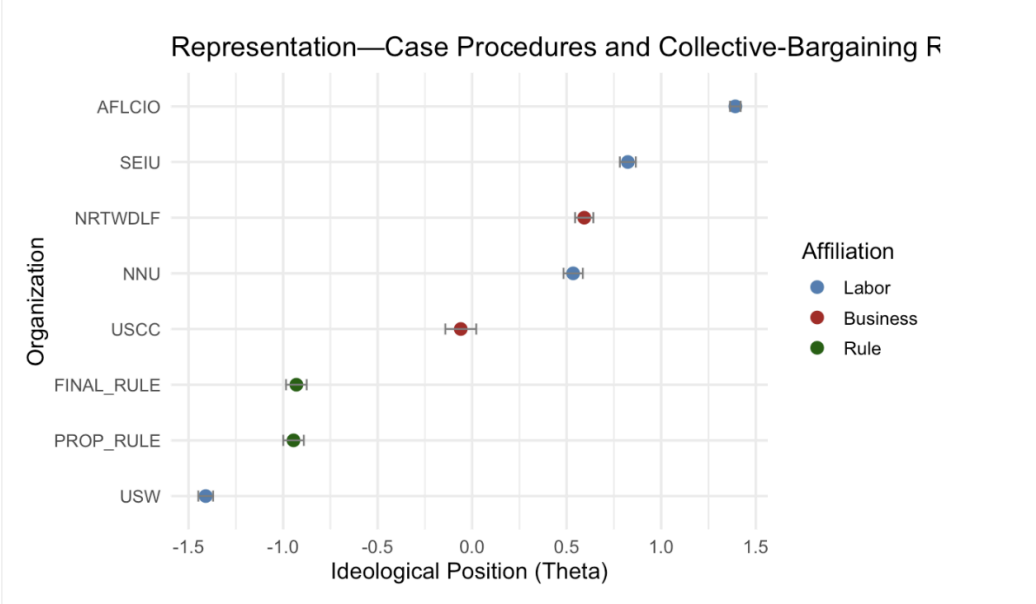

Figure 3. Representation-Case Procedures: Election Bars; Proof of Majority Support in Construction Industry Collective-Bargaining Relationships

In Fig. 3, we notice the formation of similar clusters, however, another anomaly is the National Right to Work Defence Legal Foundation (NRTWDLF), which strongly opposes rescinding the 2020 Election Protection Rule and takes a strong deregulatory and anti-union stance, but is positioned closer to the labor organizations in the graph above. This is again a result of the word frequency similarities between the documents. Wordfish's inability to consider the context in which words are used sometimes leads to odd clustering patterns such as this.

In this figure, the final rule moves very slightly rightward, with a theta value of -0.93, while the proposed rule was positioned at a theta value of -0.94, indicating a more labor-leaning final rule.

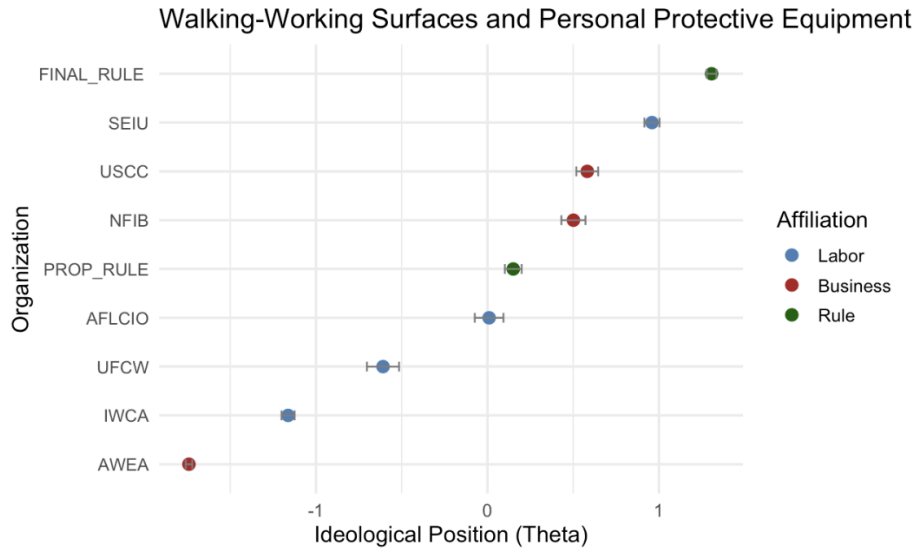

Figure 4. Walking-Working Surfaces and Personal Protective Equipment (Fall Protection Systems)

In Fig. 4, notably, the ideological positions of labor and business groups do not form distinct clusters, likely due to shared technical language in safety-focused comments and variation within

each group type. This case illustrates that issue-specific comments can blur traditional ideological boundaries, even within a business vs. labor framework, given shared use of technical terms. However, it is still evident that the final rule aligned most closely with labor organizations, particularly SEIU and AFL-CIO, and less with NFIB and NAM.

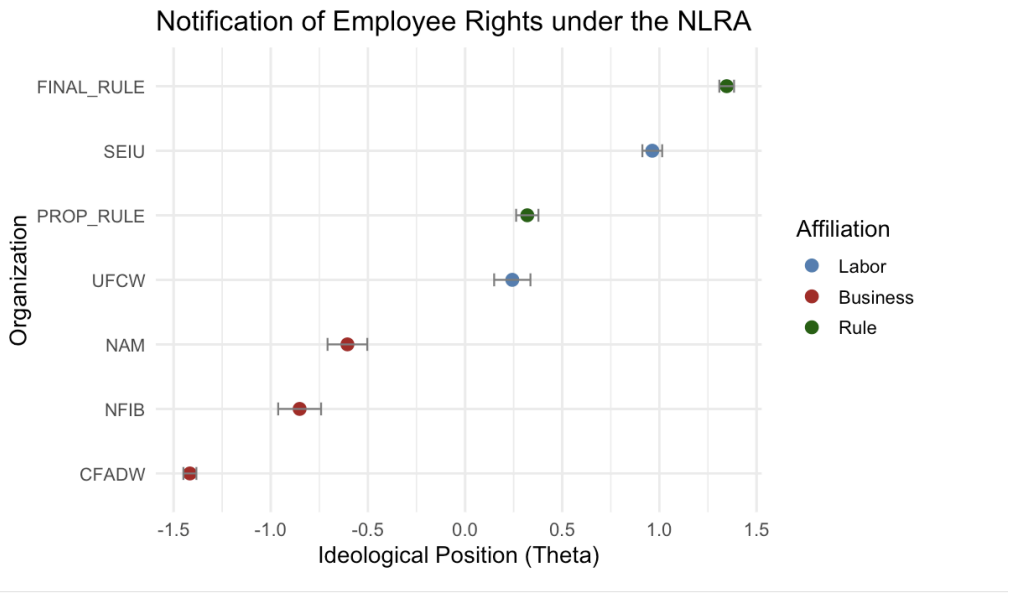

Figure 5. Notification of Employee Rights under the National Labor Relations Act

Here we can clearly make out that the final rule aligned more with the labor organisations like UFCW and SEIU.

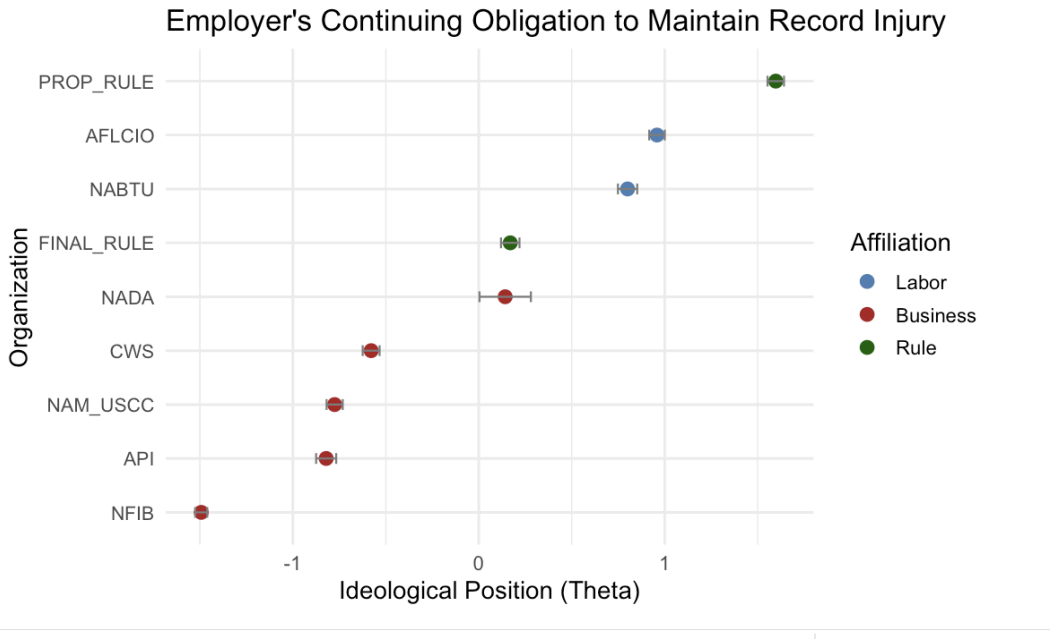

Figure 6. Clarification of Employer's Continuing Obligation to Make and Maintain an Accurate Record of Each Recordable Injury and Illness

We can see here in Fig. 6 that the proposed rule was initially well-aligned with labor organizations like AFL-CIO and North America's Building Trades Unions, however, the positions of the final rule shifted leftward towards the business organizations. We can infer that business organizations won in this case.

QCA results

The above graphs and wordfish were used to populate the "Outcome" field of the QCA input to indicate whether an organization was in their particular case or not. Along with that are 5 other independent variables that I take into consideration; these have been explained above in the methods section under "Operationalization of Key Concepts and QCA variables". Given our 5 considerations, there are a total of 32 configurations considered in the truth table below:

```
 OUT: output value
   n: number of cases in configuration
incl: sufficiency inclusion score
 PRI: proportional reduction in inconsistency
```

	Labor_Group	High_Resource	High_Quality_Comment	High_Salience_Rule	Citations	OUT	n	incl	PRI	cases
1	0	0	0	0	0	0	1	0.000	0.000	25
2	0	0	0	0	1	?	0	-	-	
3	0	0	0	1	0	0	1	0.000	0.000	6
4	0	0	0	1	1	0	1	0.000	0.000	7
5	0	0	1	0	0	?	0	-	-	
6	0	0	1	0	1	0	2	0.500	0.500	19,24
7	0	0	1	1	0	?	0	-	-	
8	0	0	1	1	1	?	0	-	-	
9	0	1	0	0	0	0	4	0.250	0.250	20,26,28,39
10	0	1	0	0	1	?	0	-	-	
11	0	1	0	1	0	0	4	0.000	0.000	3,4,30,31
12	0	1	0	1	1	1	1	1.000	1.000	15
13	0	1	1	0	0	?	0	-	-	
14	0	1	1	0	1	0	6	0.833	0.833	23,34,35,36,37,38
15	0	1	1	1	0	1	1	1.000	1.000	11
16	0	1	1	1	1	0	3	0.333	0.333	5,12,29
17	1	0	0	0	0	?	0	-	-	
18	1	0	0	0	1	1	1	1.000	1.000	18
19	1	0	0	1	0	?	0	-	-	
20	1	0	0	1	1	?	0	-	-	
21	1	0	1	0	0	?	0	-	-	
22	1	0	1	0	1	?	0	-	-	
23	1	0	1	1	0	0	1	0.000	0.000	13
24	1	0	1	1	1	0	1	0.000	0.000	14
25	1	1	0	0	0	1	3	1.000	1.000	21,22,27
26	1	1	0	0	1	?	0	-	-	
27	1	1	0	1	0	1	1	1.000	1.000	32
28	1	1	0	1	1	0	1	0.000	0.000	9
29	1	1	1	0	0	?	0	-	-	
30	1	1	1	0	1	0	4	0.500	0.500	16,17,40,41
31	1	1	1	1	0	?	0	-	-	
32	1	1	1	1	1	0	5	0.600	0.600	1,2,8,10,33

Figure 7. Truth Table

‘Incl’ can be considered the Consistency or Inclusion Score, which shows the proportion of cases in that configuration that actually exhibit the outcome (i.e., OUT = 1). This is the consistency threshold, most QCA models use incl ≥ 0.75 or 0.80 as a minimum threshold to consider the configuration a meaningful pathway to the outcome. E.g.. Row 15 has incl 0.833, which means 5 of the 6 cases with this configuration had out=1.

Similarly, PRI (Proportional Reduction in Inconsistency) measures how well the configuration avoids also leading to the absence of the outcome. Higher PRI (close to 1) indicates the configuration is not simultaneously a cause of failure (OUT = 0). Therefore, we focus on outcomes that have both a high consistency score and a high PRI.

Another thing to note is the number of remainders in our dataset, or the number of configurations with no cases (n = 0). Out of the 32 configurations, we have 14 remainders or rows without any cases, which may lead to limited diversity in our analysis. But since our dataset is quite small (~ 41 cases) and we are considering 5 conditions (32 configurations), our remainder levels can be overlooked for now, but a more robust analysis would require more cases and more diversity.

Next, we perform minimization, which identifies the simplest combination(s) of causal conditions (a "causal recipe") that are sufficient to produce the outcome, here, having the final rule align with the organization's suggestions.

```
M1: Labor_Group*High_Resource*~High_Quality_Comment*~Citations + ~Labor_Group*High_Resource*~High_Quality_Comment*High_Salience_Rule*Citations +
    Labor_Group*~High_Resource*~High_Quality_Comment*~High_Salience_Rule*Citations +
    ~Labor_Group*High_Resource*High_Quality_Comment*High_Salience_Rule*~Citations -> Outcome

                                                                                                  inclS   PRI   covS   covU   cases
-------------------------------------------------------------------------------------------------------------------------------------
1                     Labor_Group*High_Resource*~High_Quality_Comment*~Citations                  1.000  1.000  0.200  0.200  21,22,27; 32
2   ~Labor_Group*High_Resource*~High_Quality_Comment*High_Salience_Rule*Citations                 1.000  1.000  0.050  0.050  15
3  Labor_Group*~High_Resource*~High_Quality_Comment*~High_Salience_Rule*Citations                 1.000  1.000  0.050  0.050  18
4   ~Labor_Group*High_Resource*High_Quality_Comment*High_Salience_Rule*~Citations                 1.000  1.000  0.050  0.050  11
-------------------------------------------------------------------------------------------------------------------------------------
                                                                                             M1  1.000  1.000  0.350
```

Fig 8. QCA "Recipes"

The minimization algorithm has compressed our results into 4 winning configurations:

1. Labor_Group*High_Resource*~High_Quality_Comment*~Citations
    - A labor group that is high-resource, submitted a low-quality comment, and did not cite sources, was still successful.
    - Covers Cases: 21, 22, 27, 32
    - Suggests that in some situations, resources and identity can outweigh technical quality or evidence
2. ~Labor_Group*High_Resource*~High_Quality_Comment*High_Salience_Rule*Citations

- A business organization that is high resource, submitted a low-quality comment on a high salience rule, and includes citations or external sources in its comments was successful.
- Case: 15
- This suggests that high-resource business organizations may not always need to have good-quality comments to win, even when the topic is salient

3. Labor_Group*~High_Resource*~High_Quality_Comment*~High_Salience_Rule*Citations
    - A low-resource labor group with a low-quality comment, in a high-salience rule, that cites sources was successful.
    - Case: 18
    - This suggests that a low-resource labor group may win in situations where the rule is less salient and their comments are well informed with citations.
4. ~Labor_Group*High_Resource*High_Quality_Comment*High_Salience_Rule*~Citations
    - A non-labor group, high-resource, with a high-quality comment, in a high-salience rule, without citations, was still successful.
    - Case: 11
    - Suggests that strong comments and resources matter when the discussion is salient, even without formal citations

Overall model metrics like incl (1.00) and PRI (1.00) suggest perfect consistency, that is, all these pathways always lead to the winning outcome in our dataset, and that none of these pathways also lead to failure. This could be a result of our small sample size. The covS or Solution Coverage metric (0.35) suggests that the above solutions explain 35% of the observed successful cases.

This is fairly solid, especially in political and social data, where multiple unobserved factors often limit coverage.

Discussion

The results from the QCA provide a very balanced view of the influence of business and labor organizations with different resources and commenting strategies. Patterns like solution 1, with the highest number of supporting cases in our dataset (4), suggest that at times well well-resourced labor organisations may not need to submit high-quality comments in order to win. The supporting cases (21, 22, 27, 32) are from across the different regulations we looked at, suggesting that salience and the technical nature of the rule topic may not necessarily be a factor. It is possible that this also means that certain labor organizations, much like business organizations, have some level of structural power or credibility that, regardless of comment quality, they are heard. Another important thing to note is that these labor organizations spanned USW, AFL-CIO, and UFCW, organizations that usually provide high-quality comments across regulatory conversations, which might have contributed to building their reputation and structural power.

Solution 2 aligns more closely with the hypothesis of this study. This solution suggests that a high-resource business organization may not always need a high-quality comment, even when the topic is popular, to be able to win. While this supports the hypothesis of the study, that businesses are favored even in spaces meant to regulate them, there is only one supporting case for this in our dataset, case 15. To be able to make a more conclusive inference, we may need a larger dataset to observe the changes in the level of consistency in this configuration with more cases.

Solution 3 suggests that when a rule is low salience and doesn't attract much attention, labor organizations, even small ones, can win without high-quality comments. This is quite obvious as these agencies are responsible for protecting labor rights, and issues that are not that contested are likely regular, everyday rules that probably don't require much consultation anyway.

Solution 4 relates to my second hypothesis that groups with greater resources are likely to reach preference attainment. Business organizations with high resources and high-quality comments on salient rules seem to achieve their objectives. In my dataset, I could not find similar results for labor organizations.

Row 32 of the truth table looks at the configuration of a labor organization, which is high resource, provides a high-quality comment, with citations, on a high salience rule. Row 30 looks at the same configuration for a low salience rule, and in both case,s the only result available is a loss with a roughly 50-60% consistency rate. This means that 50-60% of the time in our dataset, a highly resourced labor organization submitted a high-quality comment on a rule (low or high salience) and lost. While these results do not have the consistency and PRI high enough to be minimized into a recipe, they provide empirical evidence that labor organizations, even ones with perceived power through resources and reputation, seem to not achieve their objectives even in spaces where agencies are meant to share similar ideologies with labor groups and work in favor of protective labor policies.

These QCA results provide satisfactory evidence and support for further exploration of business elite and labor organizations' influence within the sphere of labor policy in the US. There are certainly visible patterns suggesting that the salience of rules and the quality of comments play a role in influencing legislation through the notice and commenting process.

An important step in making this analysis more robust would be to develop a stronger preprocessing method for Wordfish that helps make documents more efficient and helps focus the attention only on words that likely express political ideologies to obtain more accurate theta values that help determine winners and losers.

Furthermore, this qualitative comparative analysis only provides a foundation for QCA within this topic. The combination of independent variables, while supported by previous literature, may need to be considered with other features to help draw a clearer understanding of patterns. This would involve more Theory-Driven Refinement of the QCA method, possibly to help answer questions such as:

The organizations that seem to win without high-quality comments in some (say low salience) cases, do they usually provide high-quality comments in other commenting spaces? This would be a metric of their perceived credibility and average comment quality

Further, are they part of a coalition? Maybe that's why they can compromise on their comment quality?

I am inclined to believe that the technical aspect of the topic being discussed is also an important consideration. For example, a more technical topic like regulation 4, Walking-Working Surfaces

and Personal Protective Equipment (Fall Protection Systems), would naturally require a more detailed technical comment and this would also make the wordfish analysis more confusing as all comments will likely use certain technical terms, the frequency of which may blur the theta scores and lead to inaccurate ideological positions. Thus, I would also want to add an independent variable that factors in the technical nature of a rule. A similar helping variable can be a rule topic variable that also helps narrow the rule area and provide a better understanding of whether influence remains the same if the topic of the rules differs. For example, do labor and business interest group influence patterns remain the same within different rule topics like worker safety or anti-discrimination?

Conclusion

The findings reveal several important dynamics. First, as hypothesized, business groups saw considerable favorable conditions within this regulatory space as well. However, the story is more nuanced: labor organizations were also successful under certain conditions, particularly when commenting on low-salience rules or when they submitted well-cited, high-quality feedback. Still, business groups enjoyed a structural advantage, evident in cases where they won despite submitting lower-quality comments or in highly contested rules. Second, resources consistently mattered. High-resource groups, whether labor or business, were more likely to win, supporting long standing theories about the importance of capacity and access in policy influence. Surprisingly, comment quality and citation inclusion, while important, did not guarantee success for labor organizations, suggesting that even credible participation may be insufficient without political alignment or institutional leverage.

This study also contributes to the development of methods to study influence and validate the effectiveness of softwares like wordfish. While not a foolproof method, wordfish proves to be an efficient way to understand ideological positions and is effective in cases like this where a QCA variable can be derived from the wordfish analysis.

The findings of this study highlight the unequal terrain of the notice and commenting process. While open to all, the capacity to generate influence remains unevenly distributed and agencies should consider mechanisms to amplify the voices of under-resourced stakeholders. Perhaps, introducing standardised commenting formats can enable stakeholders to compete more fairly with all of them having to abide by similar commenting guidelines, for example, including external sources to justify a point, conducting a cost-benefit analysis where feasible and affordable etc. Further agencies should set clearer guidelines on how comments are evaluated, or increase transparency around how comments affect final rule changes.

For interest groups, especially labor organizations, the results suggest that participation alone is not enough. Even well-resourced labor groups with high-quality, evidence-based submissions failed to sway policy in many high-salience rules. This underscores the need for coalitional strategies, long-term reputation building, and engagement outside the formal comment process, including political mobilization, media engagement, and backchannel communications. Groups may also benefit from coordinated commenting strategies, to replicate the volume and cohesion advantages often enjoyed by business associations.

Future research in this avenue should consider a larger dataset encompassing more rules and agencies would allow for stronger generalizations. While the six regulations in this thesis were strategically chosen for diversity, expanding the sample could improve the robustness of QCA patterns and the external validity of Wordfish estimates. Apart from expanding this research to different topics within labor regulations, the study can also span across other policy domains, such as climate regulation, education, or healthcare, to assess whether the patterns identified here hold across issue areas.

Moreover, the inclusion of additional variables, such as technical complexity of the rule, coalition membership, or media visibility, could refine the QCA models and account for unexplained variation. Some of the anomalies in my findings, such as labor losses despite high-quality engagement, may be better understood by incorporating these context-sensitive factors.

In conclusion, this thesis provides an empirical lens into the mechanics of influence in the regulatory state. It shows that business dominance is real but conditional, that labor influence persists but is uneven, and that institutional design matters in shaping who gets heard. While the notice-and-comment process holds potential as a democratizing tool, it also mirrors the broader imbalances in American political life. To ensure equitable representation in policy making, we must not only open doors to participation but also equalize the weight of the voices that walk through them.

References

W. D. Costain et al. "Interest Groups as Policy Aggregators in the Legislative Process." Polity, 14 (1981): 249 - 272. https://doi.org/10.2307/3234547.

Lucig H. Danielian et al. "The heavenly chorus: interest group voices on TV news." American Journal of Political Science, 38 (1994): 1056-1078. https://doi.org/10.2307/2111732.

M. Ambrosius et al. "State Legislators' Perceptions of Business and Labor Interests." Legislative Studies Quarterly, 13 (1988): 199. https://doi.org/10.2307/439821.

Andreas Dür et al. "The Question of Interest Group Influence." Journal of Public Policy, 27 (2007): 1-12. https://doi.org/10.1017/S0143814X07000591.

H. Klüver et al. "Measuring Interest Group Influence Using Quantitative Text Analysis." *European Union Politics*, 10 (2009): 535 - 549. https://doi.org/10.1177/1465116509346782.

J. Yackee et al. "A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy." *The Journal of Politics*, 68 (2006): 128 - 139. https://doi.org/10.1111/j.1468-2508.2006.00375.x.

Andreas Dür et al. "Interest Groups in the European Union: How Powerful Are They?." *West European Politics*, 31 (2008): 1212 - 1230. https://doi.org/10.1080/01402380802372662.

J. Potters et al. "Interest groups: A survey of empirical models that try to assess their influence." *European Journal of Political Economy*, 12 (1996): 403-442. https://doi.org/10.1016/S0176-2680(96)00008-0.

Appendix

Acronyms:
AFL-CIO: American Federation of Labor and Congress of Industrial Organizations
SEIU: Service Employees International Union
NRF: National Retail Federation
NFIB: National Federation of Independent Business
NAM: National Association of Manufacturers
AWBC: Association of Women's Business Centers
EPI: Economic Policy Institute
UFCW: United Food and Commercial Workers International Union
NEA: National Education Association
USCC: US Chamber of Commerce
NTFTI: National taskforce on trades women's issues
NPWF: National PArtnership for women and Families
NADA: National Automobile Dealers Association
NNU: National Nurses United
NRTWDLF: National Right to Work Defence Legal Foundation
USW: United Steel Workers
AWEA: American Wind Energy Association
IWCA: International Window Cleaning Association
ICC: Intercontinental Chemical Corporation
CFADW: Coalition for a Democratic Workplace
API: American Petroleum Institute
CWS: Coalition for Workplace Safety
NABTU: North America's Building Trades Unions

Table 1: 30 most influential words for Regulation 1

Description: df [60 × 4]

	Side <chr>	word <chr>	beta <dbl>	psi <dbl>
814	Left	economic	−3.913451	−1.88365891
1488	Left	often	−3.001928	−1.73309890
1957	Left	healthcare	−2.446066	−0.95777659
1990	Left	janitors	−2.404246	−1.32578714
2182	Left	client	−2.364531	−1.57818071
2058	Left	officers	−2.321485	−1.79155977
2249	Left	trujillo	−2.321485	−1.79155977
1425	Left	annual	−2.291934	−1.51439564
1959	Left	fight	−2.290700	−1.91878480
1976	Left	fissured	−2.290700	−1.91878480
1978	Left	weil	−2.290700	−1.91878480
1894	Left	building	−2.223898	−0.76234177
1053	Left	janitorial	−2.199857	−1.32952016
679	Left	lower	−2.180059	−1.82345755
847	Left	respondents	−2.112842	−0.51392192
1309	Left	posted	−2.100222	−1.93844655
1340	Left	week	−1.971403	−1.83251208
1307	Left	franchise	−1.894913	−0.12261110
1553	Left	statistics	−1.870643	−1.56978492
1182	Left	already	−1.864464	−1.27724495
815	Left	impact	−1.859791	−0.21752445
118	Left	million	−1.845765	−0.85713165
1474	Left	franchisee	−1.836120	0.03278221
1314	Left	llc	−1.769780	−1.67390130
1418	Left	member	−1.769780	−1.67390130
1986	Left	hospital	−1.731050	−0.68899277
1944	Left	david	−1.687425	−1.27535294
2055	Left	industry	−1.651227	−0.48645449
1947	Left	seiu	−1.650302	−0.10278240
806	Left	sector	−1.633210	−1.38954874
1005	Right	information	2.886351	−0.43434153
107	Right	appropriate	2.457930	−1.72544403
94	Right	revised	2.457630	−1.85866911
134	Right	issued	2.457529	−1.72503409
544	Right	appeals	2.434461	−1.70156925
1292	Right	gov	2.424262	−1.97891931
264	Right	accordingly	2.418527	−1.81896914
2598	Right	code	2.397523	−1.95195244
80	Right	subsection	2.389049	−1.65576165
2698	Right	columbia	2.369653	−0.94320717
579	Right	approval	2.367205	−1.76744329
1667	Right	november	2.327841	−1.27639395
2263	Right	manager	2.207214	−1.94769226
997	Right	filed	2.197403	−0.98282596
2727	Right	enfd	2.146913	−1.55415200
48	Right	mandatory	2.120827	−0.99078872
56	Right	subjects	2.117085	−1.16963075
96	Right	statutory	2.115358	−1.07271649
78	Right	adjudication	2.113254	−1.52274847
1828	Right	ownership	2.107977	−1.67200592
2750	Right	sub	2.103272	−1.84997060
2751	Right	nom	2.103272	−1.84997060
139	Right	subordinate	2.101301	−1.15500106
483	Right	cleaned	2.092849	−1.25255701
180	Right	inquiry	2.089876	−1.65528173

1–55 of 60 rows

Table 2: Input table for QCA

Description: df [41 × 8]

Case <int>	Organisation <chr>	Labor_Group <int>	High_Resource <int>	High_Quality_Comment <int>	Citations <int>	High_Salience_Rule <int>	Outcome <int>
1	AFLCIO	1	1	1	1	1	1
1	SEIU	1	1	1	1	1	1
1	NRF	0	1	0	0	1	0
1	NFIB	0	1	0	0	1	0
1	NAM	0	1	1	1	1	0
1	AWBC	0	0	0	0	1	0
1	EPI	0	0	0	1	1	0
2	AFLCIO	1	1	1	1	1	0
2	UFCW	1	1	0	1	1	0
2	NEA	1	1	1	1	1	0
2	NFIB	0	1	1	0	1	1
2	USCC	0	1	1	1	1	1
2	NTFTI	1	0	1	0	1	0
2	NPWF	1	0	1	1	1	0
2	NADA	0	1	0	1	1	1
3	AFLCIO	1	1	1	1	0	1
3	SEIU	1	1	1	1	0	1
3	NNU	1	0	0	1	0	1
3	NRTWDLF	0	0	1	1	0	0
3	USCC	0	1	0	0	0	0
3	USW	1	1	0	0	0	1
4	AFLCIO	1	1	0	0	0	1
4	AWEA	0	1	1	1	0	0
4	IWCA	0	0	1	1	0	1
4	ICC	0	0	0	0	0	0
4	NFIB	0	1	0	0	0	0
4	UFWC	1	1	0	0	0	1
4	USCC	0	1	0	0	0	0
5	CFADW	0	1	1	1	1	0
5	NFIB	0	1	0	0	1	0
5	NAM	0	1	0	0	1	0
5	UFCW	1	1	0	0	1	1
5	SEIU	1	1	1	1	1	1
6	NFIB	0	1	1	1	0	1
6	API	0	1	1	1	0	1
6	NAM	0	1	1	1	0	1
6	USCC	0	1	1	1	0	1
6	CWS	0	1	1	1	0	1
6	NADA	0	1	0	0	0	1
6	NABTU	1	1	1	1	0	0
6	AFLCIO	1	1	1	1	0	0

41 rows