Teaming Up Across Political Divides: Evidence from Climate Regulations

Dahyun Choi*

August 20, 2023

Why do interest groups with contrasting interests and policy goals work together? In this paper, I argue that interest groups prioritize high-quality implementation of policies even when it means compromising on their policy preferences. To test this argument, I introduce original measurement strategies that reveal systematic patterns in which firms and environmental groups invest in joint efforts to improve the implementation of greenhouse gas emissions standards. The analysis, using public comments spanning 2010-2020, demonstrates that comments written by joint efforts of environmental groups and firms contain more information that can contribute to the quality of policy implementation greater than individual efforts alone. Although the compromise is biased toward the firms' interests, environmental groups can exercise meaningful influence over the finalized policy outcome by inducing more participation from the firms. These findings highlight the hidden dynamics of regulatory politics, wherein divergent political goals are reconciled for high-quality policy implementation.

*Ph.D. Candidate, Department of Politics, Princeton University. Email: dahyunc@princeton.edu

1. Introduction

Why do political actors collaborate despite having competing interests and policy goals? The question of how political actors influence a policy has been a central topic in political economy. During the prior several decades, much progress has been made in understanding the process of developing policy when political actors with vested interests compete in varying institutional contexts (e.g., Hirsch and Shotts 2015, 2012; Baron and Ferejohn 1989; Krehbiel 2010; Crawford and Sobel 1982; Gilligan and Krehbiel 1989). One prominent argument is that actors use *policy-specific* expertise to effectively achieve a particular political goal. However, there are no clear explanations as to why and how political actors compromise their contrasting policy preferences "within a team," despite abundant empirical evidence pointing to the formation of "interest-diverse" coalitions (e.g., Nelson and Yackee 2012; Baumgartner et al. 2009; Dwidar 2022; Heaney and Leifeld 2018; Lorenz 2020; Phinney 2017).

This paper is motivated by several consistent empirical patterns in climate politics that classical accounts of policymaking literature do not explain. While the U.S. Chamber of Commerce opposed passing cap-and-trade legislation during the 111th Congress, several Chamber members joined the U.S. Climate Action Partnership (USCAP), a coalition of industry and environmental stakeholders that attempted to hammer out a workable compromise that could attract the necessary votes to become law (Livermore and Revesz 2015).[1] The Environmental Defense Fund (EDF), one of the mainstream nonprofit environmental advocacy groups, explicitly mentions on its website that it saw the need to partner with mainstream businesses since the 1980s. The group is actively partnering with

[1]See U.S. Climate Action Partnership, About U.S., http://www.us-cap.org/about-us/ (declaring USCAP's "pledge to work with the President, the Congress, and all other stakeholders to enact an environmentally effective, economically sustainable, and fair climate change program"); see also Eric Pooley, The Climate War: The Believers, Power Brokers, and the Fight to Save the Earth 142,170 (2010) (quoting Duke Energy executive Jim Rogers, a member of USCAP, responding to criticism of his participation by coal mining executive Robert Murray of Murray Energy: "Legislation is coming. We can help shape it, or we can sit on the sidelines and let others do it").

Walmart and FedEx.[2] Another example is the American Council for an Energy-Efficient Economy (ACEEE), one of the nonprofit coalitions supporting climate action. More explicitly, its Ally Program has listed utilities, manufacturers, and other energy industries as partners, such as the American Chemical Council, and Xcel Energy, in addition to a group of environmental and consumer leaders.[3]

These partnerships are puzzling given the contrasting policy preferences of firms and environmental groups. A closer analysis of business strategies in climate change reveals that restrictions on firms' polluting behaviors pose a significant challenge to particular industries. Although some firms (e.g., Shell, BP) have begun to diversify into other energy sources that produce less greenhouse gas emissions, none of these alternative energy sources can provide business opportunities on the same scale as those of oil and coal production (Stokes 2020). Contrary to industries' fear of adverse consequences from regulations, previous studies indicate that stringent regulations would primarily benefit environmental groups (Cheon and Urpelainen 2013; Bernauer and Caduff 2004; Keohane et al. 1998; Aidt 1998). However, despite the divergent effects of regulations leading to different policy preferences, firms and environmental groups collaborate closely.

I argue that the concern over the quality of policy implementation is the reason behind collaborative efforts between political actors who have contrasting policy preferences. In regulatory politics, the pursuit of *policy preference* is accompanied by concerns for *quality of policy implementation*. This holds true within the context of climate regulations, where the primary objective is achieving target emission reductions. Here, the instrumental motive of policy outcomes themselves becomes less significant (Hirsch 2022; McCarty 2020), as long as political actors are able to contribute to reducing uncertainties in policy implementation. Although the divergent policy goals are reconciled, political actors prefer

[2]See the website of Environmental Defense Fund, https://www.edf.org/partnerships/business-and-industry. EDF has collaborated with over 30% of Fortune 100 companies

[3]Please see the website of ACEEE for further details. https://www.aceee.org/aceee-ally-program. Allies receive benefits from ACEEE, including public recognition via ACEEE's website, early access to ACEEE research reports, and access to a network of energy efficiency experts, leaders, and decision-makers.

a compromise with a high-quality of policy implementation instead of their own preferred policies with a low-quality implementation.[4]

To analyze why political actors work together despite unaligned preferences, I draw upon the theoretical framework of McCarty (2020) and Alchian and Demsetz (1972) to incorporate the dynamics of rulemaking for which regulatory officials need quality information to make reasonably good policy decisions. Most regulations are created by bureaucrats (Warren 2018; Shipan 2004), a process that is particularly true for environmental policymaking, where very few environmental laws have been passed (e.g., Rothenberg 2018; Lazarus 2014). Focusing on climate regulations, I show that firms and environmental groups, which have competing interests, invest in joint efforts to provide informative texts (defined as abundant analytical evidence and scientific reasoning) so that regulators can make fine-grained and technical judgments (Breyer 1982; Hawkins and Thomas 1989).

My theory provides micro-foundations for the argument that interest group competition in regulatory policymaking is centered on the provision of expertise (Epstein et al. 2014; Carpenter and Moss 2014; Huber and Shipan 2002; Weingast 1984). Existing research on interest group politics is focused on financial resources, such as PAC contributions or lobbying expenditures, as a measure of political power. However, the primary resource of power in the regulatory context is information.[5] In this paper, I describe systematic measurements of information, placing particular emphasis on expertise. This approach contrasts with existing literature on rulemaking, which has primarily focused on analyzing the frequency of submissions or the types of political actors involved in the notice-and-comment period. I accomplish this by conducting an analysis of 15,883 publicly submitted comments on greenhouse gas emissions standards between 2010 and 2020. I first filter organization/entity comments for comparability and classify comments by five types to

[4]Please see Choi (2023) for equilibrium characterization of the game where agents with contrasting preferences work together.

[5]The role of information in the regulatory process has been discussed in a wide range of literature. Magat et al. (2013) elaborates that higher quality information supporting a proposed regulation reduces opponents' ability to modify the regulations. Moreover, the timing of when information is received can influence the rulemaking decisions (Ingram and Ullery 1977).

capture who participates in rulemaking. To this end, I retrieve the history of environmental groups' websites using Wayback Machine, and reference IRS Form 990 tax returns from the charitable foundations funded by Fortune 500 and S&P 500 corporations (Bertrand et al. 2020), and Cory et al. (2021)'s classification framework. I provide descriptive patterns of comments on emission standards that joint coalitions have continuously submitted.

I then dive into the political implication of joint efforts on climate regulations by firms and environmental groups. In the regulatory process, the information that regulators need is sometimes held only by the business interests they seek to regulate. For instance, polluting firms are better positioned to know details concerning the environmental risks created by their production processes (Coglianese and Lazer 2003; Wagner 2003). Hence, their inherent information advantage over the government and other political actors results in compromised policy outcomes that are relatively favorable to the firms. By incorporating text embedding methods with a *Paragraph Vector* framework, I show that comments from environmental groups with business partners are relatively skewed to business-friendly topics compared with comments from environmental groups that lack business partnerships. However, given that they represent a compromised outcome, the extent of the issue slant in comments from partnerships is comparatively less pronounced than the slant in comments from business interests alone.

I further explore the benefits that environmental groups and business interests gain from strategic partnerships by quantifying information using named entity recognition techniques. Public comments written by strategic partnerships of firms and environmental groups contain more specific evidence and analytical reasoning compared with comments composed individually by each group, which is consistent with my theoretical predictions. Specifically, a collaboration with business partners substantially augmented the volume of information present in the comments associated with environmental groups, even after controlling for different group characteristics. Lastly, I employ information theory to quantify the political influence of strategic partnerships on finalized policy outcomes. I find that comments produced through collaborative efforts between firms and environmental groups exhibit a closer statistical distance to the finalized policy relative to comments composed by single entities. As a robustness check, I examine the citation patterns among EPA officials. The results reveal that EPA officials tend to cite comments written by strategic partnerships more frequently than they cite other types of comments. These findings provide further support for my argument on why political actors with conflicting interests engage in collaboration, and how the enhanced quality of information that results from strategic partnerships is translated into political influence in regulatory politics.

This article makes both theoretical and empirical contributions to the study of coalition lobbying in policymaking (e.g., Bertrand et al. 2020; Dwidar 2022; Junk 2019; Phinney 2017; Heaney and Lorenz 2013; Nelson and Yackee 2012; Hula 1999), with implications for understanding regulatory politics in which interest groups with conflicting interests prioritize a high quality of policy implementation. By examining the coalition of polluting firms and environmental groups and their effects on climate regulations, I also contribute to the empirical literature on the influence of interest groups on climate politics (e.g., Cory et al. 2021; Colgan et al. 2021; Culhane et al. 2021; Brulle and Downie 2022; Lerner and Osgood 2022; Sautner et al. 2020; Urpelainen and Van de Graaf 2018).

The rest of the paper is organized as follows. The next section discusses the broader literature on regulation politics and outlines theoretical expectations regarding the strategic partnership of business interests and environmental groups. I then describe my dataset and empirical strategies and provide empirical evidence for my arguments. The final section discusses the implications of strategic partnership in environmental politics, as well as the contribution to broad literature on interest group politics.

2. Interest Groups Working Together in Regulatory Politics

Scholars have emphasized the influence that interest groups have over regulatory policymaking. Regulators have significant discretion in formulating regulations (McCarty

2017), and interest groups consider various actions to influence regulators' policy choices that are in their favor. Interest groups directly lobby bureaucrats (You 2017), serve on federal advisory committees (Balla and Wright 2001; Moffitt 2014), lobby legislators who wield oversight authority over bureaucrats (Hall and Miler 2008; Epstein and O'halloran 1995; McCubbins and Schwartz 1984), and participate in the notice and comment process (Gordon and Rashin 2021; Libgober et al. 2020; Yackee and Yackee 2006; Haeder and Yackee 2015; McKay and Yackee 2007; Furlong and Kerwin 2005).

Interest groups frequently engage in these political activities via formal partnerships or ad-hoc coalitions (Nelson and Yackee 2012; Baumgartner et al. 2009; Hula 1999; Heinz et al. 1993). They invest as teams in any coordinated efforts, with the objective of advancing their interests. To explain why lobbying together is a more advantageous strategy compared to lobbying alone, scholars have analyzed the size of coalitions (Nelson and Yackee 2012) or the types of their interests (e.g., broad versus narrow) represented in the coalition (Mahoney 2007). A recent growing body of work relates lobbying success to the effect of the composition of coalition such as organization types (e.g., trade association and sectoral firms), partisan identities, or interest diversity (e.g., organizations representing diverse industries) (Dwidar 2022; Heaney and Leifeld 2018; Lorenz 2020; Phinney 2017). However, the dynamics of how competing interests compromise a policy "within a side" and what incentivizes them to work together despite such compromises are rarely addressed. To bridge the gap, I propose a theoretical prediction wherein compromises between political actors emerge endogenously due to interest groups' concern for high-quality policy implementation.

2.1. Theory: Investing in Team Efforts for Improving the Quality of Policy Implementation

My focus is participation in the notice and comment process because this stage is the most common way for interest groups to get their voices heard regarding agency policies

(Baumgartner et al. 2009; Yackee and Yackee 2006; Baumgartner and Jones 2010).[6] The process of rulemaking is centered on improving the implementation of policy after an agenda is fixed (You 2017). Thus, it requires fine-grained, technical judgment concerning how major operations should be designed. Therefore, information, namely expertise, plays a vital role in regulatory politics (Libgober et al. 2020; Breyer 1982; Hawkins and Thomas 1989), and political actors with specialized knowledge of the complex policy arena have an advantage in this competition (Epstein et al. 2014).

Given the nature of regulatory policymaking, I argue that the demand for high-quality policy implementation is dictated through cooperative specialization despite differences in policy preferences. The competing political goals are reconciled to the extent that political agents are incentivized to contribute to joint products to improve the quality of policy implementation. On the *policy preference side*, political actors have asymmetric capacities in the sense that their different areas of expertise have varying impacts on regulators (Berry and Wilcox 2015; Yackee and Yackee 2006). These differing abilities at developing policy proposals influence the way political actors compromise within a team. The high-capacity group tends to be more engaging due to its superior resources (e.g., the impact of information, staff expertise, and funding) compared with the low-capacity group, and as a result, the imbalanced capabilities lead to a compromised policy outcome biased toward the high-capacity group. However, although the compromised outcome relatively favors a high-capacity group, a low-capacity group gains advantages by inducing more participation from the high-capacity group than otherwise, assuming that preferences over *policy outcome* and *quality of policy implementation* are inseparable, and assuming it elicits concessions from high-capacity groups' extreme policy preferences.[7]

[6] Under the Administrative Procedure Act (APA) of 1946, agencies typically must provide the notice and comment period in which a proposed policy is open for public review. During this stage, all interested parties are invited to provide written comments regarding the content of the proposed rule posted by agencies.

[7] Please see McCarty (2020) and Choi (2023) for further details concerning the joint policy production where agents with asymmetric capabilities work together. Agents are willing to invest more effort when a compromise is closer to their ideal policy. When inputs from agents are substitutes, a reduction of the inputs of one agent increases the marginal productivity of another. When inputs are complements, increased

On the *quality of policy implementation* side, a jointly generated product exceeds the sum of its individual contributions. Both agents can benefit from exchange and production in accord with the comparative advantage and save resources for gathering or processing information for crafting a proposal (Alchian and Demsetz 1972). Because the rulemaking process is focused on improving the implementation of policy, groups are motivated to invest in a joint effort to achieve a high quality of implementation. A group prefers high-quality policy implementation with a compromised outcome over an ideal policy that has a low-quality implementation. Therefore, despite the differences in policy preferences, both groups benefit from investing in joint efforts.

2.2. Empirical Evidence: Strategic Partnerships between Firms and Environmental Groups in Climate Regulations

I examine the strategic partnerships between polluting firms and environmental groups for empirical implications as the dynamics of environmental regulatory policymaking squarely represent the properties of the theory outlined in the previous section.

First, polluting firms and environmental groups have asymmetric capacities because the regulatory system depends heavily on information supplied by the regulated entities. The information regulators need is often held only by the industries or firms that they should regulate (e.g., McCarty 2017; Wagner 2003; Coglianese and Lazer 2003). Regulated entities possess private information about costs, compliance, or the industry-level effects that would be useful for policymakers to know (Gailmard and Patty 2012). Particularly, in the realm of climate policymaking, regulators are poorly positioned to gather information about business operations, and the best source of information about mitigation costs or the feasibility of different reduction approaches is the very firms that regulators seek to regulate. Although epidemiological research by government or environmental scientists can reveal as much about the health effects of pollutants, firms typically know more

inputs by one agent increase the productivity of another. The analysis of this paper assumes that inputs from political actors are complements.

about what they produce, as well as how they produce it (Coglianese 2007; Coglianese and Lazer 2003).[8] Therefore, business interests have an information advantage about which pollution control measures will be effective in their facilities and which measures would yield unexpected costs or consequences.

Such information advantage of polluting firms in environmental regulations implies asymmetric capabilities of business interests and environmental groups. And correspondingly, the imbalances in capacities within a team present an intuitive pattern, where the compromised policy outcome is biased towards the preferences of the higher capacity group. Although it is impossible to exactly divide each group's individual contributions to a jointly designed policy outcome (Alchian and Demsetz 1972), we can empirically demonstrate whether the compromised policy outcome is biased in favor of a high-capacity group. If the compromised policy outcome is relatively skewed towards the topic favored by a high-capacity group compared with the topic emphasized by a low-capacity group, we can infer that the compromised outcome favors the high-capacity group.

For my analysis, I leverage the fact that polluting firms have strategically highlighted R&D and technological issues in the climate debates. Abundant qualitative evidence suggests that business actors attempt to reframe climate policy and weaken EPA's justification for emission cuts by strategically discussing R&D and technological issues (Grumbach 2015; Downie 2017).[9] To give an example, ExxonMobil highlights its contributions to climate actions with advertorials citing "our industry-leading investments in research and development," such as the Global Climate and Energy Project at Stanford University, which implies that current solar or wind technologies are inadequate (Supran and Oreskes 2021). According to related witnesses and testimonies, business interests strategically use sci-

[8] Coglianese and Lazer (2003) suggests that the EPA could not have regulated 160 industries without business actors involved in constructing regulatory standards.

[9] Still, large firms have not provided emissions reduction targets despite saying they want to reduce their impact on climate change. They have made R&D and technology commitments but have struggled to cut emissions. Eavis, P., & Krauss, C. (2021, May 12). What's Really Behind Corporate Promises on Climate Change? The New York Times. https://www.nytimes.com/2021/02/22/business/energy-environment/corporations-climate-change.html

entific research and technology to undermine efforts aimed at reducing emissions or to emphasize the uncertain costs associated with climate policies (Schlichting 2013). This use of science to weaken antipollution efforts leads to my first hypothesis, namely, that comments written by strategic partnerships would emphasize R&D and technological issues more than comments written by environmental groups alone. However, the extent of the slant toward R&D topics in the comments would be less pronounced than what is observed in comments authored solely by business interests because comments written as an outcome of strategic partnerships are a compromise between the two.

HYPOTHESIS 1. *(Compromised Policy Outcome) Comments from strategic partnerships between firms and environmental groups would be slanted toward discussing R&D and technology topics, compared with comments written by environmental groups lacking business partners.*

Second, although the compromised outcome is biased toward the firms' preferred policies, environmental groups can derive benefits by collaborating with these firms. Benefits accrue because firms are motivated to help generate a higher output when the compromise aligns more with their preferences.[10] Given that firms' contributions as high-capacity groups can have a greater positive effect on the quality of policy compared with the contributions of environmental groups, environmental groups are willing to make concessions to achieve a higher-quality policy implementation.

Investing in joint efforts is more efficient than devoting separate, additive efforts in multiple ways, not only for environmental groups but also for firms (Alchian and Demsetz 1972). For instance, firms can better frame their private information in conjunction with environmental groups' expertise in climate mitigation strategies, community-level knowledge (Bolden et al. 2018), or scientific research presented by environmental groups that concern the likely impact of further pollution (Bromley-Trujillo et al. 2014). And environmental groups can access private information that firms hold concerning the types of pollutants firms produce or the processes of generating those pollutants. On the basis of

[10]This is consistent with theoretical predictions in McCarty (2020) and Choi (2023)

this inference, I posit that comments formulated by collaborative efforts between firms and environmental groups contain the comprehensive scientific reasoning and specific information sought by regulators to develop and implement technical aspects of a policy, as compared to other forms of comments written separately by each group. The nature of collaborative comments leads to my second hypothesis:

HYPOTHESIS 2. *(Augmented Expertise): Comments crafted through collaborative efforts between firms and environmental groups contain a greater amount of scientific evidence and specific information compared with comments written separately by either environmental groups or business interests.*

Lastly, regulators who implement environmental regulations require an understanding of various solutions to reducing pollutants and greenhouse gas or the unexpected consequences of alternative regulatory standards (Coglianese 2007). Therefore, expertise is a key factor in policy implementation and regulators value the specialized knowledge that reveals the intricacies of the policy landscape. Given that comments arising from the joint efforts of firms and environmental groups are more informative than other types of comments, I hypothesize that the comments produced by the collaboration of firms and environmental groups will have a greater impact on the finalization of the policy outcome compared with comments written independently by either business interests or environmental groups. Hence, my final hypothesis is the following:

HYPOTHESIS 3. *(Political Influence): Comments from joint efforts are more likely to influence policy amendments than other types of comments, among comparably resourced comments.*

Another potential explanation is the availability of resources. Interest groups possess diverse resources and capacities (Yackee and Yackee 2006; Berry and Wilcox 2015), as previously mentioned. Therefore, the establishment of strategic partnerships or the production of high-quality comments may depend on these factors. To address this concern, I construct a variable to control for group characteristics, such as staff size. Data for this variable are collected from various sources, including *InfluenceWatch*, which provides descriptions of political actors involved in public policy issues, and from firms' websites, LinkedIn, Indeed, Buzzfile, Rocketreach, or Glassdoor.[11]

3. Data and Stylized Facts

I use an original dataset containing 15,883 comments officially submitted on *Greenhouse Gas Emissions Standards* from 2011 to 2020; the dataset does not have duplicates.[12] The policy comments were submitted for the EPA's regulatory review of the Greenhouse Gas Emissions Standards under sections 111 and 112 of the Clean Air Act, for which the EPA opened notice-and-comment periods seven times.[13] The year 2011 was chosen as a starting point because it immediately follows the new rules in which the EPA expanded emission regulations to a wide range of industries. The 10-year time period ensures that I am able to observe how both Republican and Democratic administrations respond to policy comments. As noted on the website of the Environmental Defense Fund,[14] the history of strategic partnerships with business interests traces back to the 1980s. Between 2011 and 2020, these partnerships have consistently remained unchanged in terms of temporal variation between firms and environmental groups. Comments from individuals without an organizational affiliation tend to be simple endorsements focused on support for or opposition to a proposed policy. To compare comments that provide substantive information, comments from individuals who lack any association with entities or organizations are dropped from the main analysis. Ultimately, using company/organization identifiers and automated text analysis, I filter 903 comments submitted by companies, entities, or organizations and I use these filtered

[11] When employment size is indicated in ranges, the upper bound is coded as the staffing size of the group.

[12] Regulations.gov includes data including the proposed policy, finalized amendments, and the comments associated with them. All rules and associated comments are linked by a docket number. A docket number is a unique identifier created by agencies that follow a regulation throughout its rulemaking process.

[13] Following is the list of starting dates the EPA posted for each notice and comment period: 1) November 30, 2011, 2) May 13, 2013 , 3) July 22, 2014, 4) January 5, 2015, 5) November 17, 2015, 6) April 1, 2016, and 7) August 9,2017.

[14] Please see Figure A.3

comments as the basis of my analysis.[15]

Comments are classified by five categories: 1) environmental groups with business partnerships, 2) environmental groups without business partnerships, 3) business associations (e.g., trade associations), 4) single businesses, and 5) others such as universities or government agencies.[16] One interesting pattern to note about this collection of comments is that recognizable polluting firms (e.g., Exxon, BP, Ford, or General Motors) have submitted relatively few comments by themselves. Most of the single firms that participated in the rulemaking process by themselves are "green firms" or small local businesses. The classification is operated by two measurement strategies. First, I provide the conservative measure of strategic partnerships between firms and environmental groups based on explicitly visible evidence. I retrieve the history of environmental groups' websites for the prior decade using the Wayback Machine, and code if environmental groups have explicitly posted polluting firms as partners. [17] Next, I construct a more generous measure by incorporating relatively invisible flows such as corporate donations into the explicitly visible channels, relying on IRS Form 990 series (Bertrand et al. 2020). Additionally, I reference the classification framework of Cory et al. (2021) to double-check the validity of the memberships lists that I collected from other sources.[18] The main analysis presented in this paper is based on the most conservative measure of partnerships between polluting firms and environmental groups constructed from explicit evidence- environmental

[15]There is no systematic correlation between the number of comments by type and participation year.

[16]I used three criteria to identify environmental groups. First, these groups are required to have a mission primarily relating to climate change and public policy. Second, the groups should be membership-based organizations. Finally, the group's membership should include diverse categories of political actors, such as citizens, consumers, and environmentalists. For instance, even though it is introduced as a pro-climate coalition in the press, the group is categorized as a business association if the membership was limited to firms.

[17]The measurement strategy focuses solely on partnerships between environmental groups and firms operating within polluting industries such as energy, transportation, oil, or coal. It does not take into account partnerships between environmental groups and green firms within renewable energy or green technology industries. Although there are a few instances of environmental groups collaborating with green firms, partnerships with polluting firms are more widespread.

[18]Unfortunately, Cory et al. (2021) classification covers approximately one hundred firm-centered climate coalitions. So it was not enough to fully validate the strategic partnerships of firms and environmental groups examined in this analysis.

groups' websites. In total, I have 541 unique entities in my data. The summary statistics are provided in Appendix.[19]

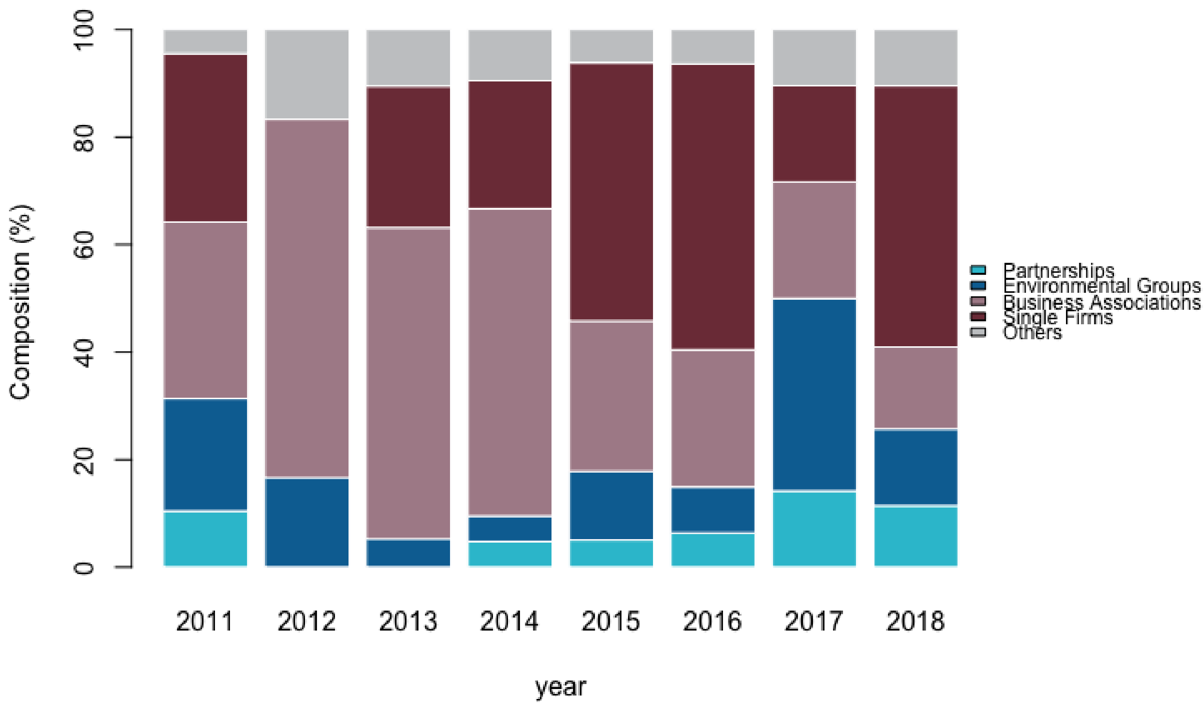


FIGURE 1. Comment Participation With Time

No comments were submitted by organizations in 2019 and 2020. EPA did not open the notice-and-comment period in 2012 and 2023, but comments were still submitted.

Figure 1 presents the composition of the comments across time. On the whole, policy comments by business associations and single firms represent the plurality of comments most of the time.[20] With time, there has been a gradual decrease in the percentage of comments from business associations and a stable trend in the percentage of comments from the partnerships between firms and environmental groups. The increase in the percentage of comments from single firms might mean an increase in participation from "green" firms. Although the frequency of joint coalitions' participation in rulemaking seems

[19] Please see Table A.3 in Appendix

[20] This observation is consistent with Golden (1998)'s finding that a huge percentage of comments are from business interests.

to be smaller than that based on other types of comments, the information conveyed by joint coalitions to regulators tends to be richer than the information of other types of comments. The next section substantiates this statement empirically.

4. Empirical Evidence

In this section, I provide empirical evidence that supports the theoretical argument. To do so, I use a text embedding method and computational techniques.

4.1. Compromised Policy Outcome Biased Towards a High-capacity Group

To investigate whether firms' information advantage leads to a compromised, business-preferred policy outcome, I construct two measures to capture the prevalence of the topics favorable to business interests: machine learning-based metrics of 1) R&D and technology coverage and 2)socioeconomic consequence coverage. There is an abundance of qualitative evidence that business interests strategically frame their climate communication by highlighting either scientific uncertainty or their contributions to R&D and technology (e.g., Supran and Oreskes 2021; Downie 2017; Grumbach 2015; Schlichting 2013). If the comments produced through collaborative efforts between firms and environmental groups primarily focus on or exhibit a bias towards business-friendly measures instead of emission reduction, we can deduce that the resulting compromised policy outcome is skewed in favor of a high-capacity group.

Measuring issue slant towards R&D and Technology

Count-based metrics convey little information concerning the context in which words are used. To handle this limitation, I apply a text embedding method that allows words to encode meaningful information about analogies. Political science research has used *Word2Vec* which embeds words in a low-dimensional vector space using neural network structure (e.g., Rodriguez and Spirling 2022). This method results in a set of vectors whereby proximity in vector spaces implies similar meaning context-wise, and vectors distant from each other have different meanings. For instance, “diligent" and “industrious" would be close together, whereas “diligent" and “lazy" would be relatively distant from each other. On the basis of embedding methods, I allow the algorithm to assign each word to a vector in a shared space during the training stage, and these assignments create clusters of words that are semantically connected. As a result, the more similar the context, the closer two words are located in geometric space.

Built on this advance in modern natural language processing technique, I use *Paragraph Vector* proposed by Le and Mikolov (2014), an unsupervised framework that learns continuously distributed vector representations at the comment level. In the *Paragraph Vector* framework, each document is mapped to a unique vector while each token is also mapped to another unique vector. They are then averaged to predict the next words in each sentence. Similar to Word2Vec's continuous-bag-of-words model, this approach is based on a distributed memory model whereby document vectors can be acquired by the task of predicting a word based on an average in consideration of context and full document levels. I construct a model with a window size of five, and I do not consider words that are observed less than five times in the entire corpus.[21]

As explained earlier, a key feature of word embeddings is that the difference between word vectors in the geometric space conveys meaning. For instance, the difference between the two vectors, $\overrightarrow{R\&D} - \overrightarrow{Reductions}$, identifies an issue dimension in the space by taking the difference between the normalized vector across a set of research words and the average normalized vector across a set of emission words: [22]

[21]The analysis reported in this paper was implemented by Doc2Vec Gensim and python3 on December 29, 2022. The parameters epoch is specified as 200. Typically epochs are set to be between 50 and 200.

[22]The vocabularies are geometrically close vocabularies in the embedding spaces trained on comments. Please see the Appendix for more details concerning R&D and Technology vocabularies and emission reduction vocabularies. The vector dimensionality of the analysis presented in the paper is 200, and the Appendix provides a robustness check using models with the dimensionality of 1,000, and 10,000.

$$\overrightarrow{R\&D} - \overrightarrow{Reduction} = \frac{\sum_n \overrightarrow{R\&D_n}}{|N_{R\&D}|} - \frac{\sum_n \overrightarrow{Reduction_n}}{|N_{Reduction}|}$$

Therefore, the vector difference corresponds to the issue slant towards the R&D direction and can be substantively interpreted as a degree to which a proposal is leaning towards the issue of R&D instead of emission cuts. Note that word vectors and document vectors live in the same space by the way that *Paragraph vector* is constructed. By the geometry of vector space, I measure the cosine of the angle between the inferred vectors of the issue slant and each document vector. The connotation of this approach is to measure the similarity of a comment to the dimension of the issue slant towards R&D and technology.[23]

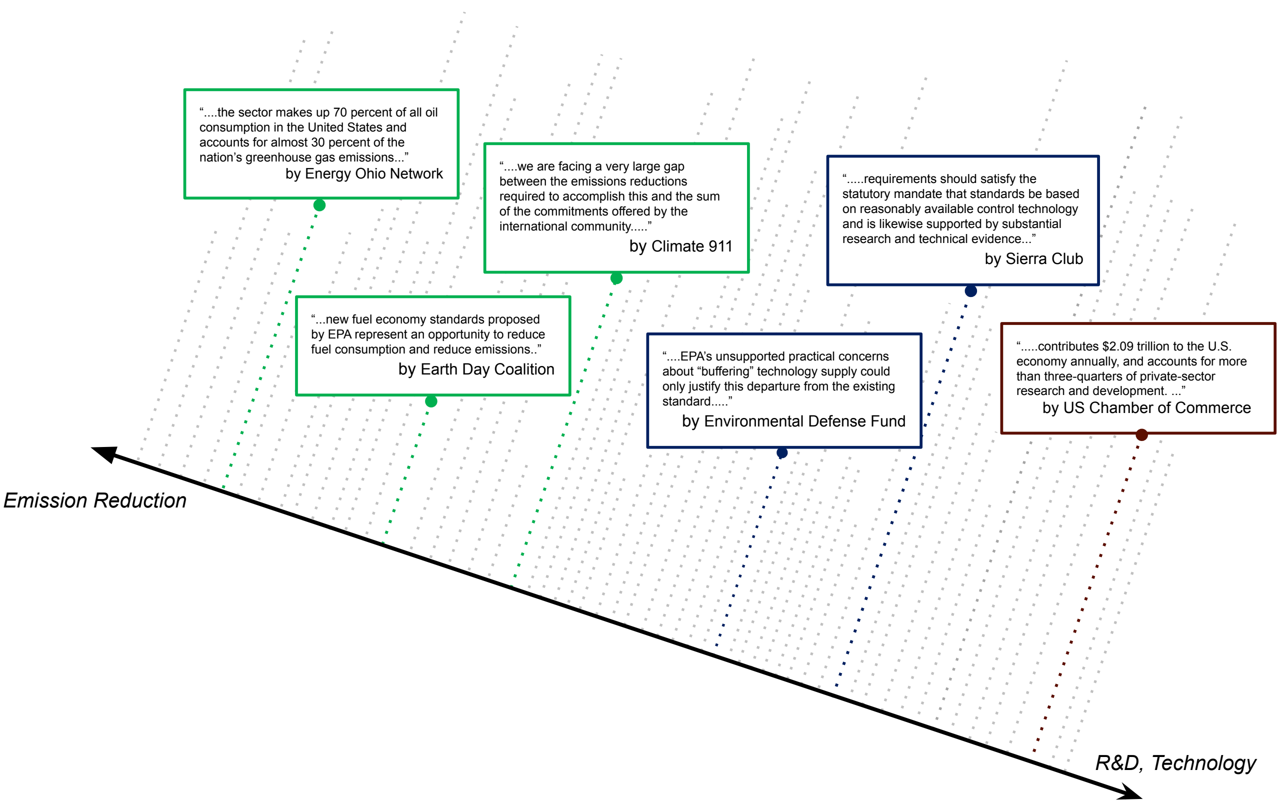


FIGURE 2. Schematic illustration of vector projection

[23]Please see the equation B.3 for the mathematical formula.

The similarity score, from -1 to 1, indicates the emphasis in a document on R&D compared with the emphasis on reductions. A score close to 1 suggests a tendency to emphasize R&D, whereas a negative score implies a skew toward emission reductions. Figure 2 depicts a schematic representation of the vector projection used in this method. It is evident that comments submitted by environmental groups in collaboration with business partners, such as the Sierra Club or Environmental Defense Fund, exhibit a tendency towards *R&D* and technology-related aspects compared with comments from environmental groups that lack business partnerships. However, comments from partnerships are relatively less skewed compared with comments from business interests, which demonstrates a notable bias toward *R&D* directions.

I use the similarity score for each comment i submitted by k in a time period t as a dependent variable and run an ordinary least squares regression[24] Specifically, I estimate the following model:

$$\text{Similarity Score}_{ikt} = \alpha + \beta_1 \text{Strategic Partnership}_i + \delta Z_k + \tau_t + \epsilon_{ikt}$$

, where Z denotes the group-level control variable and τ are year-fixed effects. The specification controls for group-level characteristics because there might be a systematic difference in research capacities due to staff size.[25] I also include commenter-fixed effects so that the effects are partially identified off within commenter variation. The error term is ϵ_{ikt}.

The first column of Table 1 examines comments from environmental groups, both with and without business partners, and the second column is focused on comments from environmental groups with business partners, business associations, and individual firms. The reference category for the second column is business associations. The last column identifies a correlation between the slant towards *R&D* and technology and the types of

[24]The cosine similarity score used in Table 1 is measured with 6 vocabularies. For robustness checks, the same analyses are repeated with a different number of vocabularies, 1,2,3, and 9. Please see details in the Appendix.

[25]The summary statistics of comments are given in the Appendix.

TABLE 1. Regression Models Examining the Issue Slant toward R&D versus Greenhouse Gas Reductions

Sample	Partnerships + Environmental Groups	Partnerships + Business Association + Single Firms	Whole Sample
	(1)	(2)	(3)
Partnership	0.128**	−0.062***	0.016**
	(0.049)	(0.016)	(0.008)
Single firms		0.008	0.028***
		(0.013)	(0.005)
Business associations			0.042***
			(0.009)
Others			0.031***
			(0.007)
Staff Size	✓	✓	✓
Year FE	✓	✓	✓
Commenter FE	✓	✓	✓
Mean Outcome	0.06	0.092	0.085
Observations	225	683	903
R2 Adj.	0.318	0.456	0.066

*p < .1; **p < .05; ***p < .01. In the first column, the reference category is *environmental groups*, while in the second column, it is *business associations*. For the third column, the reference category remains *environmental groups*. Standard errors are clustered by notice and comment periods in parentheses.

comments in the entire dataset. Comments from environmental groups without business partners serve as the reference category.

The results show a positive, statistically significant effect of strategic partnerships on the slant toward R&D and technology when environmental groups partner with polluting firms (Column 1). More notably, comments from partnerships with environmental groups tend to be inclined towards emission reduction compared with comments from business associations or individual firms (Column 2). The magnitude of the coefficient related to partnerships in Column 1 is significantly greater than that in Column 2, indicating that environmental groups are more willing to make concessions to reach a compromise while the joint products are moderate to prevent extreme policies from business interests. In general, strategic partnerships between polluting firms and environmental groups exhibit a positive inclination towards *R&D* and technology (Column 3), but compared to environmental groups without business partners, its coefficient magnitude is considerably smaller than that of business interests. This empirical evidence lends support to the *Compromised Policy Outcome* hypothesis; the policy goals of firms and environmental groups are reconciled and the outcome favors the high-capacity group, to generate a compromise. The full results, including all control variables, are presented in Appendix. For robustness check, I construct another measure to capture the prevalence of the topic, a frequency-based metric of *R&D* and technology coverage. The details concerning the analysis are also presented in Appendix.

4.2. Achieving a Higher-Quality Proposal for a Higher-quality Policy Implementation

In this section, I examine whether environmental groups and firms achieve high-quality policy implementation despite a compromised policy. Focusing on the role of expertise in regulatory politics, I investigate the effect of strategic partnerships between firms and environmental groups by the amount of technical and analytical information in the comment.

To construct a measure of information quality, I apply an information retrieval technique to extract technical and informative chunks from unstructured raw text documents. The primary problem to be tackled when measuring information is the identification of scientific entities or languages that convey specific information. Although crowdsourcing is one method for performing manual, human-oriented tasks, the expertise required to extract scientific evidence or analytical facts makes crowd-sourcing impractical (Bonney et al. 2014, 2009). Therefore entity recognition techniques have been widely used in academic disciplines to quantify information (e.g., Liu et al. 2021; Hong et al. 2020). This technique operates by locating and classifying proper nouns into categories, such as organizations (e.g., companies, government organizations, committees), local-level knowledge (e.g., cities, countries, rivers) or measurement.[26] In total, eighteen categories

[26]The analysis presented in the paper is implemented by SpaCy v3.0, an open-source library for advanced language processing, on December 27, 2022. This transformer-based pipeline has an accuracy of 89.8.

are used to measure the amount of scientific information.[27]

Nevertheless, to the limited extent that information is available, the emission inventories indicate that fugitive or vented emissions from gas processors and compressors can be extremely high. For example: Merit Energy Company's ORG Anschutz Ranch East Gas Plant in Wyoming GPE reported 276 tons QUANTITY of fugitive VOCs in 2013 DATE . Merit Energy Company's Halfmoon Battery ORG reported 178.1 tons QUANTITY of vented and fugitive VOC NORP emissions in 2013 DATE . A fraction of that total came from fugitives (5.5 tons QUANTITY) while most was vented (172.6 tons QUANTITY). DCP Midstream's ORG East Texas Gas Plant ORG (RN102805272) leaked and vented 326 tons QUANTITY of VOCs in 2012 DATE and 206.5 tons QUANTITY in 2013 DATE . Exxon Mobil's ORG King Ranch Gas Plant ORG (RN102488517) released 306.87 tons QUANTITY of fugitive and vented VOC emissions in 2012 DATE and 142.63 tons QUANTITY in 2013 DATE ,

A. Comment Submitted by Clean Air Council

We encourage stricter controls for emissions and are concerned about keeping our country air clean from industrial wastes going into our air near homes, schools, and animal habitats. even during drilling, fracing, flaring are changing our air quality at a fast and increasing rate. We know of families near gas sites that complain of the continual odors and having headaches, nose-bleeds, throat issues and breathing issues from exposure. I have felt adversely affected by being near these sites within fifteen minutes TIME . Thank you for your attention to this important matter concerning our health.

B. Comment Submitted by Citizens For Clean Water

FIGURE 3. Illustration of Information Retrieval Techniques for Public Comments

Figure 3 illustrates the application of the information retrieval technique to comments. The colored boxes represent the technical details identified by this approach. Each box is marked to display the named entities identified by the technique. For instance, the example demonstrates that the named entity recognition technique successfully captures organizations discussed in the comment submitted by the Clean Air Council, such as Merit Energy Company or Exxon Mobile, as well as various locations such as King Ranch Gas Plant, East Texas Gas Plant, or Wyoming. Furthermore, the technique identifies quantities of emissions (e.g., 326 tons) and specific dates. However, in the comment submitted by Citizens for Clean Water, there are only a few colored boxes because the comment does not include any specific or scientific evidence. During the validation process, the frequency

[27] Eighteen classes include PERSON, NORP, FAC, ORG, GPE, LOC, PRODUCT, EVENT (Named hurricanes, battles, wars, sports events, etc.), WORK OF ART (titles of books, songs, etc), LAW (Named documents made into laws), LANGUAGE (any named language), DATE (absolute or relative dates or periods), TIME (times smaller than a day), PERCENT(percentage, including "%"), MONEY (monetary values, including unit), QUANTITY (measurements, as of weight or distance), ORDINAL("first", "second", etc.), CARDINAL (numerals that do not fall under another type). Please see the Appendix for further details.

of false-positive identifications is noticeably smaller than the frequency of false-negative identifications, suggesting that the named entity recognition provides a conservative measure of expertise. Additional details regarding human validations are presented in the Appendix.

I use the number of all the colored boxes in each comment as a measure of expertise and estimate the effect of a strategic partnership on it. Formally, the dependent variable is a count variable that represents the number of detected named entities in each comment. Negative binomial models are presented in the main analysis, considering the count data. Quasi-Poisson models are used as a robustness check, and the analysis is presented in the Appendix.

TABLE 2. Negative binomial model estimating the quantity of information

Sample	Partnerships + Environmental Groups		Partnerships + Business Association + Single Firm		Whole Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Partnership	0.705***	2.670***	0.760***	1.878***	0.864***	0.730***
	(0.149)	(0.366)	(0.141)	(0.582)	(0.148)	(0.135)
Single Firm			−0.355***	−0.033	−0.200*	−0.334***
			(0.090)	(0.538)	(0.117)	(0.089)
Business associations					0.134	
					(0.123)	
Environmental groups						−0.134
						(0.123)
Others					0.167	0.033
					(0.167)	(0.152)
Issue Slant	1.301	−0.123	−0.429	0.069	−0.237	−0.237
(*R&D* and Technology)	(0.791)	(1.426)	(0.530)	(1.479)	(0.462)	(0.462)
Staff Size	✓	✓	✓	✓	✓	✓
Commenter FE		✓		✓		
Year FE	✓	✓	✓	✓	✓	✓
Observations	225	225	683	683	903	903

*p < .1; **p < .05; ***p < .01. Standard errors are clustered by notice and comment periods in parentheses.

Table 2 presents the estimation results with marginal effects in the main entries and standard errors in parentheses.[28] I include commenter-fixed effects so that the results

[28]To economize the space, I present the estimation results only for the key variable of interest throughout the paper. The full results, including the control variables, are presented in the Appendix.

are robust to commenter-level time-invariant confounders. I also control the issue slant toward R&D and technology constructed in the previous section because that topic is likely to be accompanied by technical details. The results suggest that comments from strategic partnerships generally have more technical information than comments from other entities. *Partnership* appears to be positive and significant at 0.01 level across all models. Both environmental groups and firms benefit from strategic partnerships; comments from environmental groups partnering with firms contain a higher quantity of information versus comments from environmental groups without business information (Columns 1 and 2). Similarly, comments from partnerships are likely to include more technical information than comments from firms without environmental group partners (Columns 3 and 4). Overall, we observe that the magnitude of the partnership effect is significantly larger when the reference category is environmental groups (Column 5) as opposed to when the reference category is business associations (Column 6). These findings lend confidence to the theoretical expectations that both firms and environmental groups derive advantages from investing in joint efforts. Additionally, the results suggest that environmental groups can attract greater participation from business interests, resulting in higher output. This finding is consistent with multiple robustness checks, including quasi-Poisson models. Please see the Appendix for further details.

4.3. Political Influence of Strategic Partnerships on Regulatory Outcome

I investigate the political influence of strategic partnerships on environmental regulations by examining whether comments from the joint efforts of firms and environmental groups affect policy amendments during the notice and comment period. Specifically, I estimate the effects of *Partnerships* on two dependent variables: (1) the divergence scores from information theory and (2) a binary variable that indicates whether a comment was cited by EPA officials in the final rule after the notice and comment period.

Quantifying Political Influence Using Information Theory

In this section, I examine if the increased quantity of knowledge translates into political power in regulatory politics by capturing distribution similarity. The intuition of this analysis is to examine how likely is it that a comment and policy amendment come from the same probability distribution. I particularly use divergence scores from information theory as relative entropy captured via divergence score denotes how close two samples are to each other. Given that the vectors in this context indicate probability distributions, the cosine angle is inappropriate because it fits for vector space modeling. Therefore, I employ Jensen-Shannon (JS) divergence score as a metric of statistical distance.[29] JS divergences have already been widely used in social science research as a similarity measure of sparse data.[30] Divergence scores close to 0 indicate a closer statistical distance, implying that two samples are likely to be from the same probabilistic distribution.

A finalized rule is generally a hundred-page document, whereas policy comments tend to focus on a few provisions of a proposed policy. Capturing the statistical distance between each comment and a huge corpus of the entire policy would underestimate the influence of each comment on rulemaking, because a finalized rule is sparse and particular provisions are supposed to be examined during the notice-and-comment process. Therefore, I construct a set of clauses updated after the notice-and-the-comment period and use the set as a basis of analysis to quantify the influence of comments on finalized policy outcome. If a policy amendment is likely to be from the same distribution of comments by partnerships of environmental groups and firms, we can infer that the joint efforts of firms and environmental groups exercise political leverage over climate regulations. There might be some concerns that this analysis would capture linguistic similarity or legal

[29] The Kullback-Leibler (KL) measure is inappropriate in this context as it is an asymmetric measure, leading to different scores for A to B and B to A. The algebraic reason is that $D(P||O)-D(O||P)$ is equal to $\sum_i^n ln(\frac{P_i}{O_i})(P_i+O_i)$ and there is no reason for this to be 0. Please refer to the Appendix for further details.

[30] Please see Section D in the Appendix for the mathematical proofs justifying the use of JS divergence as a test statistic and the detailed procedure of this analysis.

formalism between comments and policies, instead of their influence on policy changes. To address this concern, I control the JS divergence score to a proposed policy posted by EPA officials before the notice and comment period. The model specification is similar to the one estimated in the previous section, with the exception that I include administration fixed effects because Republican politicians are generally considered business-friendly and prioritize policies that put business interests over environmental concerns.

Capturing Political Influence Using Citations by EPA officials

After the notice and comment period, EPA officials consider the comments submitted on a proposed policy and decide whether to revise the regulations accordingly when issuing a final rule.[31] When posting the finalized amendments, EPA officials add supplementary information; they provide a broad executive summary and explanations on the regulatory background of final standards. In addition, EPA officials summarize the significant comments, and they respond to those comments in a document that announces a final rule.

To estimate the influence of strategic partnerships on regulatory outcomes, I specifically focus on a final rule that was posted on March 12, 2018. The finalized policies posted by EPA officials take various inconsistent forms. In most cases, EPA officials make broad and generic statements that summarize the collection of comments without referencing specific commenters or comment IDs. However, for the March 2018 rule, the officials explicitly included comment IDs or commenters that regulators considered to update a proposed policy. Using this final rule as the basis for analysis, I construct a binary indicator that is coded as 1 if a comment is specifically cited by EPA officials in their response.[32]

[31] Sometimes the agency extends or reopens a comment period because it has not received enough comments. Similarly, the agency may find that people have raised new issues in their comments that were not previously considered in the initial proposed policy. As new issues or additional complexity arises, the agency may publish a series of proposed rules in the Federal Register.

[32] The purpose of opening the notice and comment period in 2017 and 208 was to make amendments to two specific provisions related to the requirements for the collection of emission components at well sites. In the final rule, the agency announced the removal of the requirement for the repair of a component within

Alternative Explanations

The primary focus of the analysis centers on the quality of policy implementation to comprehend the dynamics of regulatory policymaking. However, the decision of firms and environmental groups to collaborate could result from a multifaceted strategic interaction. An alternative explanation as evidenced by prior studies could be that regulators may find the diversity within partnerships more appealing (Lorenz 2020; Phinney 2017; Mahoney 2007) because regulators typically seek indications of broad support for a policy proposal (Esterling 2009).[33] To consider this potential scenario, I combine a unique dataset of public comments on greenhouse gas emission standards with interest group ideal point estimates, referred to as "IGscore," introduced by Crosson et al. (2020). Then, I estimate the preference gap by calculating the absolute difference between the highest IGscore of firms and the lowest IGscore of environmental groups.[34] For single entities, the absolute difference is 0.

Table 3 presents the estimation results, separately for different reference categories. In all models, *Partnership* decreases the statistical distance and its effect is statistically significant (Columns 1- 2). A finalized policy outcome tends to have a closer statistical distance to comments from joint efforts, namely more informative comments that contain a larger amount of scientific reasoning and specific evidence. This demonstrates that enhanced expertise as a result of joint efforts by firms and environmental groups translates into political power in the rulemaking process, controlling the difference between IGscores. Columns 3-4 further show that comments from strategic partnerships are more

30 days of the detection of fugitive emissions.

[33] Most literature on coalition lobbying relies on a signaling model that policymakers find diverse coalitions' signal more credible for the following reasons. Interest-diverse coalitions can synergize their advocacy tactics and network, and they send a more heterogeneous signal to legislators about the quality of a legislative proposal. Third, diverse coalitions are harder to maintain, making their legislative signals costlier. Thus, legislators have reason to believe that bills favored by diverse coalitions are more deserving of their attention and support than those favored by homogeneous coalitions, all else equal. However, it is worth pointing out that the canonical signaling models including Crawford and Sobel (1982) do not lead to policy bias but only to the reduction of uncertainty.

[34] Environmental groups tend to work with multiple business partners.

TABLE 3. Regression Models Estimating JS Divergence Scores and Citation by EPA Officials

	JS Divergence Scores (OLS)		*Citation By EPA Officials* (Probit)	
	(1)	(2)	(3)	(4)
Partnership	−0.031***	−0.015*	0.512	1.737**
	(0.003)	(0.009)	(0.679)	(0.720)
Environmental groups	−0.015**		−1.225***	
	(0.007)		(0.396)	
Single firm	−0.001	0.014***	−0.445	0.779*
	(0.003)	(0.004)	(0.346)	(0.426)
Business associations		0.015**		1.225***
		(0.007)		(0.396)
Others	−0.015***	−0.0001	−5.014	−3.789
	(0.006)	(0.001)	(228.534)	(228.534)
Absolute difference between IGscores	0.002	0.002	0.207	0.207
	(0.004)	(0.004)	(0.497)	(0.497)
Staff Size	✓	✓	✓	✓
Administration FE	✓	✓		
JS Divergence to a proposed policy	✓	✓		
Observations	903	903	181	181

*p<0.1; **p<0.05; ***p<0.01. Standard errors are clustered by notice and comment periods in parentheses (columns 1 and 2). Since our analysis focuses on single notice and comment periods that overlap both the Obama and Trump administrations, we do not have control over the years of submission. Therefore, the analysis using citation patterns by EPA officials (Columns 3 and 4) does not consider the years of submission as a controlled factor.

cited by EPA officials. Which types of comments have the stronger influence on policy amendments? If the signaling perspective holds true, a higher absolute difference between IGscores would lead to a reduced statistical distance to a finalized policy or more citations by EPA officials. However, we do not find any effect of IGscores on the two measures.[35]

The analysis using two measures of political influence provides the evidence for my *Political Influence* hypothesis. Comments from joint efforts by firms and environmental groups tend to have a closer statistical distance to policy amendments and are likely to be cited by EPA officials. The full results are presented in Appendix Table D.1. One might

[35]In Appendix, I run another analysis without any consideration of commenter types and observe a negative relationship between the absolute difference between IGscores and JS divergence scores. Please see Table D.2

assume that moderate and high-capacity groups are more likely to engage in strategic partnerships and that the capacity of individual groups would drive the result. However, although my model does point to the selection into partnerships, it is important to note that the analysis presented in this paper emphasizes the output from collaborative efforts instead of the capacities or moderation of individual groups. By working together, interest groups can generate more informative comments that cannot be attained by individual efforts. Enhanced expertise, as a result of collaboration, translates into political influence in regulatory politics, thereby contributing to better-quality policy implementation.

This finding presents a new empirical implication regarding the influence of interest groups on policy implementation. In contrast to the argument made by Yackee and Yackee (2006) that business comments are most commonly associated with policy changes, my research reveals that comments stemming from joint efforts involving experts with diverse expertise wield greater political influence during the rulemaking process. This empirical evidence reinforces the theoretical predictions put forth by the policymaking literature (e.g., McCarty 2020; Hirsch and Shotts 2012) that collaborative efforts by involved actors can increase the quality of policy implementation by conveying more informative proposals to regulators.

5. Conclusion

Interest groups play a crucial role in policymaking. Canonical models of policymaking focus primarily on how interest groups compete using their policy-relevant information to realize their political interests. Conversely, empirical evidence points to interest-diverse coalitions in which political actors with divergent interests cooperate. What incentivizes political actors to work together despite contrasting policy goals? What does a compromise look like and why would they invest in joint efforts despite compromise?

In this paper, I tackle this question by focusing on the dynamics of regulatory policymaking. Given that regulatory policymaking involves the development of technical and fine-grained details of a policy, I expect that compromise arises endogenously because involved parties are incentivized to produce high-quality policy implementation. By using unique data of public comments officially submitted on greenhouse gas emission standards and employing information retrieval techniques, I demonstrate that environmental groups and polluting firms craft public comments that incorporate a greater amount of scientific evidence and analytical information in comparison with other types of comments. On the basis of information theory and citation patterns by EPA officials, I further show that the enhanced expertise of strategic partnerships between firms and environmental groups exercises the biggest leverage on the final policy, even when controlling for the difference in ideology scores of the partnered interest groups.

Specifically, this paper contributes to the growing literature on understanding the influence of interest groups on environmental regulations. By leveraging recent developments in machine learning techniques, I uncover that regulated firms' informational advantage leads to compromised policy outcomes that align with their preferences. However, despite these concessions, environmental groups with business partners gain advantages by accessing business information and resources, as well as attracting greater participation from business interests. As a result, partnered environmental groups achieve greater political leverage in regulation politics compared with environmental groups that lack business partnerships.

This study is primarily focused on high-quality policy implementation as an explanation for the motives behind investing in joint efforts amid political rivalry. However, an alternative explanation for partnerships could be competition among environmental groups. There has been ongoing disagreement among environmentalists regarding strategies to reduce carbon emissions,[36] and some environmental groups may find it more beneficial to collaborate with business interests to amplify their voices, as opposed to

[36]See Pulkkinen, Levi. (2021, March 12). Washington climate activists disagree about how to cut carbon,https://crosscut.com/environment/2021/03/washington-climate-activists-disagree-about-how-cut-carbon

working solely with other climate activists. These dynamics, particularly in relation to how interest groups select their partners and navigate these complex relationships, are the subject for further research.

My results contribute to our understanding of unexplored dynamics of regulatory politics. There are a variety of mechanisms used by interest groups that attempt to lobby agency rulemaking. By highlighting the strategic partnership of environmental groups and business interests - an overlooked channel of influence- I contribute to efforts in capturing diverse circuits of the political influence of interest groups. A broader set of political instruments are available to interest groups in regulation politics, and a valuable direction for future research is identifying and systematically measuring the role of other less visible channels of influence. Recognizing the various channels of influence, and their magnitudes of impacts, can contribute to a better understanding of interest group politics and its implication on regulatory politics.

References

Aidt, Toke S (1998). Political internalization of economic externalities and environmental policy. *Journal of Public Economics 69*(1), 1–16.

Alchian, Armen A and Harold Demsetz (1972). Production, information costs, and economic organization. *The American economic review 62*(5), 777–795.

Balla, Steven J and John R Wright (2001). Interest groups, advisory committees, and congressional control of the bureaucracy. *American Journal of Political Science*, 799–812.

Baron, David P and John A Ferejohn (1989). Bargaining in legislatures. *American political science review 83*(4), 1181–1206.

Baumgartner, Frank R , Jeffrey M Berry, Marie Hojnacki, Beth L Leech, and David C Kimball (2009). *Lobbying and policy change: Who wins, who loses, and why*. University of Chicago Press.

Baumgartner, Frank R and Bryan D Jones (2010). *Agendas and instability in American politics*. University of Chicago Press.

Bernauer, Thomas and Ladina Caduff (2004). In whose interest? pressure group politics, economic competition and environmental regulation. *Journal of Public Policy 24*(1), 99–126.

Berry, Jeffrey M and Clyde Wilcox (2015). *The interest group society*. Routledge.

Bertrand, Marianne , Matilde Bombardini, Raymond Fisman, and Francesco Trebbi (2020). Tax-exempt lobbying: Corporate philanthropy as a tool for political influence. *American Economic Review 110*(7), 2065–2102.

Bolden, Isaiah W , Sasha K Seroy, Emily A Roberts, Lauren Schmeisser, J Zachary Koehn, Canita H Rilometo, Emerson Lopez Odango, Corrin Barros, Julian P Sachs, and Terrie Klinger (2018). Climate-related community knowledge networks as a tool to increase learning in the context of environmental change. *Climate Risk Management 21*, 1–6.

Bonney, Rick , Caren B Cooper, Janis Dickinson, Steve Kelling, Tina Phillips, Kenneth V Rosenberg, and Jennifer Shirk (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience 59*(11), 977–984.

Bonney, Rick , Jennifer L Shirk, Tina B Phillips, Andrea Wiggins, Heidi L Ballard, Abraham J Miller-Rushing, and Julia K Parrish (2014). Next steps for citizen science. *Science 343*(6178), 1436–1437.

Breyer, Stephen G (1982). Regulation and its reform. In *Regulation and Its Reform*. Harvard University Press.

Bromley-Trujillo, Rebecca , James W Stoutenborough, Kellee J Kirkpatrick, and Arnold Vedlitz (2014). Climate scientists and environmental interest groups: The intersection of expertise and advocacy. *Politics, Groups, and Identities 2*(1), 120–134.

Brulle, Robert and Christian Downie (2022). Following the money: trade associations, political activity and climate change. *Climatic Change 175*(3-4), 11.

Carpenter, Daniel P and David Moss (2014). Preventing capture: Special interest influence in regulation, and how to prevent it.

Cheon, Andrew and Johannes Urpelainen (2013). How do competing interest groups influence environmental policy? the case of renewable electricity in industrialized democracies, 1989–2007. *Political Studies 61*(4), 874–897.

Choi, Dahyun (2023). Investing in team efforts: When agents with contrasting preferences collaborate. *Working Paper*.

Coglianese, Cary (2007). Business interests and information in environmental rulemaking. *Business and environmental policy: Corporate interests in the American political system*, 185–210.

Coglianese, Cary and David Lazer (2003). Management-based regulation: Prescribing private management to achieve public goals. *Law & Society Review 37*(4), 691–730.

Colgan, Jeff D , Jessica F Green, and Thomas N Hale (2021). Asset revaluation and the existential politics of climate change. *International Organization 75*(2), 586–610.

Cory, Jared , Michael Lerner, and Iain Osgood (2021). Supply chain linkages and the extended carbon coalition. *American Journal of Political Science 65*(1), 69–87.

Crawford, Vincent P and Joel Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Crosson, Jesse M , Alexander C Furnas, and Geoffrey M Lorenz (2020). Polarized pluralism: organizational preferences and biases in the american pressure system. *American Political Science Review 114*(4), 1117–1137.

Culhane, Trevor , Galen Hall, and J Timmons Roberts (2021). Who delays climate action? interest groups and coalitions in state legislative struggles in the united states. *Energy Research & Social Science 79*, 102114.

Downie, Christian (2017). Business actors, political resistance, and strategies for policymakers. *Energy Policy 108*, 583–592.

Dwidar, Maraam A (2022). Coalitional lobbying and intersectional representation in american rulemaking. *American Political Science Review 116*(1), 301–321.

Epstein, Dmitry , Cynthia Farina, and Josiah Heidt (2014). The value of words: Narrative as evidence in policy making. *Evidence & Policy 10*(2), 243–258.

Epstein, David and Sharyn O'halloran (1995). A theory of strategic oversight: Congress, lobbyists, and the bureaucracy. *The Journal of Law, Economics, and Organization 11*(2), 227–255.

Esterling, Kevin (2009). *The political economy of expertise: Information and efficiency in American national politics*. University of Michigan Press.

Furlong, Scott R and Cornelius M Kerwin (2005). Interest group participation in rule making: A decade of change. *Journal of Public Administration Research and Theory 15*(3), 353–370.

Gailmard, Sean and John W Patty (2012). *Learning while governing: Expertise and accountability in the executive branch*. University of Chicago Press.

Gilligan, Thomas W and Keith Krehbiel (1989). Asymmetric information and legislative rules with a heterogeneous committee. *American journal of political science*, 459–490.

Golden, Marissa Martino (1998). Interest groups in the rule-making process: Who participates? whose voices get heard? *Journal of Public Administration Research and Theory 8*(2), 245–270.

Gordon, Sanford C and Steven D Rashin (2021). Stakeholder participation in policy making: Evidence from medicare fee schedule revisions. *The Journal of Politics 83*(1), 409–414.

Grumbach, Jacob M (2015). Polluting industries as climate protagonists: Cap and trade and the problem of business preferences. *Business and Politics 17*(4), 633–659.

Haeder, Simon F and Susan Webb Yackee (2015). Influence and the administrative process: Lobbying the us president's office of management and budget. *American Political Science Review 109*(3), 507–522.

Hall, Richard L and Kristina C Miler (2008). What happens after the alarm? interest group subsidies to legislative overseers. *The Journal of Politics 70*(4), 990–1005.

Hawkins, Keith and John M Thomas (1989). Making policy in regulatory bureaucracies. *Making Regulatory Policy*, 3–30.

Heaney, Michael T and Philip Leifeld (2018). Contributions by interest groups to lobbying coalitions. *The Journal of Politics 80*(2), 494–509.

Heaney, Michael T and Geoffrey M Lorenz (2013). Coalition portfolios and interest group influence over the policy process. *Interest Groups & Advocacy 2*, 251–277.

Heinz, John P , Edward O Laumann, and Robert L Nelson (1993). *The hollow core: Private interests in national policy making*. Harvard University Press.

Hirsch, Alexander V (2022). Productive policy competition and asymmetric extremism.

Hirsch, Alexander V and Kenneth W Shotts (2012). Policy-specific information and informal agenda power. *American Journal of Political Science 56*(1), 67–83.

Hirsch, Alexander V and Kenneth W Shotts (2015). Competitive policy development. *American Economic Review 105*(4), 1646–1664.

Hong, Zhi , Roselyne Tchoua, Kyle Chard, and Ian Foster (2020). Sciner: extracting named entities from scientific literature. In *Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part II 20*, pp. 308–321. Springer.

Huber, John D and Charles R Shipan (2002). *Deliberate discretion?: The institutional foundations of bureaucratic autonomy*. Cambridge University Press.

Hula, Kevin W (1999). *Lobbying together: Interest group coalitions in legislative politics*. Georgetown University Press.

Ingram, Helen M and Scott J Ullery (1977). Public participation in environmental decision-making: Substance or illusion? *Public participation in planning*, 123–139.

Junk, Wiebke Marie (2019). When diversity works: The effects of coalition composition on the success of lobbying coalitions. *American Journal of Political Science 63*(3), 660–674.

Keohane, Nathaniel O , Richard L Revesz, and Robert N Stavins (1998). The choice of regulatory instruments in environmental policy. *Harv. Envtl. L. Rev. 22*, 313.

Krehbiel, Keith (2010). *Pivotal politics: A theory of US lawmaking*. University of Chicago Press.

Lazarus, Richard (2014). Environmental law without congress. *J. Land Use & Envtl. L. 30*, 15.

Le, Quoc and Tomas Mikolov (2014). Distributed representations of sentences and documents. In

International conference on machine learning, pp. 1188–1196. PMLR.

Lerner, Michael and Iain Osgood (2022). Across the boards: Explaining firm support for climate policy. *British Journal of Political Science*, 1–24.

Libgober, Brian et al. (2020). Meetings, comments, and the distributive politics of rulemaking. *Quarterly Journal of Political Science 15*(4), 449–481.

Liu, Zihan , Yan Xu, Tiezheng Yu, Wenliang Dai, Ziwei Ji, Samuel Cahyawijaya, Andrea Madotto, and Pascale Fung (2021). Crossner: Evaluating cross-domain named entity recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Volume 35, pp. 13452–13460.

Livermore, Michael A and Richard L Revesz (2015). Interest groups and environmental policy: Inconsistent positions and missed opportunities. *Environmental Law*, 1–17.

Lorenz, Geoffrey Miles (2020). Prioritized interests: Diverse lobbying coalitions and congressional committee agenda setting. *The Journal of Politics 82*(1), 225–240.

Magat, Wesley , Alan J Krupnick, and Winston Harrington (2013). *Rules in the making: A statistical analysis of regulatory agency behavior*. Routledge.

Mahoney, Christine (2007). Networking vs. allying: the decision of interest groups to join coalitions in the us and the eu. *Journal of European Public Policy 14*(3), 366–383.

McCarty, Nolan (2017). The regulation and self-regulation of a complex industry. *The Journal of Politics 79*(4), 1220–1236.

McCarty, Nolan (2020). Team policy production. *Available at SSRN 3505655*.

McCubbins, Mathew D and Thomas Schwartz (1984). Congressional oversight overlooked: Police patrols versus fire alarms. *American journal of political science*, 165–179.

McKay, Amy and Susan Webb Yackee (2007). Interest group competition on federal agency rules. *American Politics Research 35*(3), 336–357.

Moffitt, Susan L (2014). *Making policy public: Participatory bureaucracy in American democracy*. Cambridge University Press.

Nelson, David and Susan Webb Yackee (2012). Lobbying coalitions and government policy change: An analysis of federal agency rulemaking. *The Journal of Politics 74*(2), 339–353.

Phinney, Robin (2017). *Strange bedfellows: Interest group coalitions, diverse partners, and influence in American social policy*. Cambridge University Press.

Rodriguez, Pedro L and Arthur Spirling (2022). Word embeddings: What works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics 84*(1), 101–115.

Rothenberg, Lawrence S (2018). *Policy Success in an Age of Gridlock: How the Toxic Substances Control Act was Finally Reformed*. Cambridge University Press.

Sautner, Zacharias , Laurence van Lent, Grigory Vilkov, and Ruishen Zhang (2020). Firm-level climate change exposure. *European Corporate Governance Institute–Finance Working Paper* (686).

Schlichting, Inga (2013). Strategic framing of climate change by industry actors: A meta-analysis. *Environmental Communication: A Journal of Nature and Culture 7*(4), 493–511.

Shipan, Charles R (2004). Regulatory regimes, agency actions, and the conditional nature of congressional influence. *American Political Science Review 98*(3), 467–480.

Stokes, Leah Cardamore (2020). *Short circuiting policy: Interest groups and the battle over clean energy and climate policy in the American States*. Oxford University Press, USA.

Supran, Geoffrey and Naomi Oreskes (2021). Rhetoric and frame analysis of exxonmobil's climate change communications. *One Earth 4*(5), 696–719.

Urpelainen, Johannes and Thijs Van de Graaf (2018). United states non-cooperation and the paris agreement. *Climate Policy 18*(7), 839–851.

Wagner, Wendy E (2003). Commons ignorance: the failure of environmental law to produce needed information on health and the environment. *Duke LJ 53*, 1619.

Warren, Kenneth F (2018). *Administrative law in the political sys*. Routledge.

Weingast, Barry R (1984). The congressional-bureaucratic system: a principal agent perspective (with applications to the sec). *Public choice 44*(1), 147–191.

Yackee, Jason Webb and Susan Webb Yackee (2006). A bias towards business? assessing interest group influence on the us bureaucracy. *The Journal of Politics 68*(1), 128–139.

You, Hye Young (2017). Ex post lobbying. *The Journal of Politics 79*(4), 1162–1176.

Supporting Information for

Teaming up Across Political Divides: Evidence from Climate Regulations

Dahyun Choi

May 2023

*Ph.D. Candidate, Department of Politics, Princeton University. Email: dahyunc@princeton.edu

A. Summary Statistics

Table A.1 presents the summary statistics on the entity/organization comments submitted for the Greenhouse gas emissions standards for the period 2010-2020. There were no entity comments submitted in 2019 and 2020. I extract the information from the *Regulations.gov* (https://www.regulations.gov/docket/EPA-HQ-OAR-2010-0505). There is no correlation between comment types and the year of submission.

TABLE A.1. Frequency of Comments By Year×Types

	2011	2012	2013	2014	2015	2016	2017	2018
Partnerships	14	0	0	1	22	3	19	12
Environmental Groups	28	1	1	1	56	4	48	15
Business Associations	44	4	11	12	122	12	29	16
Single Firms	42	0	5	5	210	25	24	51
Others	6	1	2	2	27	3	14	11

Table A.2 shows the summary statistics of key variables measured and constructed in the empirical analysis.

TABLE A.2. Descriptive Statistics on Key Variables

	mean	sd	median	min	max	range
Staff Size	6570.66	38276.32	141.00	1.00	543000.00	542999.00
NER score	153.61	169.51	79.00	0.00	1461.00	1461.00
Environmental Groups' IGscore	-0.88	0.75	-1.06	-2.65	1.82	4.47
Firms' IGscore	0.55	0.41	0.57	-1.20	1.78	2.98
difference between IGscores	0.10	0.37	0.00	0.00	2.25	2.25
Issue Slant to R&D	0.09	0.08	0.09	-0.20	0.33	0.53

ACEEE
American Council for an Energy-Efficient Economy

529 14th Street, N. W., Suite 600 Washington, D.C. 20045 202.507.4000 202.429.2248 www.aceee.org

To:
Air and Radiation Docket and Information Center,
Environmental Protection Agency, Mail code: 28221T,
1200 Pennsylvania Ave. N.W., Washington, DC 20460

Docket Management Facility, M–30,
U.S. Department of Transportation,
West Building, Ground Floor, Rm. W12–140,
1200 New Jersey Avenue S.E., Washington, DC 20590

From:
Therese Langer and Siddiq Khan, American Council for an Energy-Efficient Economy (ACEEE)

Re: Docket ID Nos. NHTSA-2014-0132 and EPA-HQ-OAR-2014-0827

Date: October 1, 2015

Attached please find the comments of the American Council for an Energy-Efficient Economy (ACEEE) on EPA and NHTSA's Proposed Greenhouse Gas Emissions and Fuel Efficiency Standards for Medium- and Heavy-Duty Engines and Vehicles; Phase 2.

ACEEE, a nonprofit, 501(c)(3) organization, acts as a catalyst to advance energy efficiency policies, programs, technologies, investments, and behaviors. We believe that the United States can harness the full potential of energy efficiency to achieve greater economic prosperity, energy security, and environmental protection for all its people. ACEEE carries out its mission by:

- Conducting in-depth technical and policy analyses
- Advising policymakers and program managers
- Working collaboratively with businesses, government officials, public interest groups, and other organizations
- Convening conferences and workshops, primarily for energy efficiency professionals
- Assisting and encouraging traditional and new media to cover energy efficiency policy and technology issues
- Educating consumers and businesses through our reports, books, conference proceedings, press activities, and websites

ACEEE was founded in 1980 by leading researchers in the energy field.

We appreciate this opportunity to provide comment on the agencies' proposal. Unless otherwise indicated, page references in the comments that follow refer to the proposed rule as it appeared in the Federal Register on July 13, 2015 (FR Vol. 80, No. 133).

FIGURE A.1. A Proposal submitted by Siddiq Khan, Senior Researcher and Lead, Heavy-Duty Vehicle Work, American Council for an Energy-Efficient Economy (ACEEE)

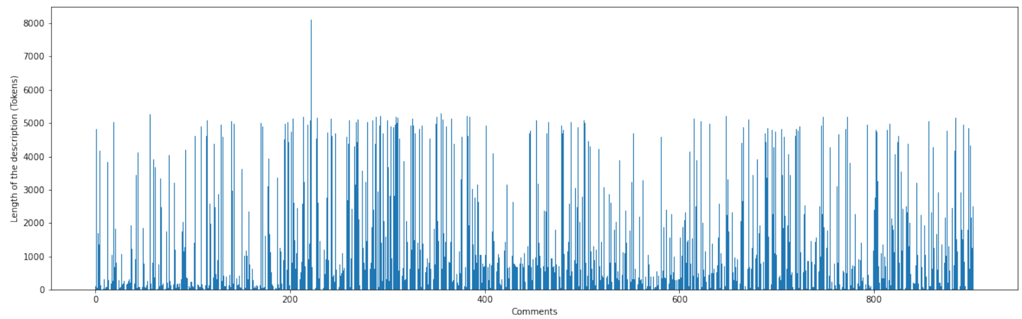


FIGURE A.2. Length of comments in the company/organization sample

TABLE A.3. Examples of Commenters in the Parternship Category

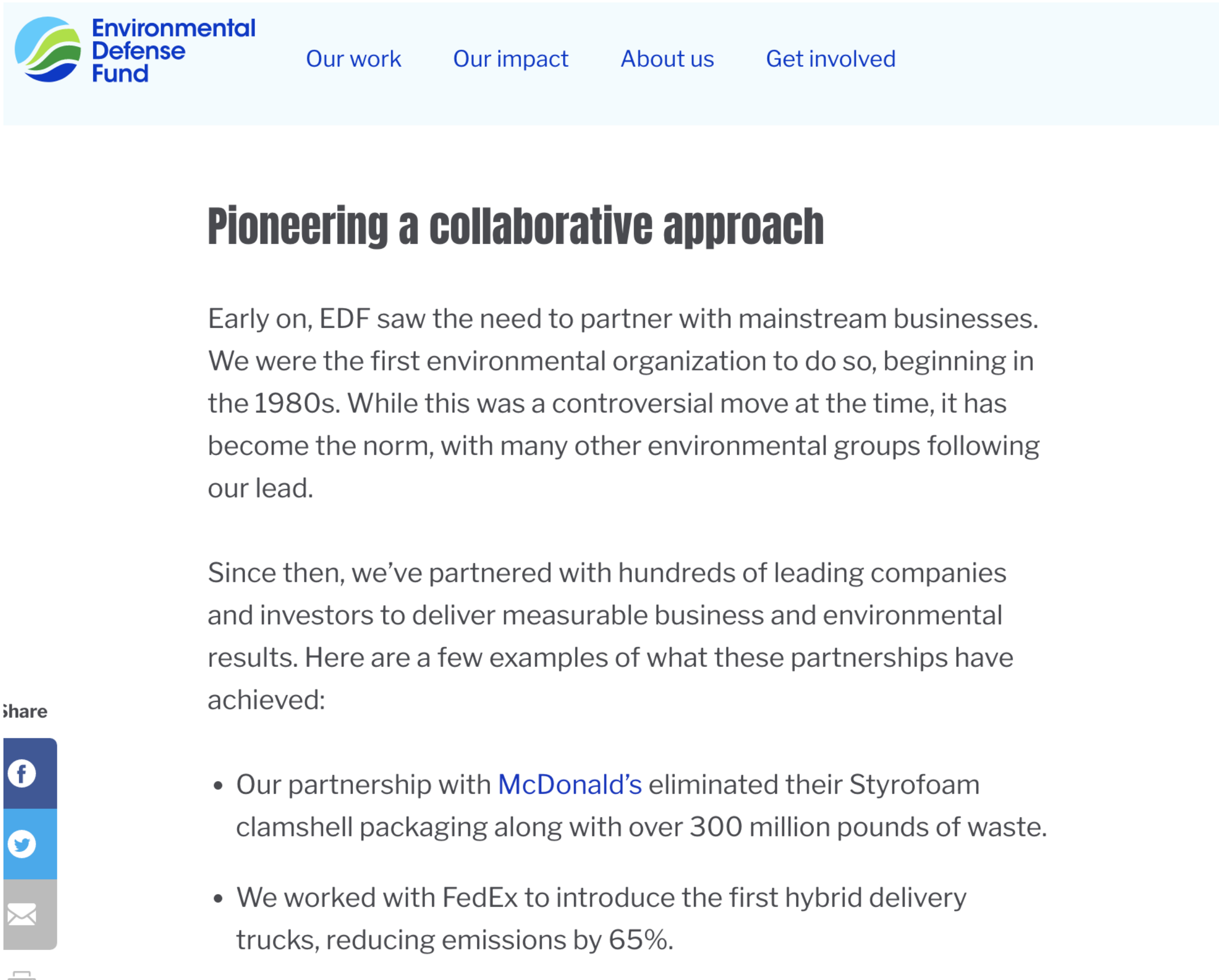
	partner1	partner2
1	Sierra Club	Amazon, Google, Patagonia, The North Face
2	Chesapeake Bay Foundation	Lockheed Martin
3	National Wildlife Federation	General Motors Company
4	Environmental Defense Fund	FedEx Corporation, Walmart, Lyft
5	Trout Unlimited	The Coca-Cola Company, Con Edison
6	Institute for Energy Research	koch
7	Clean Air Task Force	Caterpillar
8	FreedomWorks Foundation	Koch foundations
9	Resources for the Future	ConocoPhillips
10	Diesel Emissions Reduction Act (DERA) coalition	Caterpillar Inc
11	American Council for an Energy-Efficient Economy	Dow Chemical Co
12	Clean Fuels Ohio	American Electric Power

B. Compromised Policy Outcome Biased Towards Polluting Firms

B.1. Strategic use of the topic of R&D and technology by business interests

Correlation

In this section, I empirically test the qualitative evidence that business actors strategically mention more about R&D and technology to take attention away from emission standards. I compare the salience of issues focused on research and development, with indicators that reflect the firms' demand for regulation intensity and test the hypothesis that the proposed coverage of research and development is intended to distract the attention away



Environmental Defense Fund

Our work Our impact About us Get involved

Pioneering a collaborative approach

Early on, EDF saw the need to partner with mainstream businesses. We were the first environmental organization to do so, beginning in the 1980s. While this was a controversial move at the time, it has become the norm, with many other environmental groups following our lead.

Since then, we've partnered with hundreds of leading companies and investors to deliver measurable business and environmental results. Here are a few examples of what these partnerships have achieved:

- Our partnership with McDonald's eliminated their Styrofoam clamshell packaging along with over 300 million pounds of waste.
- We worked with FedEx to introduce the first hybrid delivery trucks, reducing emissions by 65%.

Share

FIGURE A.3. Environmental Defense Fund's Website- Business Partnership

from stringent emission cuts. I first perform an initial and simplistic study of this agenda-setting strategy. I define *R&D* coverage following the comparative agendas project and added specific conditions that mention research and development at least twice. I show the proportion of comments discussing *R&D* in Figure B.1. I found a strong positive correlation (γ_1= 0.8722452 for Carbon Dioxide emission, γ_2 =0.913333 for Methane emission); this suggests that mentions of research or development increase as emission increase. This positive correlation indicates the possibility of strategic use of agenda-setting by business interests.

Next, I extend these preliminary results in various ways. First I redefine the definition of *R&D* coverage by using two metrics: comment level, the number of articles that mention

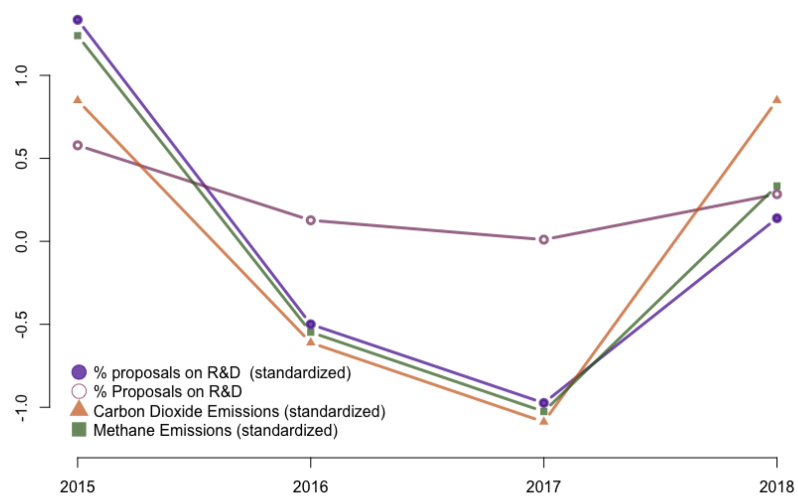


FIGURE B.1. Proportion of regulatory proposals by business interests groups that mention research, technology and development at least twice and emission (Yearly emission data is from EPA)

research, development, or technology at least twice, normalized by the total number of proposals in the time slice; and word level, the frequency of the occurrences of the research and development topics normalized by the total count of words in the time slice. Next, I refine the time resolution and use year and month levels of analysis. For monthly analysis, I use monthly average-emission data measured at ESRL Global Monitoring Laboratory.

Table B.1 describes the correlation between metrics of R&D coverages and climate change indicators at year and month levels. At all levels, there exist correlations between the proportion of agenda focused on research and development and CO2 emissions.

Granular Analysis of Granger Causality

Next, I hypothesize that these correlations are directed in fact: Greenhouse gas emissions are followed by business interests' coverage of research and development. To scrutinize this conjecture, I combine Granger Causality test with measures based on textual analysis. The crux of Granger causality is that we can identify if cause precedes effects. Therefore,

Levels	Proposal-level	Word-level
U.S. Atmospheric CO2 (Global Monitoring Laboratory, Monthly Averages)*	0.3723696	0.3364334
CO2 Emissions (EPA, Yearly)**	0.3328969	0.4501426

CO2 measurements from flask-air samples scaled by National Oceanic and Atmospheric Administration (NOAA)

Emissions are expressed in million metric tons of carbon dioxide equivalents.

TABLE B.1. Pearson's correlation between proposal coverage of the Research and Development and emission indicators

a time series X is said to Granger-cause a time series values Y if past values x_{t-i} are a statistically significant indicator in predicting y_i.

The analysis for Granger Causality is as follows. 1) I calculate the metrics at the proposal level and word-level to implement a weekly-level granular analysis from 2015 to 2018. I use the U.S. weekly mean CO2 data from Global Monitoring Laboratory. Next, Granger causality between word-metric and CO2 molfrac (ppm) is computed by fitting a linear regression model with m-lag and n-lag. CO2 emissions Granger-cause the coverage of *R&D* if the analysis finds that β is different from zero with statistical significance.

$$\text{Word-level metric}_t = \sum_{i=1}^{m} \alpha_i (\text{Word-level metric}_{t-1}) + \sum_{j=1}^{n} \beta_j (\text{CO2 molfrac (ppm)}_{t-1})$$

Table B.2 indicates the analysis results. I found 1-lag mean Carbon Dioxide values Granger-cause coverage of *R&D* agenda at both proposal-level and word-level metrics. Remarkably, the coefficients for Greenhouse gas emission are positive, which indicates that the increase in the emission is followed by an increase in R&D coverage. In the 2-lag analysis, the p-value at the proposal-level is less than 0.1 but it is not statistically significant at the word-level. To rule out the likelihood of reverse causality, I compute Grange causality in the opposite direction, and the analysis does not return statistically meaningful results.

	Proposal-level		Word-level	
	1-Lag	**2-Lag**	**1-Lag**	**2-Lag**
R&D Coverage$_{t-1}$	0.800**	0.202***	0.003	0.002
	(0.353)	(0.073)	(0.072)	(0.072)
R&D Coverage$_{t-2}$		−0.073		0.051
		(0.072)		(0.072)
Mean carbon dioxide$_{t-1}$	0.115***	0.044	0.00002***	−0.001
	(0.036)	(0.038)	(0.00000)	(0.001)
Mean carbon dioxide$_{t-2}$		0.070*		0.001
		(0.038)		(0.001)
N	190	190	190	190
Adj. R-squared	0.1079	0.117	0.127	0.122

***p < .01; **p < .05; *p < .1

TABLE B.2. Granger Causality between CO2 molfrac (ppm) and the coverage of *R&D*

The Granger Causality Tests in the reverse direction are presented below (Table B.3). These findings imply that corporate actors discuss the topic of R&D when air pollution increases.

	Proposal-level		Word-level	
	1-Lag	**2-Lag**	**1-Lag**	**2-Lag**
Mean carbon dioxide$_{t-1}$	1.000***	0.728***	0.002***	0.001***
	(0.0002)	(0.069)	(0.084)	(0.092)
Mean carbon dioxide$_{t-2}$		0.272***		0.089***
		(0.069)		(0.071)
R&D Coverage$_{t-1}$	1.767	1.601	1.0004	1.001
	(3.804)	(3.677)	(0.509)	(0.704)
R&D Coverage$_{t-2}$		1.602		0.007
		(3.681)		(2.001)
N	190	190	190	190
Adj. R-squared	0.1079	0.117	0.137	0.112

$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

TABLE B.3. Granger Causality between CO2 molfrac (ppm) and the coverage of *R&D* (Reverse Direction)

B.2. Paragraph Vector Framework

Figure B.2 is from Le and Mikolov (2014). In this schematic framework, context of three words ("the," "cat," and "sat") is used to predict the fourth word. The input words are mapped to columns of the matrix W to predict the output word. And the additional paragraph token that is mapped to a vector via matrix D. Next the concatenation or average of this vector with a context of three words is used to predict the fourth word. The paragraph vector represents the missing information from the current context and can act as a memory of the topic of the paragraph. The paragraph vectors and word vectors are trained using stochastic gradient descent and the gradient is obtained via backpropagation. At every step of stochastic gradient descent, one can sample a fixed-length context from a random paragraph, compute the error gradient from the network in Figure 9 and use the gradient to update the parameters in the model.

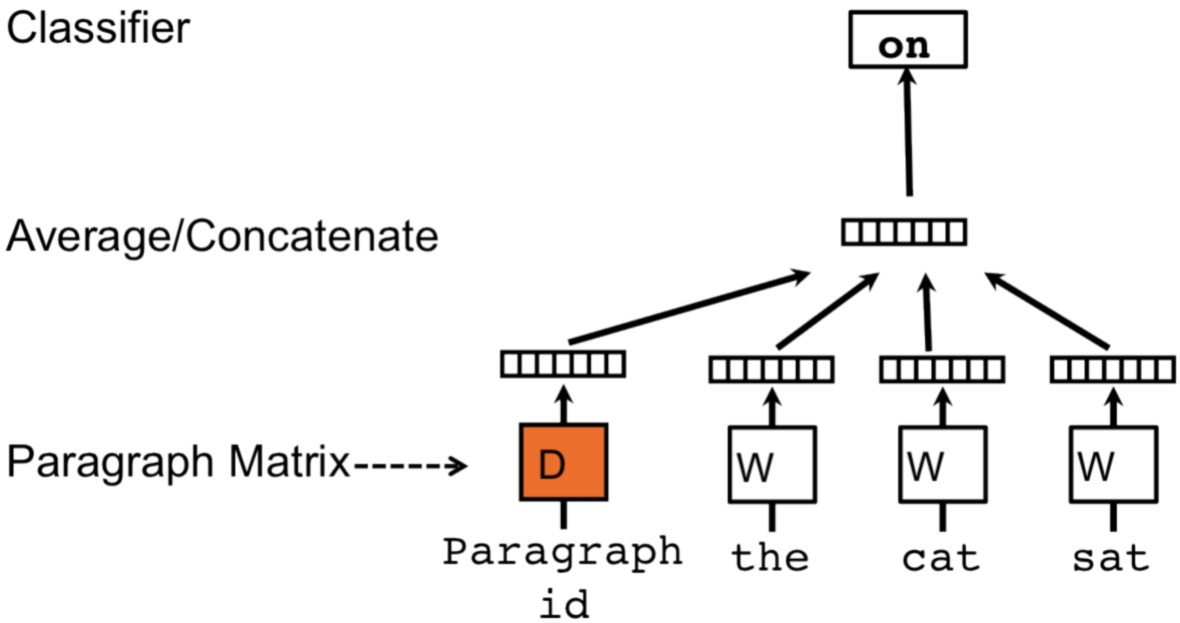


FIGURE B.2. A Framework of Learning Paragraph Vector in Le and Mikolov (2014)

B.3. Cosine Similarity

Figure B.3 below illustrates the intuition of using cosine similarity as a measure of issue slant toward RD and technology. The positive score implies that a comment is slanted in the RD and technology direction while the negative score indicates that a comment is slanted in the reduction direction. A score of 0 means a comment is balanced between the two.

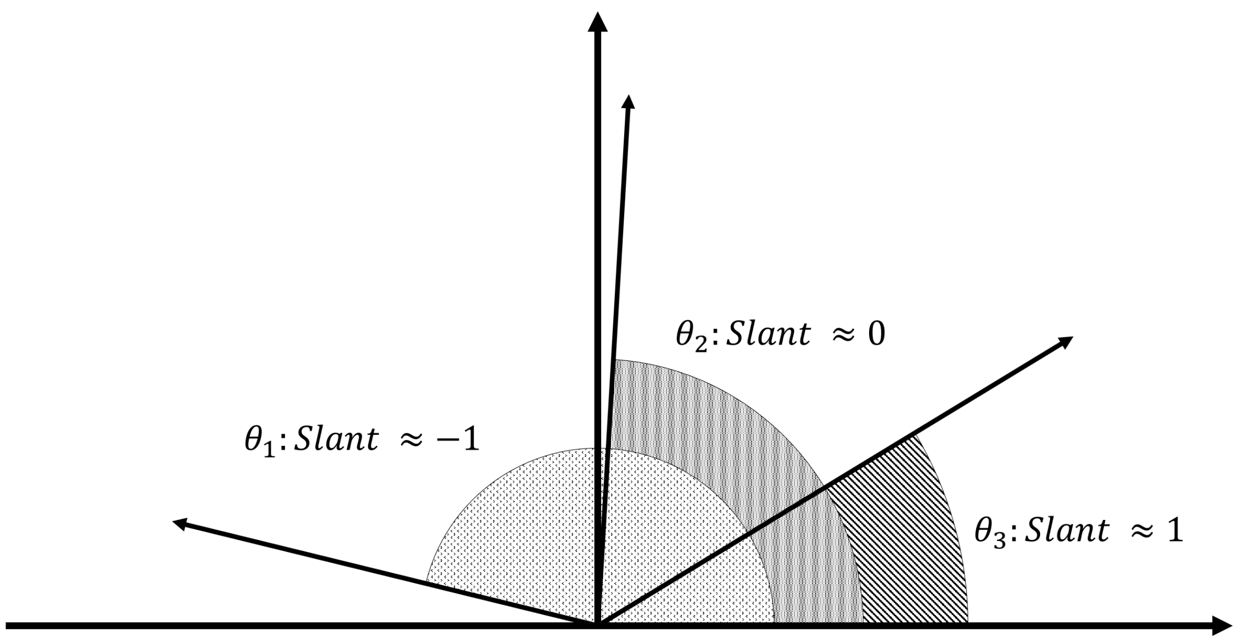


FIGURE B.3. Cosine Similarity as a measure of issue slant towards R&D and Technology

For simplicity, I denote $\overrightarrow{R\&D} - \overrightarrow{Reduction}$ as $\overrightarrow{x}$ and paragraphs vectors for each document as $\overrightarrow{y}$. This metric for non-zero vectors, $\overrightarrow{x}$ and $\overrightarrow{y}$, is defined as

$$similarity(\overrightarrow{x}, \overrightarrow{y}) = cos(\theta) = \frac{\overrightarrow{x} * \overrightarrow{y}}{||\overrightarrow{x}||||\overrightarrow{y}||} = \frac{\sum x_i y_i}{\sqrt{\sum x_i{}^2}\sqrt{\sum y_i{}^2}} \tag{A1}$$

where θ denotes the angle between vectors and $||*||$ indicates the 2-norm. The similarity score ranges from -1 to 1 and, a score close to 1 indicates that a document tends to emphasize

R&D compared to reductions. If the score is negative, this implies that the issue slant in the emission direction.

B.4. Word Embeddings Evaluations

There is no real evaluation possible except to test how Paragraph Vectors performs on the downstream task as it is an unsupervised model. Alternatively, I implemented human-based validation by examining the closest words to the words; research, development, technology and reductions to make sure that training a neural network makes sense substantively. The table below indicates the words used for constructing a global issue slant vector. For robustness check, I adjust the number of words 3,5 and 9 for each topic and the results are still consistent.

TABLE B.4. Vocabularies Associated with Each Topic

	Related Vocabularies
R& D Technology	research, development, technology, alternatives, strategies, solutions, innovations, drilling, sector, extraction
Emission Reductions	limit, address, mitigate, curb, prevent, minimize reducing, decrease, restrict, cut

Table B.5. Regression Models Examining the Issue Slant toward R&D versus Greenhouse Gas Reductions (Vocabulary size 6)

Sample	Partnerships + Environmental Groups	Partnerships + Business Association + Single Firms	Whole Sample
	(1)	(2)	(3)
Partnership	0.128**	−0.062***	0.016**
	(0.049)	0.016)	(0.008)
Single firms		0.008	0.028***
		(0.013)	(0.005)
Business associations			0.042***
			(0.009)
Others			0.031***
			(0.007)
Staff Size	−0.00004	−0.00000	0.00000
	(0.00003)	(0.00000)	(0.00000)
Constant	0.034	0.163***	0.044***
	(0.031)	(0.016)	(0.004)
Year FE	✓	✓	✓
Commenter FE	✓	✓	

*p < .1; **p < .05; ***p < .01. In the first column, the reference category is *environmental groups*, while in the second column, it is *business associations*. For the third column, the reference category remains *environmental groups*.

TABLE B.6. Regression Models Examining the Issue Slant toward R&D versus Greenhouse Gas Reductions (Vocabulary size 9)

Sample	Partnerships + Environmental Groups	Partnerships + Business Association + Single Firms	Whole Sample
	(1)	(2)	(3)
Partnership	0.052	−0.115***	0.012
	(0.054)	(0.014)	(0.008)
Single firms		−0.017	0.040***
		(0.011)	(0.007)
Business associations			0.049***
			(0.011)
Others			0.031***
			(0.009)
Staff Size	−0.00004	−0.00000***	0.00000
	(0.00002)	(0.00000)	(0.00000)
Constant	0.062*	0.197***	0.044***
	(0.032)	(0.014)	(0.005)
Year FE	✓	✓	✓
Commenter FE	✓	✓	

*p < .1; **p < .05; ***p < .01. In the first column, the reference category is *environmental groups*, while in the second column, it is *business associations*. For the third column, the reference category remains *environmental groups*. Standard errors are clustered by notice and comment periods in parentheses.

B.5. Analysis Using Simple Count-based Metrics for Robustness Check

In this subsection, I examine additional measures that capture the dynamics of a compromise between firms and environmental groups as robustness checks. I construct two measures: 1) a count-based metric that indicates how many times research, development, and technology vocabularies are mentioned, and 2) a binary variable indicating whether a comment has mentioned the vocabularies of research, development, or technologies at least three times.

TABLE B.7. Negative Binomial Model Estimating the Frequency of R&D Vocabularies- Using Count-based Metric

Sample	Partnerships + Environmental Groups	Partnerships + Business Association + Single Firms	Whole Sample
	(1)	(2)	(3)
Partnership	0.639***	1.125***	1.085***
	(0.187)	(0.358)	(0.355)
Single Firms		−0.082	−0.063
		(0.164)	(0.162)
Environmental Groups			0.309
			(0.233)
Others			−0.065
			(0.208)
Staff Size	0.0001	0.00001***	0.00000***
	(0.0003)	(0.00000)	(0.00000)
Constant	1.014***	0.120	0.225*
	(0.095)	(0.098)	(0.119)
Year FE	✓	✓	✓

*p < .1; **p < .05; ***p < .01

*p < .1; **p < .05; ***p < .01. In the first column, the reference category is *environmental groups*, while in the second column, it is *business associations*. For the third column, the reference category remains *business associations*. Standard errors are clustered by notice and comment periods in parentheses.

TABLE B.8. Probit Regression estimating the likelihood that the comment mentions R&D and technological topics

Sample	Partnerships + Environmental Groups	Partnerships + Business Association + Single Firms	Whole Sample
	(1)	(2)	(3)
Partnership	0.533***	1.125***	0.821***
	(0.098)	(0.358)	(0.253)
Single Firms		−0.082	−0.087
		(0.164)	(0.122)
Environmental Groups			0.140
			(0.164)
Others			−0.006
			(0.110)
Staff Size	0.0004	0.00001***	0.00000
	(0.0002)	(0.00000)	(0.00000)
Constant	−0.455***	0.120	−0.913***
	(0.051)	(0.098)	(0.085)
Year FE	Yes	Yes	Yes

*p < .1; **p < .05; ***p < .01. In the first column, the reference category is *environmental groups*, while in the second column, it is *business associations*. For the third column, the reference category remains *business associations*. Standard errors are clustered by notice and comment periods in parentheses.

C. Increased Quantity of Expertise for a Higher-quality Policy Implementation

	busi_coalition (N=250)	env (N=154)	others (N=66)	Partnership (N=71)	single_business (N=362)	Overall (N=903)
Count of Technical Information						
Mean (SD)	165 (173)	134 (157)	161 (167)	299 (248)	124 (136)	154 (170)
Median [Min, Max]	93.5 [0, 1130]	68.5 [0, 720]	78.5 [7.00, 737]	248 [30.0, 1460]	62.0 [0, 573]	79.0 [0, 1460]
Staff Numbers of Group						
Mean (SD)	1280 (10600)	121 (243)	8300 (48700)	430 (314)	13900 (55300)	6570 (38300)
Median [Min, Max]	50.0 [1.00, 160000]	40.0 [1.00, 1510]	245 [4.00, 390000]	600 [10.0, 800]	700 [1.00, 543000]	141 [1.00, 543000]

Note: Entries indicate the statistics for each comment type.

FIGURE C.1. Summary Statistics - A Measure Constructed by Named Entity Recognition

C.1. Validating named entity recognition techniques

As a classification task, the performance of the named entity recognition technique is usually measured in terms of classification metrics (over all the tokens) like precision, recall, F-score, accuracy, etc. I used spaCy v3.0, which introduces a transformer-based pipeline, and the testing accuracy of it is found to be around .89. After training a NER model, I conducted a human validation exercise using public comments submitted by the Clean Air Council. Figure C.2 shows pages 6-7 of the comments written by the Clean Air Council. The specific comment ID that includes the comment is EPA-HQ-OAR-2010-0505-4375, and the link is as follows: 'https://www.regulations.gov/comment/EPA-HQ-OAR-2010-0505-4375.'

The colored boxes in Figure C.3 show the words detected by the entity recognition technique, which includes organizations, geopolitical locations, or cardinality. One interesting pattern to note is that the number of false positives is relatively lower compared to that of false negatives. It is very obvious that all the colored boxes refer to very specific words, such as 'Ohio,' 'Evaporation Pond Facilities,' or '23 tons.' However, the texts in public comments contain more scientific and technical information than what is detected by entity recognition techniques. For instance, named entity recognition cannot detect *volatile organic compounds (VOCs)* or *Hydraulics,* which convey scientific implications. Moreover, given that the primary challenge is identifying 'false negatives,' while sorting out 'false positives' is much more straightforward, the expertise captured by entity recognition techniques may provide a more conservative estimate than the actual quantity of specific evidence and scientific knowledge present in the public comments as a result.

52,750 – Provisions for encouraging innovative technology

The Council supports a specific program to test new and innovative technologies in the oil and gas industry. However, this program must be crafted carefully. Innovation and experimentation comes with potential increases in pollution and environmental harm as technologies are explored and tested. This program should establish several criteria by which operators are allowed to test new technologies.

For established companies with capital to expend on innovation and significant histories as operators in the oil and gas industry, EPA should consider the following: (1) the resources of the

Clean Air Council EPA-HQ-OAR-2010-0505 7

company to determine whether it has the capacity to test new technology; (2) the number and seriousness of the violations of environmental laws and regulations in relation to the size of the company to determine its relative level of environmental safety; (3) whether the company is a member of EPA's Natural Gas STAR program; (4) whether the new and innovative technology proposed would significantly reduce potential cumulative pollution relative to the older technology or method the new innovative technology would displace; and (5) whether the technology would significantly reduce the risk of a larger, immediate pollution event relative to the older technology or method the new innovative technology would displace. EPA should also consider collateral or bonding requirements to test these technologies as well, with amounts depending on the scale of the test and the potential environmental impacts.

52,756 – Evaporation Pond Facilities

FIGURE C.2. Public Comment by Clean Air Council

For established companies with capital to expend on innovation and significant histories as operators in the oil and gas industry, EPA ORG should consider the following: (1 CARDINAL) the resources of the company to determine whether it has the capacity to test new technology; (2 CARDINAL) the number and seriousness of the violations of environmental laws and regulations in relation to the size of the company to determine its relative level of environmental safety; (3 CARDINAL) whether the company is a member of EPA ORG 's Natural Gas STAR ORG program; (4 CARDINAL) whether the new and innovative technology proposed would significantly reduce potential cumulative pollution relative to the older technology or method the new innovative technology would displace; and (5 CARDINAL) whether the technology would significantly reduce the risk of a larger, immediate pollution event relative to the older technology or method the new innovative technology would displace. EPA ORG should also consider collateral or bonding requirements to test these technologies as well, with amounts depending on the scale of the test and the potential environmental impacts. 52,756 CARDINAL — Evaporation Pond Facilities ORG The Council does not have firm data on the emissions from evaporation pond facilities. However, anecdotal evidence from Council ORG members living near multiple evaporation ponds suggests these are a significant source of VOCs and HAPs. Members have stated that rashes, allergies, and difficulty breathing coincide with evaporation pit installation, venting, and burn-off operations. EPA ORG should evaluate short-term impacts from these ponds because they may pose a significant threat to human health. Cumulative impacts are largely unknown as well. Should EPA ORG find what the Council ORG and its members suspect—that these ponds represent the types of dangers the NSPS/NESHAP rule revision are meant to guard against—then EPA ORG should consider strict rules on technologies to be used for evaporation ponds, should there even be technology in existence that would make evaporation ponds safer. 52,758 CARDINAL — Reduced Emission Completions The Council supports all efforts by the EPA ORG to require companies to perform emission completions (RECs) of wells. It is estimated that tens of thousands CARDINAL of wells will be drilled in shale deposits in Pennsylvania GPE 's portion of the Marcellus Shale LOC alone. The underlying and nearby Utica Shale ORG suggests that hydraulically fractured well drilling could expand in Ohio GPE and other parts of Pennsylvania GPE . EPA ORG 's estimate of 23 tons QUANTITY of VOCs emitted from uncontrolled hydraulically fractured wells multiplied by the thousands CARDINAL of potential wells that could be drilled and hydraulically fractured indicates the gravity of the pending air quality problem. EPA ORG estimates that many RECs would recover enough methane that would otherwise be lost to provide a net return despite the increased costs. But EPA ORG should also consider health impacts and the broader cost to society that not requiring RECs would cause. Increased VOCs would have a significant, measurable set of impacts on public health, from increased hospital visits to premature death. This is an important part of any economic impact assessment; the EPA ORG should not merely stop with an estimate on the change in profitability that any company doing the REC ORG would undergo.

FIGURE C.3. Applying Entity Recognition Technique to Public Comment by Clean Air Council

TABLE C.1. Negative binomial model estimating the quality of information (Year Fixed Effect)

Sample	Partnerships + Environmental Groups		Partnerships + Business Association + Single Firm		Whole Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Partnership	0.705***	2.670***	0.760***	1.878***	0.864***	0.730***
	(0.149)	(0.366)	(0.141)	(0.582)	(0.148)	(0.135)
Single Firm			−0.355***	−0.033	−0.200*	−0.334***
			(0.090)	(0.538)	(0.117)	(0.089)
Business associations					0.134	
					(0.123)	
Others					0.167	0.033
					(0.167)	(0.152)
Environmental Groups						−0.134
						(0.123)
Issue Slant	1.301	−0.123	−0.429	0.069	−0.237	−0.237
(*R&D* and Technology)	(0.791)	(1.426)	(0.530)	(1.479)	(0.462)	(0.462)
Year FE	✓	✓	✓	✓	✓	✓
Commenter FE		✓		✓		
Env_staff	0.0002	−0.0001	0.00000*	0.00001	0.00000	0.00000
	(0.0002)	(0.001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
Constant	4.801***	6.621***	4.801***	3.194***	4.668***	4.802***
	(0.204)	(0.445)	(0.145)	(0.606)	(0.146)	(0.132)

*p < .1; **p < .05; ***p < .01

TABLE C.2. Negative binomial model estimating the quality of information (Administration Fixed Effect)

Sample	Partnerships + Environmental Groups		Partnerships + Business Association + Single Firm		Whole Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Partnership	0.712***	0.284	0.644***	1.378**	0.786***	0.626***
	(0.144)	(0.970)	(0.127)	(0.564)	(0.137)	(0.123)
Single Firm			−0.344***	0.179	−0.158	−0.317***
			(0.090)	(0.554)	(0.112)	(0.089)
Business associations					0.160	
					(0.120)	
Others					0.161	0.001
					(0.161)	(0.147)
Environmental Groups						−0.160
						(0.120)
Administration FE	✓	✓	✓	✓	✓	✓
Commenter FE		✓		✓		
Staff Size	0.0002	−0.0002	0.00000**	0.00001	0.00000	0.00000
	(0.0002)	(0.001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
Issue Slant toward R&D	1.225	0.090	−0.276	−0.338	−0.039	−0.039
	(0.814)	(1.333)	(0.547)	(1.508)	(0.468)	(0.468)
Constant	4.766***	6.168***	5.205***	3.735***	4.988***	5.147***
	(0.130)	(0.326)	(0.089)	(0.612)	(0.101)	(0.086)

*p < .1; **p < .05; ***p < .01. In the first and second columns, the reference category is *environmental groups*, while in the third and fourth columns, it is *business associations*. For the fifth column, the reference category remains *environmental groups* while the reference category becomes *business associations*. Standard errors are clustered by notice and comment periods in parentheses.

Table C.3. Quasi-Poisson Model Estimating the Quality of Information (Year Fixed Effect)

Sample	Partnerships + Environmental Groups		Partnerships + Business Association + Single Firm		Whole Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Partnership	0.700***	2.757***	0.658***	1.337***	0.822***	0.646***
	(0.148)	(0.383)	(0.127)	(0.475)	(0.139)	(0.086)
Single Firm			−0.321***	−0.110	−0.134	−0.311***
			(0.086)	(0.405)	(0.114)	(0.086)
Business associations					0.176	
					(0.120)	
Others					0.158	−0.018
					(0.163)	(0.146)
Environmental Groups						−0.176
						(0.120)
Year FE	✓	✓	✓	✓	✓	✓
Commenter FE		✓		✓		
Staff Size	0.0002	−0.0002	0.00000	0.00001	0.00000	0.00000
	(0.0002)	(0.001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
Issue Slant toward R&D	1.241*	0.109	−0.342	0.054	−0.032	−0.032
	(0.744)	(1.543)	(0.518)	(1.352)	(0.452)	(0.452)
Constant	4.873***	6.565***	4.900***	3.736***	4.701***	4.878***
	(0.181)	(0.476)	(0.142)	(0.488)	(0.145)	(0.130)

*p < .1; **p < .05; ***p < .01. In the first and second columns, the reference category is *environmental groups*, while in the third and fourth columns, it is *business associations*. For the fifth column, the reference category remains *environmental groups* while the reference category becomes *business associations*. Standard errors are clustered by notice and comment periods in parentheses.

TABLE C.4. Quasi-Poisson Model Estimating the Quality of Information (Administration Fixed Effect)

Sample	Partnerships + Environmental Groups		Partnerships + Business Association + Single Firm		Whole Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Partnership	0.720***	0.183	0.645***	1.179***	0.802***	0.632***
	(0.141)	(0.932)	(0.124)	(0.408)	(0.136)	(0.123)
Single Firm			−0.294***	−0.008	−0.116	−0.286***
			(0.090)	(0.400)	(0.112)	(0.089)
Business associations					0.170	
					(0.119)	
Others					0.172	0.002
					(0.160)	(0.146)
Environmental Groups						−0.170
						(0.119)
Administration FE	0.026	−0.143	−0.241**	−0.066	−0.138	−0.138
	(0.140)	(0.208)	(0.107)	(0.174)	(0.086)	(0.086)
Staff Size	0.0002	−0.0003	0.00000	0.00001	0.00000	0.00000
	(0.0002)	(0.001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
Issue Slant toward R&D	1.303*	−0.142	−0.376	−0.473	−0.005	−0.005
	(0.716)	(1.311)	(0.517)	(1.275)	(0.449)	(0.449)
Constant	4.788***	6.174***	5.179***	3.947***	4.956***	5.126***
	(0.124)	(0.313)	(0.089)	(0.451)	(0.101)	(0.086)

*p < .1; **p < .05; ***p < .01

*p < .1; **p < .05; ***p < .01. In the first and second columns, the reference category is *environmental groups*, while in the third and fourth columns, it is *business associations*. For the fifth column, the reference category remains *environmental groups* while the reference category becomes *business associations*. Standard errors are clustered by notice and comment periods in parentheses.

D. Political Influence of Strategic Partnerships on Regulatory Outcome

D.1. Proof: Comparing Multinomial Distributions and Divergence Scores

The procedure for calculating JS divergence scores is as follows. I collapse the comments of each group and adopt the notion of the bag-of-words model which represents a document as a set of the count. Equivalently, a cluster of proposals is represented as a multinomial probability distribution over words. Figure 2 shows the Kernel density of each distribution. P denotes the proposals submitted by business interests while O indicates final amendments. Again, each of them is represented as multinomial distribution. P is a multinomial distribution of (p_1, p_2, ..., p_n) and O is a multinomial distribution of (o_1, o_2, o_3, ..., o_n). n is the total number of tokens in the entire text. p_i is defined as the $\frac{count(word_i, proposals)}{length(proposals)}$ and o_i is defined as the $\frac{count(word_i, finalizedrule)}{length(finalizedrule)}$.
Let x be a discrete variable, and the set of probability distribution x is parameterized by a vector p where p(x=k)= p_k.

(A2)
$$p(x|p) = \prod_{k=1}^{K} p_k$$

where $\delta(x = k)$ is an indicator function. Therefore, the joint probability of N IID samples X can be expressed as

(A3)
$$p(X|P) = \prod_{k=1} K\, p_k$$

(A4)
$$N_k = \sum_{k} \delta(x_k = k)$$

A conjugate prior for p is the Dirichlet distribution:

(A5) $$p(p|\alpha) \sim D(\alpha_1, ..., \alpha_k) = \frac{\Gamma(\sum_k \alpha_k)}{\pi_k \Gamma(\alpha_k)} \prod_k {}_k^{k-1}$$

(A6) $$\sum_k p_k = 1$$

In above equations, α_k is the hyper-parameter, a virtual count for value k. Large α is equivalent to prior knowledge about the distribution. The Dirichlet distribution has the properties of

(A7) $$p(p_1|\alpha) \sim D(\alpha_1, \alpha_2 + ..., +\alpha_k)$$

(A8) $$E[p_1] = \frac{\alpha_1}{\sum_k \alpha_k}$$

(A9) $$E[l\,og\,p_1] = \psi(\alpha_1) - \psi(\sum_k \alpha_k)$$

, where $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$. The maximum of its density is at $p_k = (\alpha_k - 1)/((\sum_k \alpha_k) - k)$. Given Dirichlet prior, the joint distribution of a set of samples X and p is

(A10) $$p(X, p|\alpha) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)}{}_k\, p_k^{N_k+\alpha_k-1}$$

and the posterior is reduced to $p(p|X, \alpha) \sim D(N_k + \alpha_k)$

Therefore, the probability that data all come from one multinomial distribution can be indicated as;

(A11) $$p(X|\alpha) = \int_p p(X, p|\alpha)$$

$$\text{(A12)} \qquad = \frac{\Gamma_k \alpha_k \prod_k \Gamma(N_k + \alpha_k)}{\prod_k \Gamma(\alpha_k)\Gamma(\sum_k N_k + \alpha_k)} \int_p D(p; N_k + \alpha_k)$$

$$\text{(A13)} \qquad = \frac{\Gamma(\sum_k \alpha_k)}{\Gamma(N + \sum_k \alpha_k)} \prod_k \frac{\Gamma(N_k + \alpha_k)}{\Gamma(\alpha_k)}$$

D.2. Proof: Jenson-Shannon divergence as a test statistic

Our primary concern is the probability that finalized rule (= Y) and proposals (=X) are from the same probabilistic distribution. Based on information-theoretic quantity of mutual information, I connect distribution similarity to hypothesis testing. This is a problem related to homogeneity and I examine if two samples X and Y are from the same multinomial distribution or different distribution. Therefore, I am interested in

$$\text{(A14)} \qquad P(same|X, Y) = \frac{P(X|Y|same)\, p(same)}{p(X, Y|same)\, p(same) + \, p(X, Y|different)\, p(different)}$$

$$\text{(A15)} \qquad = \frac{1}{1 + \frac{P(X,Y|different))}{P(X,Y|same)} \frac{P(different)}{P(same)}}$$

The quantity of $\frac{p(X,Y|different)}{p(X,Y|same)}$ is the ratio in favor of difference as shown below by Wolpert (1995)

$$\text{(A16)} \qquad \frac{p(X|\alpha)\, p(Y|\alpha)}{p(X, Y|\alpha)} = \frac{\Gamma(\sum_k \alpha_k \Gamma(M + N + \sum_k k)}{\Gamma(M + \sum_k \alpha_k)\Gamma(N + \sum_k k)} \prod_k \frac{\Gamma(M_k + \alpha_k)\Gamma(N_k + \alpha_k)}{\Gamma(\alpha_k)\Gamma(M_k + N_k + \alpha_k)}$$

By entropy approximation, the logarithm of this ratio is equal to

(A17) $$log\frac{p(X|\alpha)\,p(Y|\alpha)}{p(X,Y|\alpha)} \approx -MH(\frac{M_k}{M}) - NH(\frac{N_k}{N}) + (M+N)H\frac{M_k+N_k}{M+N})$$

(A18) $$= MD(\frac{M_k}{M}||\frac{M_k+N_k}{M+N}) + ND(\frac{N_k}{N}||\frac{M_k+N_k}{M+N})$$

(A19) $$\text{where } D(p||q) = \sum_k p_k log\frac{p_k}{q_k}$$

The equation 19 is equal to the average divergence to the mean, which is known to be Jensen-Shannon divergence.

Note: Entropy Approximation

(A20) $$\frac{\gamma(K)}{\Gamma(N+K)} \approx \frac{\Gamma(K)}{\Gamma(N+1)N^{K-1}} \approx \frac{1}{\Gamma(N+1)}$$

(A21) $$log\,p(X|\alpha=1) \approx -Nl\,ogN + N + \sum_k (N_k l\,ogN_k - N_k)$$

(A22) $$= \sum_k N_k l\,og\frac{N_k}{N}$$

(A23) $$= -NH(N_k/N)$$

D.3. What does the Jensen-Shannon (JS) divergence score capture?

Figure D.1 provides intuition behind the use of the JS divergence score with actual examples. The box at the top shows one paragraph from the policy comments submitted by the Environmental Defense Fund (EDF) and the Sierra Club, both of which are classified as environmental groups with business partners. The box at the bottom represents one paragraph from the EPA's response to the policy announcement. In their comments, EDF and the Sierra Club point out that EPA misunderstood the emission control measure of lower explosive limits (LEL). In response to these concerns, EPA acknowledged learning from the comments and recognized that emission control based on LEL detectors would not work effectively.

The purpose of the analysis using the divergence score in this section is to quantify the influence of policy comments on the policy amendments made by EPA.

B. Proposed Method for Delineating End of Initial Flowback Period

" ……..Environmental Commenters support EPA's proposal to require monitoring of methane concentrations during the initial flowback period, and to require emission controls once methane concentrations approach the lower explosive limit (LEL). We note that recent studies demonstrate that the LEL is a conservative indicator of the feasibility of separating and controlling emissions during a completion. Therefore, we respectfully request that EPA explicitly defines the precise methane concentrations that……………"

Comments by Environmental Defense Fund and Sierra Club

Commenters responded that the EPA apparently had misunderstood earlier discussions regarding use of the LEL detector. They asserted that the detector is used for safety reasons and that although the LEL detector indicates that there may be potential flammability, it does not necessarily indicate that sufficient gas is present for the separator to function. Commenters also asserted that monitoring the gas concentration does not reflect other conditions such as sand and water content and well characteristics that have a bearing on the point where the separator will operate. We also learned that some operators begin to direct the flowback to the separator immediately upon initial flowback, even though it may not maintain a gaseous phase and one or more liquid phases in the separator.

Summary of Significant Changes and Comments by EPA

FIGURE D.1. Example: Comments Submitted by the Environmental Defense Fund and EPA's response to the comment

D.4. Analysis Procedure- JS divergence scores

The procedure for calculating the statistical distance is as follows. After adopting the bag of words model, each comment and a set of policy amendments can be represented as normalized count, or equivalently a multinomial probability distribution over words KL divergence is formally defined as:

$$D(P||O) = \sum_i P_i log(\frac{P_i}{O_i})$$

where P and O are multinomial distributions of the words in comments and policy amendments and the expectation value of the log difference between the two probability distribution is computed with weights of P_i.[1]

However, the KL divergence is an asymmetric measure, resulting in the score for A to B being different from the score for B to A. The algebraic reason is that $D(P||O) - D(O||P)$ is equal to $\sum_i^n ln(\frac{P_i}{O_i})(P_i+O_i)$ and there is no reason for this to be 0. Therefore, the test statistics used in this analysis are inferred from Jensen-Shannon (JS) divergence. JS divergence is defined as

$$J(P||O)) = \frac{1}{2}(D(P||M) + D(O||M))$$

where $M = \frac{1}{2}(P + O)$ is a mean distribution. Figure D.2 represents the density plot of JS divergence scores to a finalized policy.

[1] P_i is equal to $\frac{count(word_i, doc_1)}{length(doc_1)}$. Similarly, O_i is $\frac{count(word_i, doc_2)}{length(doc_2)}$.

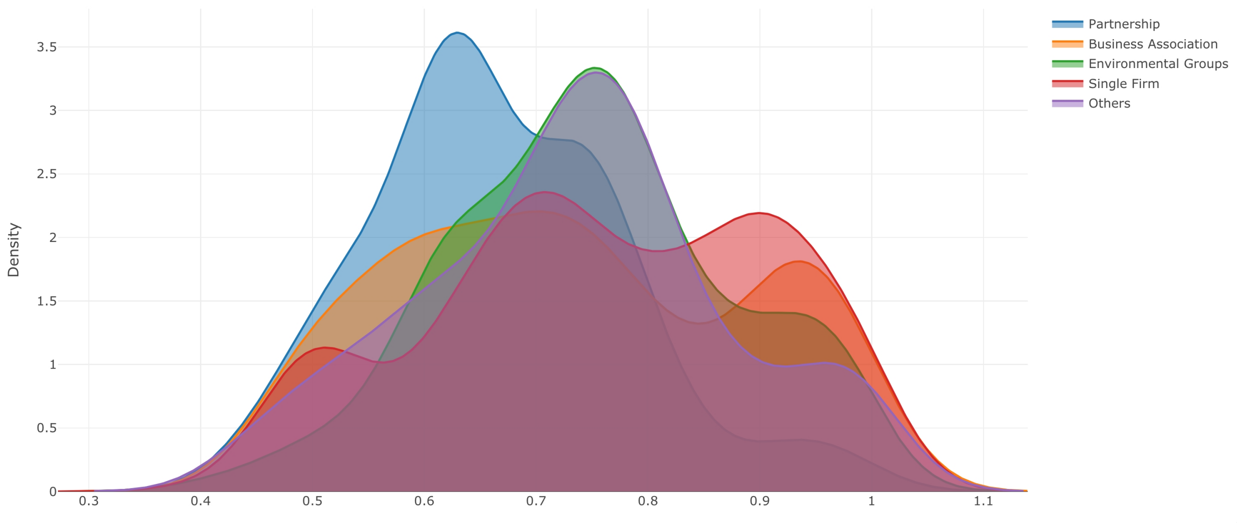


FIGURE D.2. Density Plot: JS Divergence Scores to Policy Amendments

TABLE D.1. Regression Models Estimating JS Divergence Scores and Citation by EPA Officials

	JS Divergence (OLS)		*Citation Counts* (Probit)	
	(1)	(2)	(3)	(4)
Partnership	−0.031***	−0.015*	0.512	1.737**
	(0.003)	(0.009)	(0.679)	(0.720)
Environmental groups	−0.015**		−1.225***	
	(0.007)		(0.396)	
Others	−0.015***	−0.0001	−5.014	−3.789
	(0.006)	(0.001)	(228.534)	(228.534)
Single firm	−0.001	0.014***	−0.445	0.779*
	(0.003)	(0.004)	(0.346)	(0.426)
Business associations		0.015**		1.225***
		(0.007)		(0.396)
Administration FE	−0.001	−0.001	0.869**	0.869**
	(0.001)	(0.001)	(0.380)	(0.380)
Staff Size	−0.00000	−0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00001)	(0.00001)
JS Divergence to a proposed policy	0.917***	0.917***		
	(0.014)	(0.014)		
Parity IGscore	0.002	0.002	0.207	0.207
	(0.004)	(0.004)	(0.497)	(0.497)
Constant	0.083***	0.068***	−1.381***	−2.606***
	(0.012)	(0.010)	(0.387)	(0.494)
Observations	903	903	181	181
Note:		*p<0.1; **p<0.05; ***p<0.01		

*p<0.1; **p<0.05; ***p<0.01.

TABLE D.2. Does the signaling mechanism hold without any consideration of partnerships?

	JS Divergence	
	(1)	(2)
Absolute difference between IGscores	-0.015^{***}	-0.015^{***}
	(0.001)	(0.001)
Administration	-0.001^{**}	-0.0001
	(0.001)	(0.002)
Staff Size	-0.000	-0.00000
	(0.00000)	(0.00000)
JS Divergence to a proposed policy	0.917^{***}	0.916^{***}
	(0.016)	(0.016)
Year FE	✓	
Constant	0.078^{***}	0.080^{***}
	(0.012)	(0.011)

*p < .1; **p < .05; ***p < .01

D.5. The use of rule posted on March 12, 2018 to capture citations patterns

This section further provides visualizations to justify the use of the final rule posted on March 12, 2018, for quantifying the political influence of public comments. Figure D.3 is a part of the final amendment posted by EPA on Sep 23, 2013, after the notice and the comment period [2]. However, while the response by EPA attempts to address the issues raised by commenters, it does not provide specific comment IDs or the names of commenters that were considered by EPA officials to make amendments. Similar to the rule on Sep 23, 2013, other rules were written without references to specific comments, so there have been empirical challenges in clearly identifying whose comments have been considered by EPA officials and what comments have been reflected in the final rule.

Comment: One commenter requested that the final rule specify the date upon which the determination of the potential VOC emission rate should occur for the purpose of determining whether the storage vessel is an affected facility. According to the commenter, since the EPA has stipulated controls to not be cost effective for storage vessels emitting less than 6 tpy of VOC, and emission rates for storage vessels in the oil production segment tend to decrease as production declines, the commenter believes the determination should be made near to the date upon which controls would be required in order to minimize the potential to install controls on storage vessels for which production decline has rendered controls no longer cost effective. The commenter stated that the proposed revisions would require a determination by October 15, 2013, of whether individual Group 1 storage vessels are affected facilities, and thus October 15, 2013, would be an appropriate date upon which determination of the potential VOC emission rate should be based. According to the commenter, this would remain consistent with the requirement for determining the potential VOC emission rate for Group 2 storage vessels by April 15, 2014 or 30 days after startup, whichever comes later.

The commenter appears to suggest that, like Group 2, Group 1 storage vessel affected facilities located in the natural gas processing and natural gas transmission and storage segments should also be required to determine potential VOC emissions as the trigger for installing control instead of tracking events but to do so by April 15, 2015 (instead of April 15, 2014, proposed for Group 2). According to the commenter, control of the relatively low number of Group 1 storage vessel affected facilities in these segments could likely be accommodated by this date.

Another commenter pointed out that the proposed reconsideration rule does not establish the date for a Group 1 storage vessel to determine its potential emissions. The commenter also recommended that notifications are only required for tanks that exceed the 6 tpy threshold on October 15, 2013. Although the publication date of the proposed reconsideration rule was April 12, 2013, the commenter contends that the EPA is not required to, nor should it, establish the emissions determination date for the source category of Group 1 storage vessels on that date. First, given the rapidly declining emissions at storage vessels following initial fracturing, the commenter believes that the expected emissions reduction to be gained from Group 1 storage vessels is likely to be limited. The commenter also states that the proposal date of April 12, 2013, has passed and operators may not be able to accurately back-calculate emissions from that date. Moreover, the commenter contends that emissions from many of these storage vessels will be below the 6 tpy affected source threshold as of October 2013. Given EPA's proposed approach, where storage vessel affected facilities whose emissions drop below 6 tpy remain subject to the standard, the commenter believes that many Group 1 storage vessels will be unnecessarily captured in the source category and required to indefinitely track “events” and perhaps install control devices even if their emissions never again exceed 6 tpy.

FIGURE D.3. A final rule posted on September 23, 2013

[2]The full text is available at https://www.regulations.gov/document/EPA-HQ-OAR-2010-0505-4635

One commenter described configurations at well sites that can lead to an automatic emergency well shut-in and where the rule, if applied as suggested in the preamble, could have unintended consequences. (2) Where well sites have a compressor that collects flash gas from a low pressure separator or a vapor recovery unit that collects flash gas from storage vessels, there are certain safety measures put in place in the event these compressors unexpectedly go offline. Depending on the remoteness of the well site, one safety measure available is to automatically shut in the well to prevent the release of gas from pressure relief valves. In these, and other similar emergency shut-in situations, the equipment is not depressurized so the well can be brought back into production as soon as possible. However, by requiring completion of the delayed repair during such shut-in events, equipment at this well site that have components placed on delayed repair would have to be depressurized and blown down, resulting in emissions that would not have occurred except for the delayed repair requirement and could be higher than the emissions from continuing to delay repair.

A. A Paragraph of a Final Rule

Footnotes

(1) See 40 CFR 60.5397a(h)(2) for delay of repair requirements.

(2) See Docket ID No. EPA-HQ-OAR-2010-0505-12446.

(3) See Docket ID No. EPA-HQ-OAR-2010-0505-12447.

(4) See Docket ID Nos. EPA-HQ-OAR-2010-0505-12421, EPA-HQ-OAR-2010-0505-12424, EPA-HQ-OAR-2010-0505-12430, EPA-HQ-OAR-2010-0505-12436, EPA-HQ-OAR-2010-0505-12446, EPA-HQ-OAR-2010-0505-12447, and EPA-HQ-OAR-2010-0505-12454.

(5) See Docket ID Nos. EPA-HQ-OAR-2010-0505-12430, EPA-HQ-OAR-2010-0505-12436, EPA-HQ-OAR-2010-0505-12446, EPA-HQ-OAR-2010-0505-12447, and EPA-HQ-OAR-2010-0505-12454.

(6) See Docket ID No. EPA-HQ-OAR-2010-0505-12447.

B. Footnote

FIGURE D.3. A Final rule posted on March 12, 2018

However, as shown in the case of the rule posted on March 12, 2018 (Figure D.3),[3] EPA officials explicitly cite comment IDs or commenters in the footnote, in contrast to other final rules focused on Emission Standards. Based on this evidence, we can infer whose comments were considered by EPA officials and influenced the updated policy. Therefore, the analysis, which relies on the citation patterns by bureaucrats, is mainly focused on the rule posted on March 12, 2018. While the analysis is limited to this particular rule, further exploration into responses by EPA officials after the commenting period can further enrich the discourse.

[3]The full text is available at https://www.regulations.gov/document/EPA-HQ-OAR-2010-0505-12503