

The Limits of Interest: Capture, Financialization, or Contestation in the Politics of Rule-Making for Derivatives

Online Supporting Material

Konrad Posch
University of California, Berkeley
Berkeley, California, USA
konrad.posch@berkeley.edu

Thomas Nath
Brown University
Providence, Rhode Island, USA
thomas_nath@alumni.brown.edu

J. Nicholas Ziegler
Brown University
Providence, Rhode Island, USA
j_ziegler@brown.edu

Correspondence: konrad.posch@berkeley.edu

Acknowledgements: The authors wish to thank Luke Elder, Katherine Sicher-Camp, and Vrinda Agarwal for coding assistance. Thanks as well to our chairs, discussants, panels, and audiences who provided helpful feedback on earlier presentations of this work at the Midwest Political Science Association (MPSA) 2016 Annual Meeting and the American Political Science Association (APSA) 2018 Annual Meeting.

The manuscript of the paper may be found through the Watson Institute Working Paper Series:

[URL of Manuscript]

Supporting Information A: Heterogeneity in Commenting between Rules

Above, we saw that our 18-part typology represented the variation in commenter type across the corpus. By exploring the heterogeneity of commenter frequency and mixture across rules, this section shows that the implementation process can be reasonably construed as one extended conversation because it shows that a variety of commenters commented on a variety of rules, both major and minor.

The majority of the 114 proposed rules have less than ten⁰ comments while only 12 rules had more than ten⁰ comments and only one had more than ten⁰⁰.¹ The rule with the most comments concerned Position Limits for Derivatives (76 FR 4752) and received 3335 comments. From this pattern, it is clear that not all rules receive equal commenter attention. It is also important to note that this long tail pattern of commenting is present in both the complete dataset of comments (n=36,066) and the corpus of comments with organization values (n=8262), suggesting that the corpus is representative of the complete dataset.

While the heterogeneity in volume of comments per rule points out variations in overall participation between rules, the types of commenters that participated in the top ten rules demonstrates that the most involved deliberations had considerable variety in the balance of commenters who chose to weigh in. In particular, there is a substantial, if varied presence of non-industry commenters.²

¹ The 114 total Proposed Rules in the comment database were identified by the CFTC in the data provided using a Federal Register citation of "Volume FR BeginningPage", i.e. 76 FR 4752

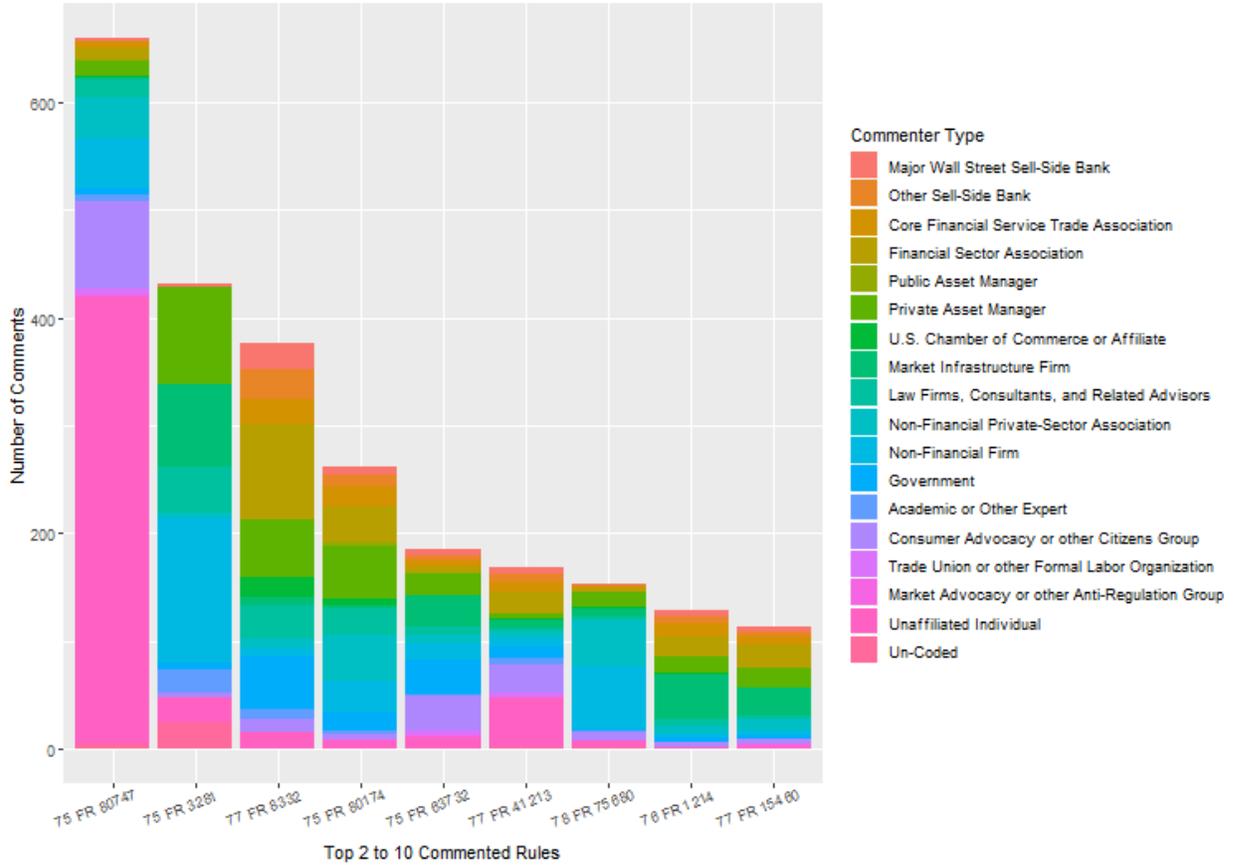
² To review our classificatory approach, when we speak of non-industry commenters, we are beginning with the 12th group on our 18 part typology (Government) and including all groups after that ("Academic or Other Expert" to "Other Individual"). Uncoded (18) is excluded from analysis.

This bifurcation of the typology into industry and non-industry commenters provides a quick short-hand for demonstrating that other groups aside industry insiders were involved in the comment process. The main manuscript discusses topic proportions by individual commenter type, providing much more detail on how distinct each commenter type truly is. In this section, we show that heterogeneity in commenter mixture on the ten most frequently commented rules shows industry and non-industry involvement.

We focus on the ten rules that attracted the highest number of comments for several reasons. First, these were the rules that treated especially consequential aspects of the derivatives trade. Second, the rule-making actions that attracted the fewest public comments, often in the single or low double digits, were often minor follow-ons or other technical actions that buttressed other more central rules. Third, these ten rules accounted for 5812 comments of 8351, or 69.6 percent of the comments whose authors supplied an organizational affiliation. And finally, in substantive terms, the CFTC's ten most frequently commented rules included those that provoked broader contestation, sometimes going well beyond the notice-and-comment process to vigorous debate in the general press and, in two instances, eventual adjudication.

Interestingly, these ten rules demonstrate that even in the largest rulemaking debates, there was a substantial, if varied, involvement of non-industry commenters. It is clear from Figure 1 that the overwhelming majority of the overall "Other Individual" commenters were concentrated in the two highest volume rules (76 FR 4752 and 75 FR 80747). However, we still see a substantial presence of non-industry commenters such as "Government" and "Consumer Advocacy or other Citizens Groups" in other rules within the top ten.

Figure 1: Most Frequently Commented Second to Tenth CFTC Proposed Rules (n=2477)



Note: The top rule (76 FR 4752) has been omitted from this display as it overwhelms the vertical scale. It has a similar cohort of commenters to 75 FR 80747.

If we examine the detailed breakdown of organizations commenting on each of the top ten rules, a number of noteworthy differences become apparent. Perhaps most intriguing, the “Major Wall Street Sell-Side Banks” are only conspicuously present in the comments for the Volcker Rule (77 FR 8332). Explicitly backed by President Obama, the Volcker Rule was the single reform plank that became most widely discussed in the broader press. Since it directly blocked banks with any sort of government guarantee from the lucrative practice of proprietary trading, a high degree of engagement by the banks was to be expected. For the other nine

Table 1: Most Frequently Commented Ten CFTC Proposed Rule: Name and Number of Comments per Rule

FR Citation	Rule Name	Comments
76 FR 4752	Position Limits for Derivatives	3335
75 FR 80747	End-User Exception to Mandatory Clearing of Swaps	659
75 FR 3281	Regulation of Off-Exchange Retail Foreign Exchange Transactions and Intermediaries	432
77 FR 8332	Prohibitions and Restrictions on Proprietary Trading and Certain Interests in, and Relationships With, Hedge Funds and Covered Funds; (Commonly known as the Volcker Rule)	377
75 FR 80174	Further Definition of “Swap Dealer,” “Security-Based Swap Dealer,” “Major Swap Participant,” “Major Security-Based Swap Participant” and “Eligible Contract Participant.”	261
75 FR 63732	Requirements for Derivatives Clearing Organizations, Designated Contract Markets, and Swap Execution Facilities Regarding the Mitigation of Conflicts of Interest	185
77 FR 41213	Cross-Border Application of Certain Swaps Provisions of the Commodity Exchange Act	168
78 FR 75680	Position Limits for Derivatives	153
76 FR 1214	Core Principles and Other Requirements for Swap Execution Facilities	129
77 FR 15460	Procedures To Establish Appropriate Minimum Block Sizes for Large Notional Off-Facility Swaps and Block Trades	113

Note: The count of comments includes comments with multiple signatories

frequently commented rules, it is nonetheless noteworthy, that the Major Wall Street Banks only submitted comments numbering in the single digits.

For the four most frequently commented rules, the level of non-industry engagement does not follow any clear sense of relative salience. For example, the definition of a “Swap Dealer” (75 FR 80174) was the kind of technical matter of critical importance to market participants, but very little direct interest to citizens groups or the broader public. As expected, the number of industry comments outweighed non-industry comments by 217 (86.1%) to 35 (13.9%). By comparison,

the Volcker Rule (77 FR 8332) itself prompted a less pronounced preponderance: 279 comments (76.2%) from industry versus 87 (23.8%) from non-industry commenters.

For the other two most frequently commented rules, the level of non-industry comments reached very surprising levels. The rule that set “Requirements for Derivatives Clearing Organizations” (75 FR 63732) established who could take ownership of the new clearinghouses through which the derivatives business would be conducted. Though clearly a technical issue, the proposed rule prompted industry to submit 99 comments (54.7%), while non-industry sources accounted for 82 comments (45.3%). The proposed rule entitled “Cross-Border Application of Certain Swaps” (77 FR 41213) regulated international transactions. In this case, industry sources submitted a total of 68 comments (42.5%), which was outweighed by non-industry comments numbering 92 (57.5%).

These observations provide additional reason to reassess the view that industry exercises disproportionate influence in the notice-and-comment process. To be sure, on some of the individual rules examined here, particularly the Volcker Rule and the Definition of a Swap Dealer, the number of business comments outweigh non-industry comments by a wide margin. On other rules of a quite technical nature, however, non-industry commenters submitted almost as many or even more comments than did industry actors themselves. Also noteworthy is the very small number of comments submitted directly by the Wall Street banks. Contrary to concepts of industry capture, there is little apparent relationship between the stakes for a given firm in the markets being regulated by an agency and the propensity of that firm to use the notice-and-comment process to shape the agency’s decisions. This conceptual significance of this heterogeneity in commenting between rules also demonstrates that our typology captures significant variation in the corpus.

Supporting Information B: LDA Topic Modelling Implementation

The Data (n=8262) and Our 18-part Commenter Typology

Our dataset includes a complete record of the 36,066 public comments submitted to the CFTC concerning Dodd-Frank over a 42-month period (January 14, 2010 to July 16, 2014).³ The data was exported from the CFTC's internal database and provided to the authors by email following a cold call to the CFTC public relations department. The CFTC's database was maintained as part of its internal compliance with the APA and includes all rule-making notices that appear in the Federal Register. Owing to the salience of the Dodd-Frank Act, the CFTC created a special internal tag for all comments submitted in response to Dodd-Frank rule-making notices. We are thus confident that our database includes all comments that the CFTC considered relevant to the Dodd-Frank reform received between January 14th, 2010 and July 16th, 2014.

In addition to the full comment text, the CFTC provided us with metadata about each comment, most significantly the organization which the commenter indicated they were affiliated with.⁴ Unfortunately, organization was an optional field in the comment form from the CFTC website leading to incomplete metadata. Unless otherwise specified, all references to “the corpus” from here onward means the group of 8262 comments⁵ with an interpretable

³ Data archived in Dryad: <https://doi.org/10.6078/D1610G>

⁴ The other metadata included: a CFTC assigned unique identifier, the date the comment was submitted, the federal register citation which the comment was submitted in response to, the first and last name of the commenter. The CFTC also sent the text extracted from documents which some commenters attached to their comment.

⁵ The CFTC provided 37,232 total entries. 36,066 of these were public comments submitted through the website or on paper while 1,165 were records of ex parte meetings. To guarantee a common data generating process to ensure comparability, we include only the comments and dropped the ex parte meetings. Of the 36,066 public comments, 9,189 had a non-blank Organization variable value. Of those 9,189, 927 had an organization value equivalent to “n/a,” “none,” or gibberish (eg. a punctuation mark) which we excluded from analysis since this was indistinguishable between attempts to speed through the webform and legitimate indications of non-affiliation. Values with enough information (e.g. “none – individual”) were retained and coded in the category “Unaffiliated Individual.” This resulted in a corpus of 8,262 public comments with an interpretable Organization value. The exact

organization value that we used to classify each comment into our typology of commenter types. This corpus forms the basis of all substantive analysis.⁶

Why LDA?

The goal of topic modelling is to uncover the latent themes, or “topics,” that underlie the actual words in a corpus of documents. Our approach uses a Latent Dirichlet Allocation (LDA) topic model to generate topics from the corpus of CFTC public comments which are then analyzed for patterns using our typology of commenter categories. As developed by David Blei, Andrew Ng, and Michael I Jordan (2003), LDA produces topics as probability distributions over all words in the corpus; that is, each word is assigned a probability of representing a certain topic. The algorithm begins with a set number of topics composed of randomly assigned words and then iteratively analyzes each word throughout the corpus. For each word, LDA analyzes (1) the proportion of assignments to a given topic across all instances of that word in the corpus, and (2) the proportion of assignments to a given topic across all words in that given document. The algorithm then produces a final output of these two distributions, topic-word proportions (1) and document-topic proportions (2), once it reaches a preset number of iterations.

We chose LDA as opposed to other topic models because it treats each document as a mixture of topics. This approach thus allows us to view the extent to which each commenter focused on a given topic by generating the proportion of topics represented within a given document and, subsequently, the entire corpus. Other scholars have used LDA to investigate

algorithm of the exclusion rules may be seen in the online supplemental material where we import and prepare the data for analysis.

⁶ Our Implementation of LDA (below) discusses how we deal with comments which list multiple organization values; the “multi-signatory issue.” This solution means the number of *cases* in the figures is 8451 although this is the same as our corpus of 8262 *comments*.

debates in a breadth of policy domains from the macroscopic – the Federal Reserve’s construction of economic policy (Fligstein et al., 2017) – to the minute –electronic monitoring of the trucking industry (Levy and Franklin, 2014).

LDA topic modelling requires two external parameters: the number of topics and the number of iterations. We chose 14 topics based on a qualitative assessment of substantive word group coherence in test runs of five to 100 topics. Similarly, we chose 12,000 iterations based on our determination of when the algorithm stabilized to the point of infinitesimally small changes to the probability distributions. While these parameters can be selected quantitatively, studies find that automated methods are not preferable to human judgment in interpreting topic coherence (Chang et al., 2009; Ramage et al., 2009). In fact, an ongoing project involving both machine learning and qualitative experts cautions that an overreliance on automated methods may at best do no better than traditional qualitative judgement and at worst lose important analytic meaning (Jefferson et al., 2019). Given the esoteric and jargon-rich nature of our corpus, our own expert judgment of topic coherence made human intuition the preferred tool for these decisions.

Our Implementation of LDA

This section covers the import and filtering of the data, the matching of the keyed table of coded organization values, the preparation of the corpus, generation of the topic model, and the creation of interpretable result tables and figures used in the findings section of main manuscript.⁷ All work was done in R (2018).

⁷ A precise explication is available in Online Supporting Information D, an HTML file produced from the Rmd syntax file that shows how we processed our data from raw CFTC export to final results

The CFTC provided us a dataset with 37,232 total entries. We dropped 1,165 of these as they were ex parte meetings and not public comments. We dropped another 26,877 cases as they did not include any organization values. Of the remaining 9,189 comments, we dropped 927 cases with an uninterpretable Organization value such as “n/a,” “none,” or gibberish (e.g. a single punctuation mark). The result is the corpus of 8262 comments on which we built our analysis.

From the corpus of 8262 comments, we exported the 4,154 unique Organization values and coded them to the 18-part typology. The final classifications from that process (described above in the data section) were used as a table of organization values keyed to our typology in order to computationally code the 8262 comments by matching the comment’s organization value to the keyed table. All 8262 comments were thus consistently computationally matched to their classification based on their organization value without consideration of any other meta-data and the final result was automatically integrated into the database.

With a corpus now coded to the typology, we created a unified text variable which was then prepared for LDA to process. The corpus of text to analyze was created by gap filling seven comments from CommentText into ExtractedText so that all 8262 comments had text in the corpus.⁸ The full UnifiedText corpus was then stemmed,⁹ removed whitespace, removed

⁸ The CFTC provided two columns from their database which included the text of comments: CommentText and ExtractedText. CommentText was created from the values which commenters pasted into the webform while ExtractedText was created from the combination of the CommentText and text extracted from attached documents (e.g. PDF, docx, etc) which the commenter may have uploaded. If the commenter did not include any attachments, ExtractedText was intended to be a copy of CommentText. We discovered that 7 values were blank in the database, and followed the same procedures as the CFTC described in order to gapfill those comments.

⁹ Stemming involves removing all suffixes and modifiers from core words to leave just the stem. For example, “require” and “requirement” both become “require.” This is standard practice in topic modelling to get just the substance of words in a corpus without additional variation which would impede comparison.

punctuation, cast to lower case, and had standard English and corpus specific stopwords removed.¹⁰

With the corpus so prepared, we used an R wrapper (Mimno, 2013) for the Java program MALLET (McCallum, 2002) to train the topic model. The cleaned text was fed into Mallet, the number of topics was set to 14 based on substantive analysis of preliminary runs,¹¹ and the model was trained for 12,000 iterations with optimization of hyperparameters every 20 iterations. The result was a set of 14 topics with every comment coded with the proportion of that comment which concerned that topic, the *topic proportion*.

Before analysis, we needed to expand the corpus to handle comments with multiple signatories. A number of comments had multiple organization types listed in the Organization value (161 with two, 27 with three, 1 with four). For these multi-signatory comments, the comment was duplicated in the corpus to count as one contribution for each type of commenter. This means that the total number of comments in the topic proportions below (n=8451) will exceed the total number of raw comments (n=8262).

We chose this approach because we are interested in who is saying what and how that differs between different types of commenters. This means it is less important to maintain exact numerical equivalency since we are comparing proportions. We chose one contribution to each category (rather than fractional based on some arbitrary formula) for the same reason that every signatory to the Declaration of Independence was equally guilty of treason even though only some (e.g. Thomas Jefferson) did the majority of writing. If you signed a letter, you are jointly,

¹⁰ The full list of stopwords, both generic and corpus specific, may be found in the online appendix.

¹¹ See discussion of the literature and reasoning about our selection of 14 topics in the “Why LDA?” section above.

severally, wholly, and equally liable for its contents even if you were potentially a lesser part of the writing process. With the multi-signatory letters dealt with, we then averaged the topic proportions for each comment into the typical proportions for each commenter type in our 18-part typology. These results are discussed main manuscript.

Works Cited

- Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003) ‘Latent Dirichlet Allocation’, *Journal of Machine Learning Research*, **3**, 993–1022.
- Carpenter, D. (2004) ‘Protection without Capture: Product Approval by a Politically Responsive, Learning Regulator’, *American Political Science Review*, **98**, 613–631.
- Chang, J., Gerrish, S., Wang, C., Boyd-graber, J. L. and Blei, D. M. (2009) ‘Reading Tea Leaves: How Humans Interpret Topic Models’. In Bengio, Y., Schuurmans, D., Lafferty, J. D., Williams, C. K. I., and Culotta, A. (eds) *Advances in Neural Information Processing Systems 22*, Curran Associates, Inc., pp. 288–296.
- Fligstein, N., Brundage, J. S. and Schultz, M. (2017) ‘Seeing Like the Fed: Culture, Cognition, and Framing in the Failure to Anticipate the Financial Crisis of 2008’, *American Sociological Review*, **82**, 879–909.
- Jefferson, A., Mills, S., Xi, X. and Hunt, M. (2019) ‘Comparing Manual Qualitative Coding to Natural Language Processing: A Case Study of Financial Decision-Making’, Association for Public Policy Analysis & Management (APPAM).
- Levy, K. E. C. and Franklin, M. (2014) ‘Driving Regulation: Using Topic Models to Examine Political Contention in the U.S. Trucking Industry’, *Social Science Computer Review*, **32**, 182–194.
- McCallum, A. K. (2002) *MALLET: A Machine Learning for Language Toolkit*.
- McCubbins, M. D., Noll, R. G. and Weingast, B. R. (1987) ‘Administrative Procedures as Instruments of Political Control’, *Journal of Law, Economics & Organization*, 243–278.
- Mimno, D. (2013) *Mallet: A Wrapper around the Java Machine Learning Tool MALLET*.
- R Foundation for Statistical Computing (2018) *R: A Language and Environment for Statistical Computing.*, Vienna, Austria.
- Ramage, D., Rosen, E., Chuang, J., Manning, C. D. and McFarland, D. A. (2009) ‘Topic Modeling for the Social Sciences’, *Workshop on Applications for Topic Models*, 4.

Woolley, J. T. and Ziegler, J. N. (2012) 'The Two-Tiered Politics of Financial Reform in the United States'. In Mayntz, R. (ed) *Crisis and Control: Institutional Change in Financial Market Regulation*, Frankfurt, New York, Campus Verlag, pp. 29–65.

Ziegler, J. N. and Woolley, J. T. (2016) 'After Dodd-Frank: Ideas and the Post-Enactment Politics of Financial Reform in the United States', *Politics & Society*, **44**, 249–280.

Supporting Information C: R Coding Syntax and Output

All processing of data was carried out using an R markdown file. That Rmd file is available as online supplemental material for readers interested in precisely how the description of the process in the manuscript section “Method: LDA Topic Modelling” was executed.

The Full HTML rendering of the Rmd Syntax may be found at:

<http://konradposch.com/research/online-appendices/posch-nath-ziegler-2019-the-limits-of-interest>

While the dataset was produced from internal CFTC servers (because the CFTC did not participate in Regulations.gov in 2014), they explicitly only provided publicly available data. Thus, the full dataset of 37,703 comments is available from Dryad (<https://doi.org/10.6078/D1610G>) upon publication along with the codebook provided by the CFTC and amended with our additional classification variables.

It should be noted that the CFTC does have additional meta-data which they declined to provide (i.e. IP addresses for commenters and other personally identifiable information) which they noted could be obtained through a Freedom of Information Act (FOIA) request. We did not pursue this avenue, but future researchers interested in, for example, the geographic distribution of commenters could request such data.

Supporting Information D: Non-Industry Commenters Include Well-Informed Pro-Reform Advocacy Groups

The argument for business's disproportionate influence in the notice-and-comment process does not rest exclusively on numerical preponderance. Several authors argue that large firms in concentrated industries are better positioned to submit high-quality substantive comments of the kind that commands attention from regulators. According to Matthew McCubbins, Roger Noll and Barry Weingast (1987), the APA was designed so that well-resourced interest groups could help members of Congress (as principals) in the task of monitoring bureaucratic agencies (as agents) as they implement legislative intent. According to Daniel Carpenter (2004), incumbent firms also have an incentive to supply reliable information to regulators in order to deepen their credibility over time.

We do not question whether incumbent firms can and often do submit comments with extensive information. Our analysis suggests, however, that non-industry commenters also seek to build their reputation for providing well-informed perspectives on the agency's plans for implementing the Dodd-Frank reforms. In the case of Dodd-Frank, the frequency of comments submitted by the most engaged organizations shows that a small number of advocacy groups also assembled considerable expertise on the rules proposed by the CFTC. As illustrated in Table 2, two of the most active commenters were small pro-reform organizations established explicitly to lobby for regulatory reform after the financial crisis. The first of these groups, Americans for Financial Reform (AFR), was itself a coalition of consumer organizations, labor unions, and other public interest groups. It was particularly active during Congressional hearings on the formulation of the Dodd-Frank Act (Woolley and Ziegler, 2012), and remained active in the post-enactment phase. In terms of frequency, it submitted over ten⁰ comments to the CFTC on

Dodd-Frank rules – more than any other organization including the key financial-services lobbying associations, major asset managers, or even the U.S. Senate. Nor was AFR alone. A

Table 2: Top Commenting Organizations (≥ 99 percentile, n = 5247)

Commenting Organization	Percentile
Americans for Financial Reform (AFR)	99.8 % (≥ 58 comments)
Better Markets, Inc.	
CME Group, Inc.	
International Swaps and Derivatives Association, Inc. (ISDA)	
Managed Funds Association (MFA)	99.5 % (≥ 26 comments)
Asset Management Group (AMG) of the Securities Industry and Financial Markets Association (SIFMA)	
BlackRock, Inc.	
Coalition of Physical Energy Companies (COPE)	
Hunton & Williams LLP	
MarkitSERV	
National Futures Associaton (NFA)	
The Depository Trust & Clearing Corporation	
The Financial Services Roundtable	
United States Senate	
Committee on Capital Markets Regulation	
Commodity Markets Council	
Congress of the United States	
Edison Electric Institute	
Encana Marketing (USA) Inc.	
Freddie Mac	
Futures Industry Association	
International Energy Credit Association	
Markit	
Minneapolis Grain Exchange, Inc.	
Sutherland Asbill & Brennan LLP, on behalf of the Federal Home Loan Banks	
The Association of Institutional INVESTORS	
The Bank of Tokyo-Mitsubishi UFJ, Ltd. Mizuho Corporate Bank, Ltd.	
Sumitomo Mitsui Banking Corporation	
Wholesale Markets Brokers' Association, Americas	

Unaffiliated Individuals(17) and Un-Coded(18) types excluded from this table in order to highlight organizational commenters. Organizations are alphabetical within percentiles.

small policy think tank, Better Markets, Inc., was the third most frequent commenter. While we have not performed systematic content analysis of the comments, the sheer number of comments submitted by these two organizations indicates that they were accumulating substantial expertise on the issues considered by the CFTC. Both organizations relied on former Congressional staffers for much of their commenting activity. Their comments were among those regularly referenced in the CFTC's final rules. And other studies have shown that the views conveyed by AFT and Better Markets, Inc. accorded with CFTC outcomes on several of the most contentious rules issued by the agency (Ziegler and Woolley, 2016).

Works Cited

- Blei, D. M., Ng, A. Y. and Jordan, M. I. (2003) 'Latent Dirichlet Allocation', *Journal of Machine Learning Research*, **3**, 993–1022.
- Carpenter, D. (2004) 'Protection without Capture: Product Approval by a Politically Responsive, Learning Regulator', *American Political Science Review*, **98**, 613–631.
- Chang, J., Gerrish, S., Wang, C., Boyd-graber, J. L. and Blei, D. M. (2009) 'Reading Tea Leaves: How Humans Interpret Topic Models'. In Bengio, Y., Schuurmans, D., Lafferty, J. D., Williams, C. K. I., and Culotta, A. (eds) *Advances in Neural Information Processing Systems 22*, Curran Associates, Inc., pp. 288–296.
- Fligstein, N., Brundage, J. S. and Schultz, M. (2017) 'Seeing Like the Fed: Culture, Cognition, and Framing in the Failure to Anticipate the Financial Crisis of 2008', *American Sociological Review*, **82**, 879–909.
- Jefferson, A., Mills, S., Xi, X. and Hunt, M. (2019) 'Comparing Manual Qualitative Coding to Natural Language Processing: A Case Study of Financial Decision-Making', Association for Public Policy Analysis & Management (APPAM).
- Levy, K. E. C. and Franklin, M. (2014) 'Driving Regulation: Using Topic Models to Examine Political Contention in the U.S. Trucking Industry', *Social Science Computer Review*, **32**, 182–194.
- McCallum, A. K. (2002) *MALLET: A Machine Learning for Language Toolkit*.
- McCubbins, M. D., Noll, R. G. and Weingast, B. R. (1987) 'Administrative Procedures as Instruments of Political Control', *Journal of Law, Economics & Organization*, 243–278.

- Mimno, D. (2013) *Mallet: A Wrapper around the Java Machine Learning Tool MALLETT*.
- R Foundation for Statistical Computing (2018) *R: A Language and Environment for Statistical Computing.*, Vienna, Austria.
- Ramage, D., Rosen, E., Chuang, J., Manning, C. D. and McFarland, D. A. (2009) ‘Topic Modeling for the Social Sciences’, *Workshop on Applications for Topic Models*, 4.
- Woolley, J. T. and Ziegler, J. N. (2012) ‘The Two-Tiered Politics of Financial Reform in the United States’. In Mayntz, R. (ed) *Crisis and Control: Institutional Change in Financial Market Regulation*, Frankfurt, New York, Campus Verlag, pp. 29–65.
- Ziegler, J. N. and Woolley, J. T. (2016) ‘After Dodd-Frank: Ideas and the Post-Enactment Politics of Financial Reform in the United States’, *Politics & Society*, **44**, 249–280.