Hall of Mirrors: Corporate Philanthropy and Strategic Advocacy

Latest version https://bradhackinen.ca/files/BradHackinen_JMP.pdf

Marianne Bertrand, Matilde Bombardini,
Raymond Fisman, Brad Hackinen, and Francesco Trebbi*

November 2018

Abstract

Politicians and regulators rely on feedback from the public when setting policies. For-profit corporations and non-profit entities are active in this process and are arguably expected to provide independent viewpoints. Policymakers (and the public at large), however, may be unaware of the financial ties between some firms and non-profits – ties that are legal and tax-exempt, but difficult to trace. We identify these ties using IRS forms submitted by the charitable arms of large U.S. corporations, which list all grants awarded to non-profits. We document three patterns in a comprehensive sample of public commentary made by firms and non-profits within U.S. federal rulemaking between 2003 and 2015. First, we show that, shortly after a firm donates to a non-profit, the grantee is more likely to comment on rules for which the firm has also provided a comment. Second, when a firm comments on a rule, the comments by non-profits that recently received grants from the firm's foundation are systematically closer in content similarity to the firm's own comments than to those submitted by other non-profits commenting on that rule. Third, when a firm comments on a new rule, the discussion of the final rule is more similar to the firm's comments when the firm's recent grantees also comment on that rule. These patterns, taken together, suggest that corporations strategically deploy charitable grants to induce non-profit grantees to make comments that favor their benefactors, and that this translates into regulatory discussion that is closer to the firm's own comments.

* Bertrand: University of Chicago Booth School of Business and NBER; Bombardini: University of British Columbia, CIFAR, and NBER; Fisman: Boston University and NBER; Hackinen: PhD Candidate, University of British Columbia; Trebbi: University of British Columbia, CIFAR, and NBER. We would like to thank Kevin Milligan and seminar participants at Harvard Kennedy School and UBC for discussion. Bombardini and Trebbi acknowledge financial support from CIFAR and SSHRC. Pietro Montanarella and Jack Vincent provided excellent research assistance.

1 Introduction

Economists and political scientists have long studied – both theoretically and empirically – the role interest groups play in the formation of laws and regulations. In the U.S., as in many democracies, there are well-established channels through which interest groups can try to influence the laws and rules that may impact their communities, their businesses, or society at large. Through means such as lobbying, grassroots campaigns, testimonies, or public advocacy, interested parties inform politicians and bureaucrats of the costs and benefits of government action.

While interest groups may have expertise on topics of direct relevance to them, they may also be tempted to present information that is tainted by their self-interest. This logic is at the core of the literature on informational lobbying (Grossman and Helpman, 2001).[1] For example, oil company representatives may have expertise in drilling, but also a strong incentive to minimize, say, the predicted environmental costs of Arctic oil exploration. Government officials must thus weigh both the quality of information and its impartiality, based in part on its source. As such, lawmakers and rulemakers may view information provided by for-profit corporations as less credible if that information is not corroborated by other groups with non-aligned (i.e., neutral or opposing) interests.

Non-profit organizations often fall into the role of interests that are non-aligned with business. Some non-profits – such as research groups, universities, and think tanks – are providers of nonpartisan, technical expertise and are commonly expected to offer more neutral input into the lawmaking and rulemaking process, with a focus on cost-benefit analysis and broader societal interests. Other non-profits – such as human services organizations, environmental protection groups, social welfare organizations, and advocacy groups – may have opposing interests to business, to the extent that laws or regulations that benefit their members (or those on whose behalf they advocate) adversely constrain business profits. Non-profit organizations are therefore expected to play an important balancing role in the informational lobbying process.

This role may be subverted, however, by the financial links between corporations and non-profits: in exchange for donations, a non-profit may (consciously or otherwise) take a perspective that is favorable to its benefactor's bottom line. If politicians and bureaucrats are more likely to implement a proposal when it is supported by interest-group diverse coalitions (as suggested theoretically in the strategic advocacy literature, e.g., Krishna and Morgan (2001), Dewatripont and Tirole (1999), Dewatripont and Tirole (2005), and empirically in Lorenz, 2017) and if such ties are undisclosed, such "coalition building via corporate giving" may distort the outcome of the

[1] By informational lobbying we refer to the broad literature on information transmission which encompasses cheap talk and costly signalling models in the context of lobbying, for example Potters and Van Winden (1992), Austen-Smith (1993), Austen-Smith (1995) and Lohmann (1995).

political process away from the public good and towards private interests.[2]

The goal of this paper is to provide systematic evidence establishing this to be an empirically relevant phenomenon. The context of U.S. Federal Regulation, with its far-reaching economic implications and its carefully documented record of communications between organizations and government agencies, offers an ideal setting to establish such evidence.

There exists anecdotal evidence that these concerns are well-founded. Across a range of issues and regulatory agencies, researchers and journalists have documented cases of companies using charitable contributions to co-opt ostensibly neutral and even non-aligned non-profits. Notably, Peng (2016) describes the efforts of telecommunications firms to win merger approvals in front of the Federal Communication Commission (FCC), in part by assembling diverse and vocal coalitions of supporters. Peng quotes Crawford (2013) on the Comcast-NBCU merger, in which *"[t]he company encouraged letters to the FCC from more than one thousand non-profits...including community centers, rehabilitation centers, civil rights groups, community colleges, sports programs, [and] senior citizen groups."* For the AT&T/T-Mobile merger, Peng similarly documents letters of support addressed to the FCC from non-profits that, at first glance, would appear to have little interest or expertise in telecommunications policy, including a homeless shelter in Louisiana, a special needs employment agency in Michigan, and the Gay & Lesbian Alliance Against Defamation (GLAAD). The non-profits were all AT&T Foundation grantees (in the case of the homeless shelter, the donation had come in just five months before the merger was announced). In no case did the non-profit disclose its AT&T funding in its letter to the FCC. In at least one case, the comments did not appear to represent the views of the non-profit membership. According to Peng, "*GLAAD's president and six board members resigned when its merger endorsement made headlines and revealed that the organization had received AT&T funds.*"

Journalists and medical experts have documented similar persuasion-via-donation in public health debates. Jacobson (2005), for example, describes a ("no-strings attached") $1 million donation from Coca-Cola Foundation to the American Association of Pediatric Dentistry (AAPD), accompanied by a shift in the tone of AAPD statements on sugary beverages, from describing soft drinks as "a significant factor" in tooth decay, to describing the scientific evidence of the relationship as "unclear." [3] Similar concerns have been raised with respect to the role of donations

[2] Implicitly we are presuming that Coasian bargaining in the political sphere does not already lead to efficient policies. To the extent that, for example, it is difficult to contract across multiple regulatory agencies and/or pieces of legislation (let alone make outright side payments), one may think of the government as aiming to set optimal policy on a rule-by-rule basis, assigning winners and losers in each instance. See, e.g., Acemoglu (2003), for a discussion.

[3] A more direct link to policy can be found in the soda industry's efforts against New York City's ban on large sugary drinks in the 2010s. In his decision to strike down the Bloomberg administration policy, the presiding judge cited amicus briefs filed by two New York non-profits (the local chapter of the NAACP and the Hispanic Federation), which argued that the ban would disproportionately affect ethnic and racial minority groups. Both non-profits were recipients of funds from Coca-Cola and PepsiCo. See "Minority Groups and Bottlers Team Up

from corporations to university research hospitals.[4]

Investigative journalists have also documented many instances of companies influencing the policy statements of "neutral" non-profits that ostensibly provide evidence-based analysis on matters of public interest. Confidential memos and documents suggest that some think tank reports are discussed with corporate donors before the research is complete, with donors potentially shaping the final reports, so that the resulting "scholarship" can be used to corroborate their separate lobbying efforts. In her 2017 book *Dark Money*, journalist Jane Mayer , provides one prominent example, documenting how the philanthropic activities of the billionaire industrialist brothers Charles and David Koch furthered their efforts to influence political discourse: "*[The Koch brothers] subsidized networks of seemingly unconnected think tanks and academic programs and spawned advocacy groups to make their arguments in the national political debate. [...] Much of this activism was cloaked in secrecy and presented as philanthropy, leaving almost no money trail that the public could trace. But cumulatively it formed, as one of their operatives boasted in 2015, a 'fully integrated network.'*" Raising concerns about such practices in general, Senator Elizabeth Warren, also a commercial law professor, observed that, "[t]his is about giant corporations who figured out that by spending, hey, a few tens of millions of dollars, if they can influence outcomes here in Washington, they can make billions of dollars."[5][6]

In this paper we show that the patterns discussed in these anecdotes hold more broadly in a setting in which we can plausibly draw a strong circumstantial connection between corporate donations and the participation of non-profits in the political and regulatory process.

We focus on the formation of federal rules and regulations. Federal agencies in the U.S. are legally required to publish proposed rules in the Federal Register and accept public comments

in Battles Over Soda." *The New York Times* March 12, 2013. Aaron and Siegel (2017) show that 95 national public health organizations received funding from Coca-Cola and PepsiCo during 2011-2015; the study does not look, however, at the effect on organizations' publicly stated positions.

[4]For example, Harris, Gardiner "Top Psychiatrist Failed to Report Drug Income." *The New York Times* October 3, 2008; Charles Piller and Jia You "Hidden conflicts? Pharma payments to FDA advisers after drug approvals spark ethical concerns" *Science News* July 5, 2018. See also Ross et al. (2000).

[5]https://www.nytimes.com/2016/08/08/us/politics/think-tanks-research-and-corporate-lobbying.html

[6]Warren also commented on the use of these practices in the rulemaking context, which we focus on in the empirical analysis below: "Unlike congressional action, agency rules are constrained by well-established judicial review standards that seek to determine whether the agency's action is supported by the evidentiary record and the authority delegated to it by Congress. Rules must be supported by "substantial evidence"; agency actions must not be "arbitrary and capricious." But corporate players are savvy. They have learned that those same judicial review standards can be used to suffocate new rules. They play a sophisticated game— leveraging their own expertise and paying outside experts with purportedly independent credentials to produce long, detailed comments filled with data and analyses, all selectively produced to serve their own interests." Discussing fixes, she also writes: "Another [principle] would be to help agencies and courts distinguish between legitimate, high-quality data and research, on the one hand, and bought-and-paid-for studies on the other, by requiring disclosure of financial arrangements and editorial relationships associated with regulatory comments." See https://www.theregreview.org/2016/06/14/warren-corporate-capture-of-the-rulemaking-process/ (accessed October 31, 2018).

on those proposals before rules are finalized and comments discussed.[7][8] While there is no legal requirement for agencies to act on feedback received in comments, the agencies themselves often attribute changes between proposed and final rules to arguments made via rulemaking. As emphasized by Sunstein (2012), public commentary is also a valuable source of feedback to preempt regulatory mistakes "when the stakes are high and the issues novel." Regulations.gov provides the largest single source for comment information on proposed rules, and was rolled out in 2003 when most agencies started a systematic effort to digitize the commenting process. By 2008, 80% of all proposed rules provided a regulations.gov link for commenting, and the fraction is about 90% as of 2018.

For the purpose of this paper, we use regulations.gov to build a comprehensive dataset including the majority of the comments submitted in the rulemaking process since the 2003. For each comment, we know the specific proposed rule the comment is in response to, as well as the content (text file) of the comment and the identity of the commenter. We may thus connect specific organizations to commentary on the same proposed regulation and its final discussion (we refer to a sequence of rule postings from proposal to final version as a "regulatory stream" or docket).

We complement the commentary data with information on corporate foundations and their beneficiaries, using data on charitable donations by foundations linked to large corporations through tax forms filed to Internal Revenue Service (IRS). The combination of these datasets allows us to explore whether (i) non-profits that benefit from corporate philanthropy are more likely to comment on the same rule as their benefactors; (ii) conditional on both providing feedback on the same regulation, the non-profits' comments are unusually similar to that of their benefactors; and (iii) co-comments by a corporate foundation's grantees lead to discussions of the rule by the regulator that use language that is more similar to the language contained in the company's comments. By exploiting the particular timing of corporate donations and comments, as well as the inclusion of firm-grantee pair fixed effects, we argue that we can plausibly draw a compelling link from funding to co-commentary and comment overlap.[9]

Our sample of firms is comprised of the companies that have appeared at any point in the 1995

[7]The Administrative Procedures Act of 1946, 5 U.S.C. 553(c) states: "...the agency shall give interested persons an opportunity to participate in the rule making through submission of written data, views, or arguments with or without opportunity for oral presentation. After consideration of the relevant matter presented, the agency shall incorporate in the rules adopted a concise general statement of their basis and purpose." https://www.law.cornell.edu/uscode/text/5/553. Accessed October 31, 2018.

[8]There are some exceptions for urgent actions or cases in which the change is so trivial that the agency does not expect comments, but in general, agencies which fail to publish a sufficiently informative proposal or fail to follow the commenting procedure can have their regulations vacated in court.

[9]Of course, this does not obviate the possibility that non-profits have time-varying policy preferences, and corporate gifts coincide with shifts in these preferences. While we cannot rule out this possibility (a critique that applies even to the Coca-Cola/AAPD example mentioned above), our approach does help to rule out the possibility that latent, time-invariant shared interests drive both donations and comment overlap.

to 2016 lists of Fortune 500 or S&P 500 (or both) for which we identify a corporate foundation, and our sample of non-profits is the set of all grantees that received at least one donation from these foundations over the period 1998-2015. Organizations (firms or non-profits) are linked by name via a fuzzy match to 981,232 rulemaking comments made on all proposed regulations on regulations.gov during the years 2003-2017. The main sample for our analysis is comprised of the 414 corporations with charitable foundations and 11,746 grantees that commented at least once during this period.

In our first set of results, we show that non-profits are more likely to comment on the same regulation as their benefactors, and that this "co-commentary" is most strongly associated with donations in the year preceding the comments, a result which survives the inclusion of firm-grantee fixed effects. The magnitude of the estimated relationship between donations and co-commentary is very large: even with firm-grantee fixed effects, our analysis implies that a donation in the preceding year is associated with nearly a doubling in the likelihood of co-commentary.

Our findings on the link between donations and co-commentary frequency point to potential influence over non-profits in their regulatory feedback. In our second set of results we examine whether, conditional on co-commentary, the content of comment-pairs from firms and non-profits linked via charitable donations tend to be more similar, relative to other comments on the same proposed rules. Using established methods of natural language processing, we generate pairwise measures of textual similarity between any two firm-non-profit comments on a given rule. Co-comments by non-profits contain content that is more similar to comments by their corporate benefactors relative to other co-comment pairs and, importantly, the timing of this relationship parallels that of our first set of findings – co-comments in the year immediately following a donation are most similar.

In addition, we show that the co-commenting relationship matters for final rules. Focusing on all comments made by corporations in our dataset, we show that, if a grantee (and particularly one receiving a recent donation) also commented on the proposed regulation, the language of the discussion of the final rule is more closely aligned with that of the corporation's comments. This result survives the inclusion of both firm fixed effects and rule (docket) fixed effects, and also holds when we measure a firm's influence based on whether it is cited by the regulators in their discussion of the final rule.

Finally, we explore whether corporations use charitable donations to encourage otherwise opposing voices to remain silent (rather than encouraging non-profits to provide supportive commentary). While it is challenging to devise a decisive test to detect the omission of comments that might otherwise have been made, we provide suggestive evidence, based on an extension of our main results on donations and co-comment frequency, that "hush money" may not be of first-order importance in our setting. More specifically, we show that the link between co-commentary and donations is strongest in areas in which a non-profit most commonly provides comments, the opposite of what one might expect if hush money played a dominant role.

Our findings first and foremost provide a contribution to the literature on the mechanisms by which interest groups seek to influence government policy (for canonical early contributions see, for example, Grossman and Helpman, 1994, 2001 and for a more recent discussion Baumgartner et al., 2009; Bertrand et al., 2014; Drutman, 2015). We differ from much of this prior work in our focus on influence via expert commentary rather than financial contributions and, much more importantly, in documenting one mechanism by which private interests may cloak biased advice by inducing its provision by a non-obviously aligned party. This has implications for how we model the process of governmental information acquisition (Austen-Smith, 1993; Laffont and Tirole, 1993), and is also of direct policy relevance. Our results suggest that calls for restrictions on financial relationships among those aiming to influence government policy may be well-founded, and that at a minimum potential conflicts-of-interest statements should be required for any organization providing input on government regulations (Peng, 2016). Our work is also related to prior research that has shown the value of coalitions of diverse interest groups in the adoption of government policy. In particular, studying bills introduced in Congress between 2005 and 2014, Lorenz (2017) shows that bills supported by interest-diverse coalitions are more likely to receive committee consideration; in contrast, Lorenz (2017) finds no association between committee consideration and lobbying coalitions' size, or their interests' PAC contributions. Generalizing beyond the lawmaking process, this work complement our findings in that it suggests that corporations can expect some return for the type of charitable "investments" we uncover in this paper. Other papers that have focused on returns to lobbying instead include Bombardini and Trebbi (2011, 2012); Kang (2016); Kang and You (2016). Finally, our paper expands on earlier work highlighting how corporations may strategically use their corporate philanthropy as an undisclosed tool of political influence. Bertrand et al. (2018) show that corporations allocate more of their charitable giving to congressional districts that are more relevant to the corporations due to the committee assignments in the House of Representatives of their elected representatives. We identify in this paper another, independent, category of "strategic CSR" (Baron, 2001) in the government arena.

The rest of the paper is organized as follows. Section 3 presents our data on U.S. federal regulation and corporate donations. Section 4 introduces parallel analyses of corporate giving and regulatory rulemaking public commentary that explores whether contributions flow to non-profits that comment on rules on which firms also comment. Section 5 presents evidence on the excess similarity between the content of comments filed by non-profits and corporations around the time corporations provide charitable grants to the non-profit. Section 6 assesses whether co-commenting by a grantee is associated with rules whose language is more aligned with that of the grantee's corporate benefactor's comments. Section 7 shifts the focus on tests for hush money. Section 8 concludes.

2 Institutional context: Rulemaking process

The rulemaking process of U.S. federal agencies provides a context in which we may observe both the presence and the content of communication by different entities with an interest in influencing the policymaker. While informational lobbying at the federal or local level does not come with statutory requirements of disclosure of the content or even the exact target of communication,[10] the rulemaking process consists of a series of codified procedures that regulate the activity of federal agencies in the production of "rules" under the Administrative Procedure Act of 1946 (APA). The subject of policy deliberation is a rule "designed to implement, interpret, or prescribe law or policy," according to the APA. The process of rulemaking may be set in motion by Congress passing a new law requiring implementation or by an agency itself, upon regularly surveying its area of legal responsibility and identifying areas that need new regulations.[11]

Figure 1 sketches the process of informal rulemaking. It starts with a Notice of Proposed Rulemaking (NPRM) including the objective of the rule and how it would modify the current Code of Federal Regulations. The NPRM is published in the Federal Register, at which point the agency specifies a period of 30 to 60 days during which the public can submit comments on the proposed rule. After comments have been received and additional information collected, the agency may proceed to publish a Final Rule in the Federal Register or issue a Supplemental Notice of Proposed Rulemaking if the initial rule was modified substantially, in which case further comments are invited. This notice-and-comment procedure is meant to include the general public and all interested parties in the crafting of the new rule. Importantly, the agency also publishes in the Federal Register a discussion of the goals and rationale of the policy, and how the comments were incorporated into the final rule in the Supplementary Information section of the final rule.

Occasionally, the process of rulemaking requires merging or splitting specific elements of a rule, issuing interim versions of the rule if the process is delayed, and more generally adapting to other external factors, including further direction from the legislative branch, and so forth. These various additional documents are typically filed in dockets maintained by the regulatory agency.

Upon finalization of the rule, comments represent part of the official record, and rules can be challenged judicially on procedural or substantive grounds based on comments filed by entities that participated in the rulemaking. Judicial review is an important constraint to rulemaking

[10]Under the Lobbying Disclosure Act of 1995, lobbying registration and reporting forms only require lobbyists to list the topic and the agency lobbied (e.g., Trade, the Senate of the United States), in addition to clients and payments. See Vidal et al. (2012); Bertrand et al. (2014).

[11]Agencies may decide to engage in rulemaking under the recommendation of congressional committees, other agencies, or following a petition from the general public.

activity in the United States in that it effectively forces regulators to attend to opinions expressed via commentary.

3 Data

This section introduces our sources and provides a brief overview of the data. For further details we refer to Appendices A and B. We begin by describing the data on charitable giving by corporate foundations, followed by the data on public comments on rulemaking. The starting point for our sample is the set of corporations that have appeared at any point during the period 1995 to 2016 in the Fortune 500 and/or S&P 500 lists, which counts 1398 firms.[12]

3.1 Charitable giving by foundations

Data on charitable donations by corporate foundations come from FoundationSearch, which digitizes publicly available Internal Revenue Service (IRS) data on the 120,000 largest active foundations in the U.S. We find 629 active foundations that can be matched by name to 474 of the initial list of 1398 firms.[13] As noted in Brown et al. (2006), larger and older companies are more likely to have corporate foundations, which results naturally from the fixed cost of establishing a foundation.[14]

Each charitable foundation must submit Form 990/990 P-F "Return of Organization Exempt From Income Tax" to the IRS annually, and this form is open to public inspection. Form 990 includes contact information for the foundation, as well as yearly total assets and total grants paid to other organizations. Schedule I of Form 990, entitled "Grants and Other Assistance to Organizations, Governments, and Individuals in the United States," specifically requires the foundation to report all grants greater than $5,000. For each grant, FoundationSearch reports the amount, the recipient's name, city and state, and a giving category created by the database.[15]

While the IRS assigns a unique identifier (Employer Identification Number, EIN) to each non-profit organization, FoundationSearch does not report this code, so we rely on the name, city and state information to match a grantee to a master list of all non-profits. This list, called the Business

[12] The initial number of firms is 1434, but we combine firms that merge during the sample, hence obtaining a smaller total number.

[13] The 629 foundations we find are linked to 474 corporations, since there are instances of multiple foundations associated with the same corporation.

[14] They also find that state-level statutes – in particular laws relating to shareholder primary and the ability of firms to consider broader interests in business decisions – predict establishment of a foundation. Various endogenous financial variables are also predictive of foundation establishment. The analysis in Brown et al. (2006) is cross-sectional, so their variables are absorbed by the various fixed effects in many of our analyses.

[15] The 10 broad categories are: Arts & Culture, Community Development, Education, Environment, Health, International Giving, Religion, Social & Human Services, Sports & Recreation, Misc Philanthropy.

Master File (BMF) of Exempt Organizations, is put together by the National Center for Charitable Statistics (NCCS) primarily from IRS Forms 1023 and 1024 (the applications for IRS recognition of tax-exempt status). The BMF file reports many other characteristics of the recipient organization, including address, assets and a non-profit sector called the National Taxonomy of Exempt Entities (NTEE). The results of the matching between all public charities, private foundations or private operating foundations (designated as 501(c)3 organizations for tax purposes) in the BMF and the recipients of charitable giving by 2014 Fortune 500 and S&P 500 companies is reported in Bertrand et al. (2018).

3.2 Comments and rulemaking

The source of data on comments, proposed, final and interim rules, as well as discussion of final rules is regulations.gov, a website through which the majority of U.S. federal agencies collect public comments in the notice-and-comment phase of rulemaking.[16] The website regulations.gov API provides a search function for document metadata.

Our research sample consists of all comments posted to regulations.gov in the years 2003-2017. We use a custom machine learning tool to extract organization names from the comment metadata. The algorithm identified 981,232 comments that appear to be authored by organizations (as opposed to private individuals) and downloaded the full text of the comments. We are particularly interested in comments submitted by non-profits and by corporations that we observe in our FoundationSearch sample. The comments are linked to corporations' and grantees' names through a custom name matching tool that implements multiple types of fuzzy matching and manual corrections.[17]

The unit of observation is what regulations.gov refers to as a docket. This is a way for agencies to organize comments that relate to a particular topic. Most straightforwardly, one may think of a one-to-one correspondence between a rule and a docket. Conceived in this way, as mentioned above, a docket will contain all comments that pertain to all versions of that rule. An example of a simple docket is FNS-2006-0044 from the Food and Nutrition Service (FNS) which contains a proposed rule (06-09136) and its corresponding final rule (E8-21293) on "Fluid Milk Substitutions in the School Nutrition Programs." All comments in this docket therefore are easily linked to this regulatory stream. There are more complex cases in which a docket contains multiple proposed rules and notices (see, for example, docket EPA-HQ-OAR-2008-0699, the Environmental Protection Agency's review of the National Ambient Air Quality Standards for Ozone). We associate all comments to the same docket given the homogeneity of the topic. The only exception is when we turn to examine the wording of the discussion of final rules as a function of corporate and

[16]For the complete list, see Appendix tables A.7 and A.8.
[17]Available from the authors upon request.

non-profit comments. There, we will consider each rule within the docket separately to ensure a finer connection between comments by corporations and the exact wording of the final rule in the docket under discussion, as the multiple rules associated within a docket are discussed and published separately in the Federal Register. We will elaborate on this distinction in Section 6, which discusses those results.

3.3 Basic data facts

Recall that our sample starts with the set of companies that appeared at least once in the Fortune 500 or S&P 500 lists between 1996 and 2015. Of the 1398 firms in that sample we find 909 that have commented at least once in the period 2003-2016.[18] This is the sample of firms that forms the basis of our regressions. We have a total of 22,654 firm comments over 5,792 dockets. Of these 909 firms 414 have a foundation.

In terms of non-profits we start from the 225,180 entities that received at least one grant from any foundation in our sample over the period 1998-2015. Our sample consists of the 11,531 of these grantees that comment at least once at any point during the period starting in 2003. We have a total of 318,841 comments in 8,729 dockets from those grantees.

There is vast heterogeneity among firms in their activity in the commenting phase. The most actively commenting firm, Boeing, provided comments on 1284 dockets. On average each firm comments on 18 dockets, but the distribution is skewed: the median firm comments on 6 dockets, while the firms at the first and third quartile comment on 2 and 17 dockets, respectively. The distribution of comments among grantees is even more skewed. On average each grantee comments on almost 5 dockets, but the median is 1 and the third quartile is 3 dockets. The most active grantee (Center for Biological Diversity) comments on 905 dockets.

Tables A.2 reports the agencies that receive the highest number of comments from grantees and firms.[19] At the top of the list for grantees are the EPA (Environmental Protection Agency), the FAA (Federal Aviation Administration) and the FDA (Food and Drug Administration). The top three agencies as recipients of grantees' comments are the FWS (Fish and Wildlife Service), the NOAA (National Oceanic and Atmospheric Administration), and the HHS (Health and Human Services Department). It is worth noticing that the EPA, the FAA and the FDA feature in the top 10 agencies for grantees as well.

Finally, we provide some information on the prevalence of commenting behavior of grantees and their co-commenting with firms in our sample. In our regressions we will often focus on "recent" donations, defined as donations from a firm to a grantee that occur in the same year, or one year prior to a public comment on a rule. Consider the set of all firms-years where the firm

[18] We only consider comments starting in 2003 because this is when the comments database is complete.
[19] Agency acronyms are listed in Appendix tables A.7 and A.8.

has commented at least once and donated recently. We can break the recipients of these recent donations in to a set of nested groups with increasingly close ties to the firm.

Firms donate to an average of 327 non-profit grantees. Of these, an average of 54 grantees ever submit a comment in our sample. Within these "commenters", 28 non-profits ever comment to one of the same agencies as the firm (not necessarily at the same time or in the same year), 8 ever co-comment on a regulation with the firm, and 1.4 co-comment with the firm that year.

In terms of expenditures, the average total amount spent on donations over a two year period is \$26 million dollars, 26% of which go to grantees that ever comment. Within commenting grantees, expenditures are biased towards the grantees with closer commenting ties to the firm. Thus, grantees that never comment to the same agency as the firm receive an average of \$103,412 each, and grantees that comment to one of the same agencies as the firm, but never on the same regulation, receive \$156,348 each, while those that ever co-comment with the firm receive an average of \$240,515 each and grantees that co-comment with the firm that specific year receive \$206,994 each.

4 Evidence based on charitable giving and non-profit commenting on regulations

This section focuses on the link between firms and non-profits through charitable grants, and establishes a relationship between firm-grantee financial ties and their tendency to comment on the same regulations.

We denote firms/foundations by $f \in F$ and grant-receiving non-profits ("grantees") by $g \in G$. Let D_{fgt} be an indicator function that takes a value of 1 if we observe a donation from firm f to grantee g in year t, and 0 otherwise. The indicator function C_{frt} is equal to 1 if firm f comments on regulation r in year t, and 0 otherwise (throughout this section and the following one, "regulation" or "rule" will refer to a docket). The indicator function C_{grt} is defined similarly and is equal to 1 if grantee g comments on regulation r in year t, and 0 otherwise. A graphical representation of this configuration is described in Figure 2.

We adopt two types of specifications: co-commenting specifications and a regulation specification.

4.1 Co-commenting specifications

We begin by relating the event of a firm and a grantee commenting on the same regulation to a financial tie between the two in the form of a charitable donation. The indicator function CC_{fgrt} is equal to 1 when donor f and grantee g comment on the same regulation r at time t, so that $CC_{fgrt} = C_{frt} \times C_{grt}$, and 0 otherwise. Our first specification explores a time-invariant link between co-commenting and donations, aggregating co-commenting to the firm-grantee pair, so that we define a new indicator CC_{fg} which is equal to 1 if we observe any co-commenting from firm f to grantee g in our sample, and 0 otherwise. That is, $CC_{fg} = I\left(\sum_r \sum_t CC_{fgrt} > 0\right)$. Similarly, the indicator variable D_{fg} indicates whether we observe any donation from f to g in our sample.

We first consider the following time-indepentent specification that relates the presence of co-commenting by firm f and grantee g to the presence of a donation within the same pair:

$$CC_{fg} = \beta_0 + \beta_1 D_{fg} + \delta_f + \delta_g + \varepsilon_{fg}. \tag{1}$$

The specification includes firm fixed effects δ_f to capture the potential bias resulting, for example, from the higher probability that large and profitable firms both donate to charities and comment on multiple regulations. Similarly, we include grantee fixed effects δ_g, to control, for example, for the fact that charities that are more successful at fundraising may on average have more resources to devote to commenting on various regulations. A positive coefficient $\beta_1$would indicate that firm-grantee pairs that are connected by donations are also more likely to comment on the same regulations.

The results are reported in Table 1. The different columns of Table 1 include different sets of fixed effects (and clustering dimensions) of increasing levels of stringency. Of particular interest is column (4), the most conservative specification, that includes both grantee and firm fixed effects. The firm fixed effects may account for the average propensity of firms to comment and to donate, which may depend on size and sector. The grantee fixed effects can capture the average level of commenting activity of the grantee, which may in turn be related to its size and overall resource endowment. Across all specifications in Table 1,we can see that a grantee and a firm are more likely to comment on the same regulation when we observe any donation from the firm to the grantee. The magnitude of this effect is large. The baseline probability of co-commenting for a firm and a grantee is 2.16%, meaning that of all the possible pairs of grantees and firms only 2.16% comment on the same rule at any point in time. This probability increases by 4 to 8 percentage points when we observe a donation connecting firm and grantee. Put differently, the presence of a donation is associated with a two- to four-fold increase in the probability of co-commenting.

Of course, this cross-sectional pattern of co-commenting may stem from the fact that firms contribute to non-profits sharing similar objectives and views, or that, more simply, operate in similar sectors. For instance, the Bayer Science & Education Foundation associated with Bayer US, a pharmaceutical company, may be more likely to donate to healthcare-related research non-profits, and both Bayer and healthcare-related non-profits may be more likely to comment on healthcare-related regulation than an average organization.

Our second specification addresses this concern, and further allows us to control for the general tendency by some firms to comment on certain issues and to contribute to non-profits that operate in related areas. It does so by focusing on the timing of donations. In particular, we examine whether co-commenting is more likely in the year immediately following the presence of a donation. For this, we turn to the following panel specification, which exploits time variation in both co-commenting and donations:

$$CC_{fgt} = \beta_0 + \beta_1 D_{fgt-1} + \delta_{fg} + \delta_t + \varepsilon_{fgt} \tag{2}$$

where $CC_{fgt} = I\left(\sum_r CC_{fgrt} > 0\right)$ indicates whether firm f and grantee g comment on the same regulation at time t, and D_{fgt-1} is equal to 1 if we observe a donation from f to g in the concurrent (t) or preceding ($t-1$) year of the comments, 0 otherwise. This specification includes firm-grantee fixed effects δ_{fg} and time fixed effects δ_t. Therefore, β_1 is estimated only employing within-pair variation over time in donations and co-commenting. In particular β_1 will detect whether, controlling for the average tendency of a certain firm f to co-comment with and donate to a specific non-profit g, we observe co-comments occurring immediately after a donation from f to g has been made.

The results are reported in Table 2, where we separate contemporaneous and lagged donations. Given the coarseness of the data along the time dimension (we only observe year of comment), it is possible for a comment to be made in, say, January of 2006 and a donation in June 2006; hence we can only be certain that the lagged-year donation took place prior to co-commenting. In Table 3, we additionally report results in which we combine the two D variables, to crate a dummy that is equal to 1 if we observe a donation at either t or $t-1$, and 0 otherwise. Our preferred specification in both Tables 2 and 3 is column (5), where we include firm-grantee pair fixed effects. This specification exclusively exploits variation within a firm-grantee pair in donations and in co-commenting. The δ_{fg} pair fixed effects control not only for the higher probability of donation and co-commenting for firms and grantees in the same sector, but also for the general ideological alignment of firm and grantee that may result in both donations and co-commenting on similar topics.

We find a robust association between donations in year $t-1$ and the likelihood of co-commenting in year t. The magnitude of effects is large in this panel specification. Co-commenting is obviously more sparse in equation (2) than equation (1): of all firm-grantee-year triples only 0.163% feature co-commenting. In column (4) of Table 3, the presence of a recent donation is associated with a quadrupling of the probability of co-commenting. In column (5) of Table 3, the presence of a recent donation is associated with a 81% increase in the likelihood of co-commenting, even after controlling for the general propensity of a specific firm to give to and as well as co-comment with a specific grantee. The example provided in the introduction, which described AT&T Foundation grantees such as GLAAD or a homeless shelter commenting on the AT&T/T-Telecom merger close on the heals of receiving donations, provide an illustration of the behavior implied by this statistical evidence (see Peng (2016) for other illustrations).

As a further robustness exercise, in Appendix Table A.3 we augment our preferred specification with a dummy for whether firm f donated to g in year $t+1$. In column (5) of that table, with the most restrictive set of fixed effects (i.e. pair fixed effects), we find that donations made immediately after the commenting period are not associated with co-commenting, whereas only immediately preceding donations are. This pattern further confirms that co-commenting seems to be more prevalent after we observe a donation from firm to grantee.

4.2 Regulation specification

In the specifications we have considered thus far, we have aggregated co-commenting across different rules within a fg pair or fgt pair-year. We now present an alternative approach that links commenting by a grantee to donations received by a firm that also comments on the same rule r. The following "regulation specification" relates the probability of commenting by a grantee on a regulation r to donations received:

$$C_{gr} = \beta_0 + \beta_1 \underbrace{I\left(\sum_f D_{fg} \times C_{fr} > 0\right)}_{DonorComment_{gr}} + \delta_g + \delta_r + \eta_{gr}$$

where C_{rg} is equal to 1 if g comments on regulation r (0 otherwise) and $DonorComment_{gr} = I\left(\sum_f D_{fg} \times C_{fr} > 0\right)$ is equal to 1 if g receives a donation from any firm that comments on r, and 0 otherwise. This specification includes regulation fixed effects δ_r, which capture how certain rules are subject to more intense commenting, and grantee fixed effects δ_g, that account for factors like resources and size of the non-profit, which may make g both more visible and more likely to comment on any regulation.

Table 4 reports estimates of β_1 under different fixed effects and clustering options. Our preferred specification in column (4) has docket and grantee fixed effects, as well as two-way clustering on these attributes. When considering all the possible combinations of grantees and rules, we find a comment in 0.039 percent of the cases. It is not surprising that this number is small, since the universe of all possible grantee-rule pairings involve non-profits, like the Red Cross, that we would not expect to comment on, say, financial regulation. Starting from this baseline probability of commenting on a specific rule, we find that the probability that the non-profit comments is three to five times higher when a donor firm commented on the same rule, a result that accords with our previous results under specification (2).

5 Quantifying the similarity in content across regulatory comments

So far the focus of the analysis has been on the propensity to comment on regulation. However, a crucial implication of our thesis that non-profits may act as strategic advocates for their corporate donors is that the content of the message delivered by non-profits to regulators may be affected by financial connections. In particular, upon receipt of (a) charitable grant(s), comments targeted to federal regulators by non-profits should be closer in content to the messages sent by their corporate benefactors (relative to the counterfactual of no corporate donations). To provide evidence in this direction, we build a portfolio of circumstantial findings with the intent of discriminating among alternative theoretical mechanisms based on how well they match the empirical regularities that we present.

To build intuition (and without intent to claim any deliberate deception by the parties involved in this particular instance), consider the example of Bank of America's donation of $150,000 to the Greenlining Institute in 2010. While Bank of America is the second largest bank in the United States by total assets and is a central player in housing finance in the country, the Greenlining Institute is a non-profit focused on improving access to affordable housing and credit to low-income families and minorities (African American, Asian American, and Latino, in particular). In 2011 both organizations commented on the Office of the Comptroller of the Currency's Credit Risk Retention (CCR) docket,[20] as part of one of the regulatory rulemaking streams initiated under the Dodd-Frank Act of 2010 (Title IX, Subtitle D, Section 941). CCR, also known as the "skin in the game" rule, imposed a 5 percent retention requirement on all mortgage loans originated by lenders in the United States to moderate "originate-to-distribute" moral hazard problems pervasive in the build-up to the 2008 financial crisis.

The main comment submitted by Bank of America[21] remarks that, in relation to relaxing the definition of qualified mortgages exempted from retention requirements on the issuing bank's balance sheet (i.e. of mortgages deemed safe enough not to warrant the restriction): "*...the PCCRA provision will cause some borrowers to be unable to obtain a loan at all. In the currently tight private residential mortgage market, borrowers already must provide significant down payments.*" The Greenlining Institute provides a similar assessment in its comment,[22] suggesting that "*by raising the barrier to affordable home ownership with an unreasonable 20% down payment requirement, we will not only keep families from rebuilding after foreclosure, but we will prohibit an entire generation of first time borrowers from owning a home, despite lower home prices across*

[20] Docket ID OCC-2011-0002

[21] Document ID OCC-2011-0002-0141

[22] Document ID OCC-2011-0002-0353

the country." In sum, both organizations appear to advocate openly for laxer definitions of the CCR exemptions, limiting the rule's bite, and allowing assets with substantially lower quality and higher risk to be exempt – an effort that ultimately succeeded in entirely defanging the rule.[23]

In this section, we provide a framework for examining the content and textual similarity of comments filed by non-profits and firms, and show that, upon receipt of a donation from a firm's foundation, comments by a non-profit are more similar to those of its donor, suggesting that the Bank of America-Greenlining example holds more broadly in the data.

We compute approximate measures of semantic similarity of pairs of public comments using Latent Semantic Indexing (LSI) with bag-of-words features. LSI is an established technique borrowed from the natural language processing literature, and it has been shown to perform well on a variety of different document classification and retrieval tasks.[24] LSI requires the conversion of text documents into vectors of word counts and applying term frequency–inverse document frequency feature extraction within each regulatory docket r. Following this preparation phase, one can compute document-level singular vectors from a singular value decomposition of the text matrices and take the cosine similarity of any pair of document vectors. This approach provides a similarity score S_{fgr} normalized by the standard deviation in each docket r and distributed between -1 and 1 for every pair of texts formed by a comment by firm f and a comment by grantee g within a given docket. To further demonstrate the validity of our approach, we show in Appendix B that our measure performs well in a classification task of separating documents from different regulations and in clustering comments from similar organizations.

Using this comment-pair similarity score as the outcome, we consider a specification of the form:

$$S_{fgr} = \beta_0 + \beta_1 D_{fgr} + \delta_f + \delta_g + \delta_r + \varepsilon_{fgr}$$

where the coefficient of interest is β_1 and D_{fgr} is indicator variable that equals 1 if firm f donates to grantee g, 0 otherwise. As the timing of such donations is a useful discriminant for interpretation of our findings, we will be careful in constructing D_{fgr} under different time horizons. The dataset we exploit for this analysis includes all possible firm-grantee pairs of comments conditional on commenting on a docket r.

We begin by exploring the sign and magnitude of the estimated coefficient β_1 when the donation indicator variable takes the value of one in the event of any grant from f to g over our entire time period. Table 5 reports estimates for β_1 across a set of four specifications with an incremental inclusion of firm, grantee, and docket fixed effects. Coefficients are clustered at the docket, firm-

[23]For a discussion, see Floyd Norris for the *New York Times*, Oct. 23, 2014, Page B1 "Banks Again Avoid Having Any 'Skin in the Game"', available at https://www.nytimes.com/2014/10/24/business/banks-again-avoid-having-any-skin-in-the-game.html

[24]See Dumais et al. (1988) and Deerwester et al. (1990). For a more recent discussion of latent semantic analysis, see Dumais (2004).

grantee, or double clustered at both levels depending on the specification. The estimates of β_1, which capture the increase in units of standard deviations of similarity across comment pairs within each r, range from 0.25 to 0.09 in the most restrictive specification (all significant at least at the 1 percent level). This indicates that pairs of comments made by firms and their grantees are more similar relative to a baseline similarity obtained by pairing comments at random within a docket.

As with our results on comment propensity in Section 4, the presence of a donation at any point in our sample period may proxy for some average similarity in the interests and beliefs of a firm and its grantee. Table 6 thus focuses on donations that take place in either the year in which the comments are filed (year t) or in the previous fiscal year ($t-1$). The point estimates are smaller in magnitude across comparable columns in Tables 6 and 5, but statistically indistinguishable. In separating explicitly contemporaneous donations and those made in the fiscal year immediately preceding the comments, as reported in Table 7, we observe that precision and magnitude of the effect come from the donations made at time $t-1$. The estimates, which capture the increase in units of standard deviations of similarity across comment pairs within each docket, range from 0.17 to 0.08 in the most restrictive specification.

Table 8 addresses the concern that the timing of donations may be spuriously related to some underlying tendency of firms and grantees working in related areas of interests, by controlling in our most restrictive specifications also for North American Industry Classification System (NAICS) 6 sector code of the firm interacted with the IRS's National Taxonomy of Exempt Entities Classification (NTEEC) code of the non-profit. As can be seen in the table, the estimated coefficient β_1 remains precisely estimated and within the confidence intervals of our baseline estimates across specifications when accounting flexibly for such industry pair controls. Finally, notice that the reduction in sample size for this table results from missing sector information for some firm-grantee pairs, and that this sample shift also does not affect the point estimates relative to the baseline specifications. More precisely, we estimate a β_1 of 0.073 in column (4) of Table 6 and of 0.074 in column (1) of Table 8, and a β_1 of 0.079 for D_{fgr} at time $t-1$ in column (4) of Table 7 and of 0.072 in column (3) of Table 8.

Finally, we present a placebo exercise that underscores the very specific timing of the link from donation to comment similarity. In particular, we modify our definition of donations to focus on the period immediately *after* the regulatory commenting phase. Table A.4 reports these results. As can be seen in the table, across specifications with incremental sets of fixed effects and industry controls, the estimated coefficient β_1 appears insignificant and smaller in magnitude relative to our base estimates.[25] This placebo exercise is informative along several dimensions.

[25]In the last column of the table, we also include donations at t or $t-1$, and show that only pre-comment donations matter, relative to donations at $t+1$.

As the donation is close in time to the commentary activity, but statistically and economically insignificant, these findings further assuage the concern that our results may be spuriously driven by some underlying tendency of firms and grantees operating in related areas. The systematic timing of excess similarity between comments' texts just following the disbursement of a charitable grant offer intuitive support to the logic of some form of suasion being exerted by the donor over the grantee.

6 Comment impact analysis: Evidence from final rule citations

While the preceding sections focus on the frequency and similarity of firm-grantee comments, we now turn to examining whether firms' comments – and the similar comments made by grantees – comments have an impact on rulemaking. As it is typically very hard to assess the effects of advocacy on policy outcomes (and in general of informational lobbying on government policy choices), we will focus here on a newly devised approximation for such outcomes by asking how the final rule was shaped by the commentary. In particular, we aim to establish that when a firm comments on a rule, the published discussion of the rule by the regulator is closer in content similarity to the firm's comments when the firm's grantees also comment on that rule.

It is important to clarify that the final regulatory text itself is written with a terminology and structure that makes it very different from comments submitted or the explanation of the rule itself offered by the regulator in the preamble to the rule. The final regulatory text is designed to formulate, amend, or repeal sections of the Code of Federal Regulations (5 U.S.C. § 551(5)). The discussion of the rule itself offers a justification and analysis of the regulator's decision making process and intended scope or interpretation of the regulation.[26] In fact, the discussion of the rule tends to be longer and reveals arguments in favor of or against specific choices that may have been brought forward by, for example, the comments from various entities, firms and grantees, in persuading the regulator. We therefore focus on this part of the final rule.

As an example consider the concern expressed by Wells Fargo, one the U.S. largest depository institutions, on a specific regulatory burden that appeared implied by the proposed rule version of the so called Volcker Rule of the Dodd-Frank Act of 2010. The Volcker Rule aimed at prohibiting depository institutions from engaging in the use of part of its depository funding for speculative trading (proprietary trading).[27] Wells Fargo expresses concern that the proposal

[26] The discussion of the rule is found in the Supplementary Information section, which is part of the preamble to the final rule and typically constitutes its most important component. See https://www.federalregister.gov/uploads/2011/01/the_rulemaking_process.pdf

[27] Docket ID OCC-2011-0014

requires transaction-by-transaction oversight: "*We also do not believe that the Proposed Rule's transaction-by-transaction approach, which would require analyzing permitted customer trading, market making, underwriting and hedging activities on a transaction-by-transaction basis, is the best way for the Agencies to implement the Proposed Rule...*"[28] The OCC addresses this concern directly and concedes some changes to the rule: "*A number of commenters expressed general concern that the proposed underwriting exemption's references to a 'purchase or sale of a covered financial position' could be interpreted to require compliance with the proposed rule on a transaction-by-transaction basis. These commenters indicated that such an approach would be overly burdensome. ...A general focus on analyzing the overall 'financial exposure' and 'market-maker inventory' held by any given trading desk rather than a transaction-by-transaction analysis.*" Importantly, also the Black Economic Council, a recent Wells Fargo grantee, is found to express concerns on the same rule on grounds of excessive complexity.[29]

We begin by defining S_{fr} the similarity score between the discussion of docket r and firm f's comment. In contrast to the score constructed in Section 5, S_{fr} measures the similarity between a comment and the discussion of the rule in a docket, rather than the similarity between the texts of two comments on a rule. S_{fr} is designed as a proxy for the salience and effectiveness of the firm's comment in shaping the regulator's decisions. As with the previous similarity measure S_{fgr}, we normalize S_{fr} by the standard deviation in each docket r, so that S_{fr} is distributed between -1 and 1 for every pair of texts.[30]

Dropping time subscripts, let us posit S_{fr} as function of the commenting effort of the firm and of grantees connected to the firm by donation:

$$S_{fr} = \beta_1 \sum_g CC_{frg} \times D_{fg} + \beta_2 \sum_g D_{fg} + \beta_3 \sum_g CC_{frg} + \delta_f + \delta_r + \varepsilon_{fr}$$

Focusing on the extensive margin of commenting behavior, we can replace all sums with indicator functions and also include firm and docket fixed effects:

$$S_{fr} = \beta_1 I\left(\sum_g CC_{frg} \times D_{fg} > 0\right) + \underbrace{\delta_f + \beta_2 I\left(\sum_g D_{fg} > 0\right)}_{\text{Firm FE}} + \underbrace{\delta_r + \beta_3 I\left(\sum_g CC_{frg} > 0\right)}_{\text{Docket FE}} + \varepsilon_{fr} \tag{3}$$

[28] Document ID OCC-2011-0014-0285)

[29] Document ID OCC-2011-0014-0024)

[30] As in some cases multiple rules may be included in a docket by regulators (including amendments, notices, etc.) and each regulatory stream can be linked to a final rule, our approach here is to take for each firm and docket the closest in similarity to the firm's comment vector. This is meant to more accurately represent the dimension of the docket the firm more closely commented about. Our results are similar when removing the lowest similarity score within a docket-firm group and then taking the mean similarity or when keeping only dockets with exactly one rule document. See Online Appendix for these robustness checks.

The variable of interest is $I\left(\sum_g CC_{frg} \times D_{fg} > 0\right)$, which is equal to 1 if we observe a donation by the firm to a grantee co-commenting on the same rule, and 0 otherwise. If there is excess similarity between rule discussion and a firm's comment when grantees connected to the firm by donation also comment on that rule, we expect β_1 to be positive. As we established in the previous two sections, such comments by non-profits occur around the time of firm donations and appear to exhibit a systematically higher textual similarity to the comments filed by the grantee's benefactors. Here, we aim to establish that corporate benefactors appear to gain in terms of S_{fr}, a proxy that at a minimum captures having the attention of the regulator, but could conceivably correlate with influence in shaping the final rule text or keeping certain provisions out.

Let us also clarify that in specification (3) the coefficient on the term $I\left(\sum_g D_{fg} > 0\right)$ cannot be separately identified from a firm f fixed effect, since it counts whether the firm ever donates to any grantee. Also the coefficient on the term $I\left(\sum_g CC_{frg} > 0\right)$ cannot be separately identified from a docket fixed effect, as it counts the average level of commenting by grantees for that rule (only firms commenting on the rule are included in the estimation and all grantees commenting on r are, by default, co-commenters of every firm also commenting on r). As β_2 and β_3 allow us to measure the direct effects of each element to the main interaction term $I\left(\sum_g CC_{frg} \times D_{fg} > 0\right)$, we include firm and docket fixed effects in our key specifications. We also experiment by removing each set (or both) in order to estimate these direct effects.

As in Section 5, we begin by exploring the sign and magnitude of coefficient β_1 when the donation indicator variable takes value 1 if there is any grant from f to g over our entire time period, and 0 otherwise. Table 9 reports estimates for β_1 across a set of five specifications with an incremental inclusion of firm and docket fixed effects for specification (3) in columns (1) to (4) and a specification with the continuous variable $\sum_g CC_{frg} \times D_{fg}$ in column (5). Coefficients are clustered by firm or docket, or double clustered at both levels depending on the specification. In columns (1) to (4), the increments expressed in terms of increases in units of standard deviation of similarity within each docket range from 4.5 to 23.7 percentage points, indicating that comments made by firms on rules that also received comments from their grantees appear closer in content to the final rule discussion.

As the presence of any donation over time is a less accurate indicator of a direct connection between firms and grantees than recent donations, Table 10 looks at donations that take place in either the year in which the comment is filed (year t) or in the previous fiscal year ($t-1$). In this specification, the point estimates of β_1 appear more precise and quantitatively sizable, with 0.173 of a standard deviation higher similarity for comments filed by firms with co-commenting grantees who were recipients of their donations in our preferred column (4). Similar results are obtained focusing on the intensive margin, as reported in column (5).

Table 11 further probes our results on rule-comment similarity by adding controls for the log number of pages of commentary filed on r by f which, even controlling for firm and docket fixed effects, turns out to be a strong predictor of similarity between rule discussion and comment by the firm. The effect of this control is intuitive, in the sense that carefully articulated comments may capture more of the attention of the regulator and translate in higher S_{fr}. The coefficient on $I\left(\sum_g CC_{frg} \times D_{fg} > 0\right)$ based on donations at t or $t-1$ remains positive and statistically significant in all specifications in Table 11. Contrasting these estimates with those based on the same variable constructed with donations at any time, included in columns (2) to (4), shows that the increase in similarity is driven by the co-commenting of a grantee that received a donation in the current or previous year, i.e., recent donations. When both variables (constructed with recent donations versus donations at any point in time) are included in columns (3) and (4), it is evident that recent donations carry the relevant variation.[31]

7 Getting paid not to comment: The role of hush money

Sections 4-6 focused on the role of donations from corporations to non-profits in generating additional messages that are more similar to the donor's position. In our final set of results, we examine whether corporations also use donations for a distinct strategic purpose: to silence opposing opinions. It is plausible to envision an informational lobbying environment in which agents supporting a specific action opposed by a counterparty may be motivated to suppress these opposing voices (and compensate the counterparty for its silence). For example, in a discussion of the strategies employed in the multi-year campaign of the tobacco industry Lando (1991) writes: "*The tobacco industry has been effective in purchasing what has been described as 'innocence by association'. Tobacco industry sponsorship of sports events is notorious. The industry has also contributed substantially to the arts, to women's groups, and to organizations representing minorities. These types of pernicious industry activities have been successful in buying the silence or the tacit support of some groups that have suffered a disproportionate share of the tobacco burden.*" Payment in exchange for inaction and silence is commonplace in the market (e.g. noncompete, nondisclosure agreements, nondisparagement clauses, etc.) and such private agreements or clauses do not represent *per se* invalid contracts or violations of free speech. They may be, however, private agreements that are undisclosed to regulators, who may interpret the silence of some parties

[31] In online Appendix table A.6 we also replaced similarity to the final rule discussion with indicators or log 1+ counts of the number of times that a firm is cited in the final rule discussion. We obtain similar qualitative results as in the analysis in this section. Specifically, when focusing on an indicator variable for being cited or not for a firm, our results indicate a positive but imprecise relationship when controlling for docket and firm fixed effects, but when focusing on number of times the firm is cited, the presence of recent donations to co-commenting non-profits is positive, significant at standard confidence levels, and robust to firm and docket fixed effects, and controlling for log pages of comments submitted and the any donations over time.

to the regulatory process as informative.[32]

The role of such "negative" strategies is thought to be crucial to the success of special interest groups in politics. Blocking unfavorable bills from ever seeing the light of day (or committee discharge) in the U.S. Congress is as much a part of lobbying as facilitating the passage of bills favorable to an industry. Similarly, interest group comments in rule making often involve aim to kill unfavorable provisions or stalling the implementation of rules. ("Nothing happening" is almost always the desirable policy outcome for incumbent industry, see Baumgartner et al., 2009.)

To test for the presence of "hush money" in rule making, we propose an extension of our empirical framework in Section 4. In particular, we modify the regulation specification in Section 4.2 as follows:

$$C_{gr} = \beta_0 + \beta_1 DonorComment_{gr} + \beta_2 DonorComment_{gr} \times ShareAgency_{gR} + \delta_g + \delta_r + \eta_{gr} \quad (4)$$

where $DonorComment_{gr}$ is equal to 1 if grantee g received a donation from a firm that also commented on the same regulation, and 0 otherwise. $ShareAgency_{gR}$ is the number (or share) of comments from g that are directed at rules under agency R over the entire sample. This new variable captures how common it is for grantee g to comment on rules from agency R.

To understand the intuition behind this test, observe that certain non-profits may have specific expertise or focus in a specific area of regulation, which we approximate by the identity of the agency overseeing the rule (e.g., the Sierra Club commenting on rules proposed by the EPA).[33] Interacting $ShareAgency_{gR}$ with the donation from a commenting firm, $DonorComment_{gr}$, aims to establish whether such donations have a differential effect on the likelihood of commenting for grantees that typically comment on rule considered by agency R, versus grantees that normally do not comment on rules by R. We argue that this interaction is useful for assessing the potential role of hush money, as within the set of issue experts (high $ShareAgency_{gR}$), it more likely that donations are made with the aim of inducing silence and muting commentary. A plausible null hypothesis supporting the presence of hush money is therefore $\beta_2 < 0$, as charitable donations may be more likely to be hush money for grantees that routinely comment on rules from R.

Our results based on this specification and reasoning suggest that hush money is not a common strategy in our setting. In Table 12 we present several specifications accounting for the nonlinearity in equation (4), adding increasingly conservative sets of fixed effects across the six columns. The evidence points clearly in the direction of donations increasing co-commenting from grantees that routinely comment on rules from the regulator proposing r. The coefficient $\beta_2 > 0$ is systematically

[32] Absence of a signal is in fact informative in games of incomplete information in which Bayesian rationality is assumed. For an applications to elections see Kendall et al. (2015).

[33] A similar approach was followed to define issue expertise of individual lobbyists from federal lobbying reports in Bertrand et al. (2014).

positive and highly statistically significant, indicating that firms are more likely to induce – rather than stifle – comments from such grantees. While this does not completely rule out the existence of hush money, it suggests that it is at a minimum less prevalent than the co-commenting behavior documented in Sections 4-6.

8 Concluding remarks

Politicians (and voters) are frequent targets of messages aimed at persuading them of the merits of specific policy positions. While in most cases the identity of senders is disclosed, allowing an assessment of the bias and interests of the originators of the message, in other cases it may be obscured, and deliberately so. These situations range from the use of dark money in U.S. electoral politics in the aftermath of the Supreme Court's decisions of Citizens United v. Federal Election Commission and McCutcheon v. Federal Election Commission to the circulation of white papers by think tanks and non-profits.

In such circumstances, a common trait identified by the qualitative literature reviewed in this article is the reliance on independent arms-length organizations to extend the credibility of the positions held by special interests. While in most cases such overlap of intent and opinion is genuine, one has to be careful in assessing those cases where such support is offered in close proximity to monetary donations from corporations to advocate non-profits. Such transfers, often in the form of charitable grants, are virtually undetectable by private citizens and civil servants without access to detailed tax forms. Thus, these transfers represent potential forms of distortion that cannot be weighted and assessed in decision making.

In order to provide a quantitative and systematic perspective to this issue, this paper studies the interaction of non-profit organizations and large corporations within the United States federal regulatory environment. We offer systematic empirical evidence underscoring several new findings in the literature on corporate philanthropy and special interest politics. The paper presents evidence that corporate foundations' charitable grants reach targeted non-profits just before those same non-profits engage in public commentary. The availability of a large set of public comments by non-profits and by corporations on a diverse set of rules and regulations, ranging from banking to environmental regulation, makes for a rich and virtually untapped empirical environment.

Importantly, the content of the messages simultaneously communicated by non-profits and by corporations appears systematically closer in terms of textual similarity in presence of a charitable contribution provided immediately before those comments are filed. While circumstantial, the evidence seems to point to potential concerns in the assessment of *prima facie* independent information on the part of targeted regulators, who may be unaware of the philanthropic grants that realize in the backdrop and may interpret similar comments stemming from different segment of the public spectrum as indicative of merit.

The paper also tries to address the issue of the benefits to large business interests in enlisting allied advocates who may be perceived as more balanced and less biased. We focus on textual similarity between the commenting firm and final rule discussion to gauge influence of comments over policymakers. It appears that the co-commenting patterns of firms and non-profits can offer additional visibility to the messages sent by the firms themselves measured in terms of comment similarity to the final rule or even likelihood of citation of a donor firm. As rates of return for political influence activities are extremely complex to measure, this is an area of statistical investigation requiring further study. Its exploration remains open to future empirical research.

References

Aaron, D. G., Siegel, M. B., 2017. Sponsorship of national health organizations by two major soda companies. American journal of preventive medicine 52 (1), 20–30.

Acemoglu, D., 2003. Why not a political coase theorem? social conflict, commitment, and politics. Journal of comparative economics 31 (4), 620–652.

Austen-Smith, D., 1993. Information and influence: Lobbying for agendas and votes. American Journal of Political Science, 799–833.

Austen-Smith, D., 1995. Campaign contributions and access. American Political Science Review 89 (3), 566–581.

Baron, D. P., 2001. Private politics, corporate social responsibility, and integrated strategy. Journal of Economics & Management Strategy 10 (1), 7–45.

Baumgartner, F. R., Berry, J. M., Hojnacki, M., Leech, B. L., Kimball, D. C., 2009. Lobbying and policy change: Who wins, who loses, and why. University of Chicago Press.

Bertrand, M., Bombardini, M., Fisman, R., Trebbi, F., 2018. Tax-exempt lobbying: Corporate philanthropy as a tool for political influence. Tech. rep., National Bureau of Economic Research.

Bertrand, M., Bombardini, M., Trebbi, F., 2014. Is it whom you know or what you know? an empirical assessment of the lobbying process. The American Economic Review 104 (12), 3885–3920.

Bombardini, M., Trebbi, F., 2011. Votes or money? theory and evidence from the us congress. Journal of Public Economics 95 (7), 587–611.

Bombardini, M., Trebbi, F., 2012. Competition and political organization: Together or alone in lobbying for trade policy? Journal of International Economics 87 (1), 18–26.

Brown, W. O., Helland, E., Smith, J. K., 2006. Corporate philanthropic practices. Journal of corporate finance 12 (5), 855–877.

Crawford, S. P., 2013. Captive audience: The telecom industry and monopoly power in the new gilded age. Yale University Press.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., Harshman, R., 1990. Indexing by latent semantic analysis. Journal of the American Society for Information Science 41 (6), 391–407.

Dewatripont, M., Tirole, J., 1999. Advocates. Journal of political economy 107 (1), 1–39.

Dewatripont, M., Tirole, J., 2005. Modes of communication. Journal of Political Economy 113 (6), 1217–1238.

Drutman, L., 2015. The business of America is lobbying: How corporations became politicized and politics became more corporate. Oxford University Press.

Dumais, S., Furnas, G., Landauer, T., Deerwester, S., 1988. Using latent semantic analysis to improve information retrieval. Proceedings of CHI 88 Conference on Human Factors in Computing Systems, 281–285.

Dumais, S. T., 2004. Latent semantic analysis. Annual review of information science and technology 38 (1), 188–230.

Grossman, G. M., Helpman, E., 1994. Protection for sale. The American Economic Review 84 (4), 833–850.

Grossman, G. M., Helpman, E., 2001. Special interest politics. MIT press.

Jacobson, M. F., 2005. Lifting the veil of secrecy from industry funding of nonprofit health organizations. International journal of occupational and environmental health 11 (4), 349–355.

Kang, K., 2016. Policy influence and private returns from lobbying in the energy sector. Review of Economic Studies 83 (1), 269–305.

Kang, K., You, H. Y., 2016. The value of connections in lobbying. working paper.

Kendall, C., Nannicini, T., Trebbi, F., 2015. How do voters respond to information? evidence from a randomized campaign. American Economic Review 105 (1), 322–53.

Krishna, V., Morgan, J., 2001. A model of expertise. The Quarterly Journal of Economics 116 (2), 747–775.

Laffont, J.-J., Tirole, J., 1993. A theory of incentives in procurement and regulation. MIT press.

Lando, H. A., 1991. Toward a comprehensive strategy for reducing the health burden of tobacco. British journal of addiction 86 (5), 649–652.

Lohmann, S., 1995. Information, access, and contributions: A signaling model of lobbying. Public Choice 85 (3-4), 267–284.

Lorenz, G. M., 2017. Prioritized interests: Diverse lobbying coalitions and congressional committee agenda-setting. Tech. rep., working paper.

Mayer, J., 2017. Dark money: The hidden history of the billionaires behind the rise of the radical right. Anchor Books.

Peng, V., 2016. Astroturf campaigns: Transparency in telecom merger review. University of Michigan Journal of Law Reform 49 (2).

Potters, J., Van Winden, F., 1992. Lobbying and asymmetric information. Public choice 74 (3), 269–292.

Ross, L. F., Norton, J., Young, S., et al., 2000. Is academic medicine for sale. N Engl J Med 343 (7), 508–10.

Sunstein, C. R., 2012. The office of information and regulatory affairs: myths and realities. Harv. L. Rev. 126, 1838.

Vidal, J. B. I., Draca, M., Fons-Rosen, C., 2012. Revolving door lobbyists. The American Economic Review 102 (7), 3731–3748.

Figure 1: Rulemaking process

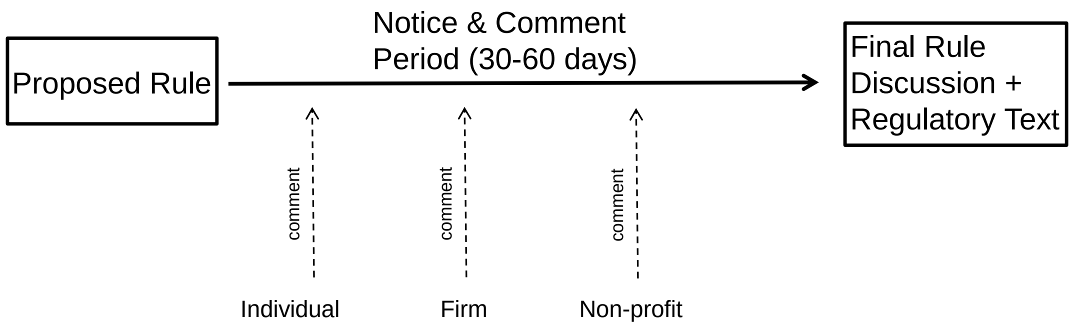

Figure 2: Co-commenting and charitable donations

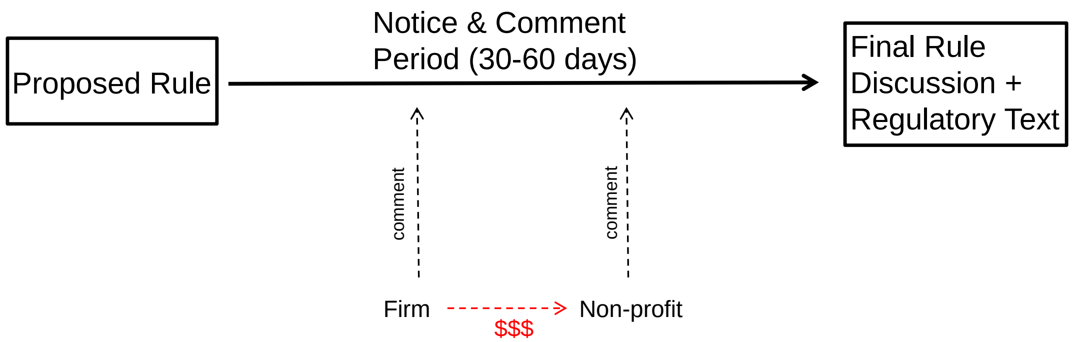

Table 1: Cross-section: co-commenting and donations

Dependent variable	Grantee g and firm f comment on same regulation×100			
Mean	2.16			
	(1)	(2)	(3)	(4)
Grantee g received donation from firm f	8.156***	6.044***	6.276***	4.033***
	(0.393)	(0.506)	(0.215)	(0.484)
Fixed Effects				
Firm		Y		Y
Grantee			Y	Y
SE Clusters	Grantee	Firm	Grantee	Firm+Grantee
Observations	11,111,716	11,111,716	11,111,716	11,111,716

Notes: The dependent variable is equal to 100 if grantee g and firm f comment on the same regulation in any year between 2003 and 2016. The independent variable is equal to one if grantee g received a donation from firm f in any year between 2003 and 2016. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 2: Panel: Co-commenting, contemporaneous and lagged donations

Dependent variable	Firm f and grantee g commented on the same regulation in year $t \times 100$				
Mean			0.163		
	(1)	(2)	(3)	(4)	(5)
Firm f contributed to grantee g in year t	0.746***	0.614***	0.587***	0.451***	-0.010
	(0.040)	(0.041)	(0.041)	(0.096)	(0.042)
Firm f contributed to grantee g in year $t-1$	0.964***	0.819***	0.798***	0.649***	0.188***
	(0.042)	(0.044)	(0.044)	(0.111)	(0.045)
Fixed effects					
Year	Y	Y	Y	Y	Y
Grantee		Y		Y	
Donor			Y	Y	
Grantee-Firm Pair					Y
SE Clusters		Grantee	Firm	Grantee +Firm	Firm×Grantee Pair
Observations	125,918,520	125,918,520	125,918,520	125,918,520	125,860,865

Note: The dependent variable is equal to 100 if grantee g and firm f comment on the same regulation in year t. The independent variable is equal to one if grantee g received a donation from firm f either at year t (respectively, $t-1$). Standard errors are clustered at the level indicated in each column under "SE Clusters".*** p<0.01, ** p<0.05, * p<0.1

Table 3: Panel: Co-commenting and contemporaneous or lagged donations

Dependent variable	Firm f and grantee g commented on the same regulation in year t				
Mean			0.163		
	(1)	(2)	(3)	(4)	(5)
Firm f contributed to grantee g in year t or $t-1$	1.153***	0.960***	0.927***	0.728***	0.132***
	(0.028)	(0.065)	(0.110)	(0.121)	(0.037)
Fixed effects					
Year	Y	Y	Y	Y	Y
Grantee		Y		Y	
Donor			Y	Y	
Grantee-Firm Pair					Y
SE Clusters		Grantee	Firm	Grantee+Firm	Firm×Grantee Pair
Observations	136,400,199	136,400,199	136,400,199	136,400,199	136,331,013

Notes: The dependent variable is equal to 100 if grantee g and firm f comment on the same regulation in year t. The independent variable is equal to one if grantee g received a donation from firm f at year t or $t-1$. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 4: Commenting on regulations

Dependent variable	Grantee g commented on regulation $r \times 100$			
Mean	0.039			
	(1)	(2)	(3)	(4)
Grantee g received donation from any firm commenting on r	0.210***	0.157***	0.181***	0.122***
	(0.012)	(0.011)	(0.010)	(0.014)
Fixed effects				
Grantee		Y		Y
Regulation			Y	Y
SE Clusters	Grantee	Grantee	Regulation	Grantee +Regulation
Observations	144,628,498	144,628,498	144,628,498	144,628,498

Notes: The dependent variable is equal to 100 if grantee r comments on regulation r. The independent variable is equal to one if grantee g received in any year 2003-2016 a donation from a firm that commented on r. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 5: Similarity - Any Donation

Dependent variable	Similarity of comments by grantee g and firm f on same regulation			
	(1)	(2)	(3)	(4)
Grantee g received donation from firm f	0.249*** (0.068)	0.161*** (0.027)	0.140** (0.065)	0.088*** (0.022)
Fixed Effects				
Docket		Y		Y
Firm			Y	Y
Grantee			Y	Y
SE Clusters	Docket	Docket	Firm+Grantee	Firm+Grantee +Docket
Observations	301,602	301,602	300,817	300,792

Notes: The dependent variable is a similarity index between the comment of firm f and the comment of grantee g on regulation r, divided by the standard deviation of similarity of all comments relative to r. The independent variable is equal to one if grantee g received a donation from firm f between 2003 and 2016. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 6: Similarity - Two years

Dependent variable	Similarity of comments by grantee g and firm f on same regulation			
	(1)	(2)	(3)	(4)
Grantee g received donation from firm f at t or $t-1$	0.172***	0.158***	0.041	0.073***
	(0.058)	(0.037)	(0.039)	(0.023)
Fixed Effects				
Docket		Y		Y
Firm			Y	Y
Grantee			Y	Y
SE Clusters	Docket	Docket	Firm+Grantee	Firm+Grantee +Docket
Observations	301,602	301,602	300,817	300,792

Notes: The dependent variable is a similarity index between the comment of firm f and the comment of grantee g on regulation r, divided by the standard deviation of similarity of all comments relative to r. The independent variable is equal to one if grantee g received a donation from firm f in the year when the comment appears or the year before. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 7: Similarity - Cont and lag

Dependent variable	Similarity of comments by grantee g and firm f on same regulation			
	(1)	(2)	(3)	(4)
Grantee g received donation from firm f at t	0.046	0.052	-0.036	0.012
	(0.046)	(0.034)	(0.035)	(0.017)
Grantee g received donation from firm f at $t-1$	0.169***	0.141***	0.085**	0.079***
	(0.054)	(0.035)	(0.039)	(0.027)
Fixed Effects				
Docket		Y		Y
Firm			Y	Y
Grantee			Y	Y
SE Clusters	Docket	Docket	Firm+Grantee	Firm+Grantee +Docket
Observations	301,602	301,602	300,817	300,792

Notes: The dependent variable is a similarity index between the comment of firm f and the comment of grantee g on regulation r, divided by the standard deviation of similarity of all comments relative to r. The independent variable is equal to one if grantee g received a donation from firm f in the year when the comment appears (respectively, the year before). Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 8: Similarity - Sector Control

Dependent variable	Similarity of comments by grantee g and firm f on same regulation			
	(1)	(2)	(3)	(4)
Grantee g received donation from firm f at t or $t-1$	0.074***	0.058***		
	(0.026)	(0.021)		
Grantee g received donation from firm f at t			0.020	0.010
			(0.023)	(0.021)
Grantee g received donation from firm f at $t-1$			0.072**	0.067**
			(0.030)	(0.030)
Fixed Effects				
Docket	Y	Y	Y	Y
Firm	Y	Y	Y	Y
Grantee	Y	Y	Y	Y
NAICS code × NTEEC code		Y		Y
SE Clusters				
Firm	Y	Y	Y	Y
Grantee	Y	Y	Y	Y
Docket	Y	Y	Y	Y
NAICS code × NTEEC code		Y		Y
Observations	162,735	162,735	162,735	162,735

Notes: The dependent variable is a similarity index between the comment of firm f and the comment of grantee g on regulation r, divided by the standard deviation of similarity of all comments relative to r. The independent variable is equal to one if grantee g received a donation from firm f in the year when the comment appears or the year before. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 9: Rule-Comment Similarity - Any Donation

Dependent variable	Similarity of rule discussion and comment by firm f on same regulation r				
	(1)	(2)	(3)	(4)	(5)
At least one grantee co-commenting and receiving donation from firm f in any year	0.237*** (0.065)	0.173*** (0.038)	0.177*** (0.059)	0.045 (0.056)	
Log number of grantees co-commenting and receiving donation from firm f in any year					0.055 (0.040)
Fixed Effects					
Docket		Y		Y	Y
Firm			Y	Y	Y
SE Clusters	Docket	Docket	Firm+Dock	Firm+Docket	Firm+Docket
Observations	5,538	5,145	5,367	4,965	4,965

Notes: The dependent variable is a similarity index between the comment of firm f and the discussion of regulation r, divided by the standard deviation of similarity of all comments relative to r and discussion of regulation r. The independent variable is equal to one if there is at least one grantee g co-commenting on regulation r and receiving a grant from firm f in any year. Column 5 reports the coefficient on the logarithm of one plus the number of such grantees. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 10: Rule-Comment Similarity - Recent Donation

Dependent variable	Similarity of rule discussion and comment by firm f on same regulation r				
	(1)	(2)	(3)	(4)	(5)
At least one grantee g co-commenting and receiving donation from firm f in year t or $t-1$	0.274***	0.261***	0.189***	0.173***	
	(0.101)	(0.052)	(0.066)	(0.062)	
Log number of grantees co-commenting and receiving donation from firm f in year t or $t-1$					0.112**
					(0.052)
Fixed Effects					
Docket		Y		Y	Y
Firm			Y	Y	Y
SE Clusters	Docket	Docket	Firm+Grantee	Firm+Docket	Firm+Docket
Observations	5,538	5,145	5,367	4,965	4,965

Notes: The dependent variable is a similarity index between the comment of firm f and the discussion of regulation r, divided by the standard deviation of similarity of all comments relative to r and discussion of regulation r. The independent variable is equal to one if there is at least one grantee g co-commenting on regulation r and receiving a grant from firm f in year t or $t-1$. Column 5 reports the coefficient on the logarithm of one plus the number of such grantees. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 11: Rule-Comment Similarity - Robustness

Dependent variable	Similarity of rule discussion and comment by firm f on same regulation r			
	(1)	(2)	(3)	(4)
At least one grantee g co-commenting and receiving donation from firm f in year t or $t-1$	0.118**		0.186***	0.131**
	(0.058)		(0.066)	(0.061)
Log number of pages of comments submitted by firm f	0.405***	0.406***		0.405***
	(0.024)	(0.024)		(0.024)
At least one grantee g co-commenting and receiving donation from firm f in any year		0.027	-0.028	-0.027
		(0.055)	(0.059)	(0.058)
Fixed Effects				
Docket	Y	Y	Y	Y
Firm	Y	Y	Y	Y
SE Clustering	Firm+Docket	Firm+Docket	Firm+Docket	Firm+Docket
Observations	4,385	4,385	4,965	4,385

Notes: The dependent variable is a similarity index between the comment of firm f and the discussion of regulation r, divided by the standard deviation of similarity of all comments relative to r and discussion of regulation r. The independent variables are the same as in tables 9 and 10. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table 12: Commenting on regulations

Dependent variable	Grantee g commented on regulation $r \times 100$					
Mean	0.039					
	(1)	(2)	(3)	(4)	(5)	(6)
$DonorComment$	0.086***	0.058***	0.042***	0.035***	0.024	-0.000
	(0.014)	(0.007)	(0.009)	(0.008)	(0.017)	(0.010)
$DonorComment_{gr}$	0.150***		0.167***		0.150***	
$\times NumberComments_{gR}$	(0.027)		(0.009)		(0.028)	
$DonorComment_{gr}$		2.560***		2.540***		2.517***
$\times ShareComments_{gR}$		(0.149)		(0.185)		(0.232)
Fixed effects						
Grantee	Y	Y			Y	Y
Regulation			Y	Y	Y	Y
SE Clusters	Grantee	Grantee	Regulation	Regulation	Grantee +Regulation	Grantee +Regulation
Observations	144,628,498	144,628,498	144,628,498	144,628,498	144,628,498	144,628,498

Notes: The dependent variable is equal to 100 if grantee r comments on regulation r. The $DonorComment_{gr}$ is equal to one if grantee g received in any year 2003-2016 a donation from firm that commented on rule r. Standard errors are clustered at the level indicated in each column under “SE Clusters”. *** p<0.01, ** p<0.05, * p<0.1

A Appendix: Regulation comments

A.1 Overview

Our data on regulatory comments comes from regulations.gov. Under the Administrative Procedures Act (APA), federal agencies must provide a means for the public to submit comments on proposed rules and other regulatory changes. Regulations.gov is a shared platform that is now used by most federal agencies to facilitate submission and public review of comments. Information about submitted comments, including the original text and attachments, can be viewed through a web browser. The site also provides an API that allows more efficient data access, particularly for collecting simple comment metadata such as the title of the comment and posted date.

Our sample starts with the the complete collection of metadata for all comments posted to regulations.gov in the years 2003-2017 (inclusive). This is a total of 6,871,697 unique documents. From these, we identify 981,232 comments that appear to be authored by organizations rather than private individuals ("org comments"). We download the complete text for all org comments using common file formats, giving us about 90% of comment text for the org comment sample.

A.2 Collecting metadata

The regulations.gov API provides a search function for document metadata. We retrieved the metadata for all public submission documents posted since the site came online in 2003, and include all years up to and including 2017. Some agencies have begun digitizing older comments and posting them to regulations.gov retroactively. But an EPA spokesperson stated (in personal email correspondence) that this work is currently incomplete, and that the text of some older comments will never be released digitally since the submitters were not aware of this possibility at the time. Thus we consider data on pre-2003 comments on *regulations.gov* unreliable and do not include them.

A.3 Identifying org comments

Authorship information can appear in three different metadata fields: "title", "organization", or "submitterName". Comments appear to fall into two main types: those that contain "organization" and/or "submitterName" information, and those that only contain authorship information in the title. First, we drop all comments that have "submitterName" information, but no organization. These appear to be written by private individuals. For the remaining comments, we look for an organization name in either the organization field or the title (if the organization field is blank). We use a custom neural network-based classifier to extract organization names from the selected field (classification is necessary for the organization field because it contains many false positives such as "self" or "none"). The classifier converts each title string to ASCII characters and predicts whether each character is part of an organization string. Contiguous chunks of characters with predicted probability greater than 0.5 are counted as organization names. The classifier is multi-layer bi-directional Gated Recurrent Unit (GRU), implemented in *PyTorch*[34]. Code is available on the Brad Hackinen's github page[35]. The classifier is trained on almost 9000 manually constructed training examples. This training set was constructed iteratively by starting with easy to parse titles, fitting the neural network, estimating the classifier's uncertainty from the total entropy of the character-level predicted probabilities, reviewing a sample of high-entropy titles, adding them to the training set, and repeating until the error rate was acceptably low. We also manually classified an additional set of 1000 random titles as a test set. The results of the test are shown below. 93% of titles are classified without error. 83% of titles with an organization are extracted exactly correctly, while 98.5% of titles with no org are extracted correctly (in other words, the classifier avoids 98.5% of false positives).

Table A.1: Organization name extraction accuracy

Sample	Count	Character Accuracy	String Accuracy
All test titles	1000	0.970	0.928
Test titles containing org	371	0.935	0.830
Test titles with no org	629	0.991	0.985

Notes:Character accuracy is the average fraction of characters classifier correctly in each title. *String accuracy* is the fraction of titles with every character correctly classified

A.4 Collecting comment text

Comments on regulations.gov can have comment text in two locations: a "text" field in the comment metadata, or in one or more attachments. The "text" field contains text that submitters have entered on a web form. It is often as brief as "see attached". Most substantial text is contained in the comment attachments where submitters can upload PDFs, word documents, other other file formats. We download all attachments of the following formats: PDF, MS Word 8, MS Word 12, and simple .txt files. The majority of attachments are in PDF format.

We use the XpdfReader *pdftotext* [36] command-line utility to extract text from most PDFs. Some PDFs contain only images of each page. In this case we must fall back on Optical Character Recognition (OCR), which we implement with a combination of *GhostScript*[37] (to render page

[34] https://pytorch.org/
[35] https://github.com/bradhackinen/subex
[36] https://www.xpdfreader.com/pdftotext-man.html
[37] https://www.ghostscript.com/

images) and *Tesseract-OCR*[38]. We use *Apache Tika*[39] to extract text from MS Word formats, and the *chardet*[40] Python package to detect formatting of simple text files. All the tools are open source.

B Appendix:Construction of comment similarity measures

In sections 5 and 6 of the paper we compare the content of firm comments with grantee comments and regulator discussion text. In the first case, our goal is to capture similarities between in the policies advocated for (or against) in by different commenters. In the second, it is to measure how much attention the regulator has paid to different comments. Complete solutions to these problems (in the sense of replicating what a literate and informed human could deduce from reading the text) are currently beyond the frontier of natural language processing (NLP) technology. Instead, we approximate these notions with a simple and robust method of text analysis called Latent Semantic Indexing (or sometimes Latent Semantic Analysis) with bag-of-words features. The basic recipe is as follows: After extracting and cleaning the comment text (to remove headers, page numbers, etc), each comment is converted into a vector of word counts. Very rare and very common words are dropped completely, and the remaining counts are weighted by a standard term-frequency-inverse-document-frequency (tf-idf) function to emphasis the words that are most useful in distinguishing between documents in each regulation. These weighted count vectors are then summarized by computing document-level singular vectors from a singular value decomposition of the feature-document matrix (this is the "latent" part of LSI, and generally improves the performance beyond using the raw feature vectors). Finally, the pairwise document similarity is computed as the cosine similarity between the document LSI vectors. The rest of this section explains these steps in greater detail, and describes a docket classification test we conducted to verify that the measure is informative.

B.1 Sample construction

We perform our analysis at the docket level. For each docket where at least one firm or one grantee comments, we load all organization comment text documents (initially treated as separate even if they are from the same author), and also discussion text from all linked rule documents. If there are at least three documents in total, we process the text and perform LSI to compute similarity measures.

[38]https://github.com/tesseract-ocr
[39]http://tika.apache.org/
[40]https://pypi.org/project/chardet/

Comment text is "cleaned" in by a Python script that attempts to identify and remove addresses and other header material that appear before the body text, tail sign-off and other material that appear after the body text, as well as repeated headers and footers (including page numbers) that appear on multiple pages. The script does not always succeed in removing the desired material (the comments are too varied in format to cover every possible case), but it is intended to remove some noise from the data.

Regulator discussion text is identified in the following way: First we load all rules that follow one or more comments in the docket (see appendix X on Federal Register document linking) and construct a separate discussion text document for each Federal Register rule document. We immediately drop Agency, Action, Dates, Summary, Addresses, Contact sections, as well as all appendices and tables of contents. Then we search for the strings "comment" and "letter" in all paragraphs and footnotes, and count a paragraph or footnote as discussion text if it appears under the same 2-level header as an instance of those strings. In other words, if the word "commenters" appears in the third paragraph under the heading "SUPPLEMENTARY INFORMATION: V. Discussion of Final Rule", every paragraph and footnote located under that heading will be included.

B.2 LSI implementation

LSI is essentially the application of singular-value decomposition (SVD) to a document-feature matrix. We follow a standard approach in constructing this document-feature matrix from word counts, and use on the excellent Gensim[41] python package for efficient implementation of these steps. First, each document is converted to lower case and words are stemmed (meaning removing common prefixes and suffixes, including pluralization so that "House" and "houses" both become "hous"). This step increasing the probability that closely-related words will be matched across documents. Next we identify every sequence of alphanumeric characters that are unbroken by white-space or other punctuation (except "-") as a word and count the number of occurrences of each word in each document. We drop all words that appear in more than 70% or less than 20% of documents (this seemingly arbitrary step is important for good results with LSI and the numbers were chosen based on experiment in a docket classification test task). Finally, we re-weight the word counts in each document using term-frequency-inverse-document frequency (tf-idf) weighting with the following formula:

$$w_{ij} = f_{ij} ln(\frac{D}{d_i})$$

where $f_{ij}$is the count of word i in document j, $d_i$is the number of documents containing word i, D is the total number of documents in the docket. The matrix of w_{ij}entries then form a $(W \times D)$ feature-document matrix M (where W is the number of distinct words).

[41]https://radimrehurek.com/gensim/index.html

Recall that SVD decomposes the matrix M into the product of three matrices: $M = U \Sigma V^*$where U is $(W \times W)$ and V is $(D \times D)$. We use an algorithm[42] that can compute the first k singular values and associated columns of U and V. If $k < min(W, D)$ then the resulting decomposition forms a rank-k approximation of M. The word "latent" in "Latent Semantic Analysis" refers to the idea that compressing the full feature-document matrix to a lower-dimensional approximation squeezes synonyms into the same singular vectors and improves overall quality of the document model. In practice, researchers have found that values of k around 200-400 appear work well in large samples of documents. However, k is bounded above by the number of separate documents D, and we have many dockets with fewer than 300 comments. As a general solution, we choose k according to the following formula:

$$k = min(D - 2, 50)$$

So the LSI vectors have higher rank in large dockets, but we keep the maximum value a bit low so that the approximations are not wildly different in dockets of different sizes. Our object of interest is the resulting $(D \times k)$ matrix V. We describe each row as a document LSI vector.

B.3 Similarity measures

Once the document LSI vectors are computed, estimating the similarity between comments from firms and grantees is straightforward. We compute organization-level vectors by summing the LSI vectors for all documents associated with that organization, and define the pairwise comment similarity as cosine similarity of the organization-level vectors.

B.4 Rule similarity

Estimating the similarity between the rule discussion and an organization's comment(s) is only slightly more complicated. In the case that there are multiple rules linked to a docket, we first construct all the comment-rule pairs and keep only those for which the comment was posted before the rule was published. Then we perform the same summing procedure to aggregate document LSI vectors associated with multiple sources of comment text submitted by the same organization, and compute similarity with the rule as the cosine similarity between the rule LSI vector and the organization-level vector.

[42]https://pypi.org/project/sparsesvd/

C Appendix: Additional tables and figures

We report here various additional figures and tables mentioned in the text.

Table A.2: Top Agencies by Number of Comments

Top 30 agencies in firms comments	Number of comments	Top 30 agencies in grantees comments	Number of comments
EPA	8099	FWS	76404
FAA	3870	NOAA	69171
FDA	1942	HHS	60969
OSHA	1245	CMS	47215
PHMSA	745	EPA	13556
NHTSA	724	ED	5105
CMS	721	FDA	4773
EERE	709	FAA	3485
DOT	541	FNS	2821
OCC	466	FSIS	2436
FMCSA	451	APHIS	2232
IRS	444	HUD	1910
NLRB	366	IRS	1733
USTR	336	CFPB	1361
CFPB	328	AMS	1310
EBSA	302	OSHA	1192
HHS	276	FHWA	1095
USCG	222	SSA	1064
FWS	208	NHTSA	1001
AMS	181	EERE	936
HUD	163	DOT	925
APHIS	152	BOEM	909
FSIS	144	ICEB	861
TSA	129	DOJ	824
FRA	109	USCG	750
FHWA	108	OMB	748
LMSO	102	FMCSA	708
BOEM	95	DOS	667
BIS	94	OPM	649
EIB	91	NLRB	616

Notes: This table reports the 30 top agencies as ranked by the number of comments they receive by firms (first two columns) or by grantees (last two columns).

Table A.3: Co-commenting in time-varying sample- Future donations

Dependent variable	Firm f and grantee g commented on the same regulation in year $t \times 100$				
Mean			0.163		
	(1)	(2)	(3)	(4)	(5)
Firm f contributed to grantee g in year $t+1$	0.557***	0.452***	0.447***	0.339***	-0.016
	(0.038)	(0.049)	(0.081)	(0.087)	(0.042)
Firm f contributed to grantee g in year t or $t-1$	0.866***	0.715***	0.699***	0.543***	0.142***
	(0.032)	(0.051)	(0.098)	(0.104)	(0.040)
Fixed effects					
Year	Y	Y	Y	Y	Y
Grantee		Y		Y	
Donor			Y	Y	
Grantee-Firm Pair					Y
SE Clusters		Grantee	Firm	Grantee +Firm	Firm×Grantee Pair
Observations	125,918,520	125,918,520	125,918,520	125,918,520	125,860,865

Notes: Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Similarity - Future Donation

Dependent variable	Similarity of comments by grantee g and firm f on same regulation					
	(1)	(2)	(3)	(4)	(5)	(6)
Grantee g received donation from firm f at $t+1$	-0.010 (0.050)	0.039 (0.030)	0.030 (0.058)	0.027 (0.036)	0.010 (0.031)	-0.030 (0.036)
Grantee g received donation from firm f at t or $t-1$						0.069*** (0.025)
Fixed Effects						
Firm	Y	Y	Y	Y	Y	Y
Grantee	Y	Y	Y	Y	Y	Y
Docket		Y		Y	Y	Y
NAICS code × NTEEC code					Y	Y
SE Clustering						
Firm	Y	Y	Y	Y	Y	Y
Grantee	Y	Y	Y	Y	Y	Y
Docket		Y		Y	Y	Y
NAICS code × NTEEC code					Y	Y
Sample with sector codes			Y	Y	Y	Y
Observations	300,817	300,792	175,660	175,643	162,735	162,735

Notes: Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Rule-Comment Similarity - Robustness

Dependent variable	Similarity of rule discussion and comment by firm f on same regulation r			
	(1)	(2)	(3)	(4)
At least one grantee g co-commenting and receiving donation from firm f in year t or $t-1$	0.106*		0.159**	0.116*
	(0.063)		(0.070)	(0.063)
Log number of pages of comments submitted by firm f	0.404***	0.405***		0.404***
	(0.025)	(0.025)		(0.025)
At least one grantee g co-commenting and receiving donation from firm f in any year		0.025	-0.070	-0.023
		(0.060)	(0.062)	(0.060)
Fixed Effects				
Docket	Y	Y	Y	Y
Firm	Y	Y	Y	Y
SE Clusters	Firm+Docket			
Observations	4,385	4,385	4,965	4,385

Notes: The dependent variable is a similarity index between the comment of firm f and the discussion of regulation r, divided by the standard deviation of similarity of all comments relative to r and discussion of regulation r. This table drops the rule within the docket that is least similar to the firm's comment. Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Rule-Comment Similarity - Robustness

Dependent variable	Citation of firm f's name in Discussion of rule r			
	Cited(Y/N) (1)	Log (1+Citations) (2)	Cited (Y/N) (3)	Log(1+Citations) (4)
At least one grantee g co-commenting and receiving donation from firm f in year t or $t-1$	0.017 (0.013)	0.043* (0.024)	0.021 (0.014)	0.059** (0.028)
At least one grantee g co-commenting and receiving donation from firm f in any year			-0.005 (0.011)	-0.032 (0.022)
Log number of pages of comments submitted by firm f			0.028*** (0.007)	0.049*** (0.013)
Fixed Effects				
Docket	Y	Y	Y	Y
Firm	Y	Y	Y	Y
SE Clusters	Firm+Docket	Firm+Docket	Firm+Docket	Firm+Docket
Observations	4,965	4,965	4,385	4,385

Note: Standard errors are clustered at the level indicated in each column under "SE Clusters". *** p<0.01, ** p<0.05, * p<0.1

Table A.7: List of Agencies on regulations.gov (A-F)

ACF	Children and Families Administration	DOI	Interior Department
AHRQ	Agency for Healthcare Research and Quality	DOJ	Justice Department
AID	Agency for International Development	DOL	Employment Standards Administration
AMS	Agricultural Marketing Service	DOS	State Department
AOA	Aging Administration	DOT	Transportation Department
APHIS	Animal and Plant Health Inspection Service	EAB	Economic Analysis Bureau
ARS	Agricultural Research Service	EAC	Election Assistance Commission
ASC	Appraisal Subcommittee	EBSA	Employee Benefits Security Administration
ATBCB	Archit. and Transportation Barriers Compliance Board	ED	Education Department
ATF	Alcohol, Tobacco, Firearms, and Explosives Bureau	EDA	Economic Development Administration
ATSDR	Agency for Toxic Substances and Disease Registry	EEOC	Equal Employment Opportunity Commission
BIA	Indian Affairs Bureau	EERE	Off. Energy Efficiency and Renewable Energy
BIS	Industry and Security Bureau	EIB	Import Export Bank of the United States
BLM	Land Management Bureau	EOIR	Executive Office for Immigration Review
BOEM	Ocean Energy Management Bureau	EPA	Environmental Protection Agency
BOP	Prisons Bureau	ESA	Employment Standards Administration
BOR	Reclamation Bureau	ETA	Employment and Training Administration
BPD	Public Debt Bureau	FAA	Federal Aviation Administration
BSEE	Safety and Environmental Enforcement Bureau	FAR	Federal Acquisition Regulation System
CCC	Commodity Credit Corporation	FBI	Federal Bureau of Investigation
CDC	Centers for Disease Control and Prevention	FCIC	Federal Crop Insurance Corporation
CDFI	Community Development Financial Institutions Fund	FDA	Food and Drug Administration
CFPB	Consumer Financial Protection Bureau	FEMA	Federal Emergency Management Agency
CMS	Centers for Medicare Medicaid Services	FFIEC	Federal Financial Institutions Exam. Council
CNCS	Corporation for National and Security Service	FHWA	Federal Highway Administration
COE	Engineers Corps	FINCEN	Financial Crimes Enforcement Network
COLC	U.S. Copyright Office, Library of Congress	FISCAL	Bureau of the Fiscal Service
CPSC	Consumer Product Safety Commission	FMCSA	Federal Motor Carrier Safety Administration
CSREES	Coop. State Research, Education, and Extension Service	FNS	Food and Nutrition Service
DARS	Defense Acquisition Regulations System	FRA	Federal Railroad Administration
DEA	Drug Enforcement Administration	FS	Fiscal Service
DHS	Homeland Security Department	FSA	Farm Service Agency
DOC	Commerce Department	FSIS	Food Safety and Inspection Service
DOD	Defense Department	FSOC	Financial Stability Oversight Council
DOE	Energy Department	FTA	Federal Transit Administration

Table A.8: List of Agencies on regulations.gov (F-Z)

FTC	Federal Trade Commission	OJP	Justice Programs Office
FWS	Fish and Wildlife Service	OMB	Management and Budget Office
GIPSA	Grain Inspection, Packers and Stockyards Adm.	ONRR	Natural Resources Revenue Office
GSA	General Services Administration	OPM	Personnel Management Office
HHS	Health and Human Services Department	OPPM	Procurement and Property Management, Office of
HHSIG	Inspector General, Health and Human Serv Dept	OSHA	Occupational Safety and Health Administration
HRSA	Health Resources and Services Administration	OSM	Surface Mining Reclamation and Enforcement Office
HUD	Housing and Urban Development Department	OTS	Thrift Supervision Office
ICEB	Immigration and Customs Enforcement Bureau	PBGC	Pension Benefit Guaranty Corporation
IHS	Indian Health Service	PCLOB	Privacy and Civil Liberties Oversight Board
IRS	Internal Revenue Service	PHMSA	Pipeline and Hazardous Materials Safety Adm.
ITA	International Trade Administration	PTO	Patent and Trademark Office
LMSO	Labor-Management Standards Office	RBS	Rural Business-Cooperative Service
MARAD	Maritime Administration	RHS	Rural Housing Service
MMS	Minerals Management Service	RITA	Research and Innovative Technology Administration
MSHA	Mine Safety and Health Administration	RUS	Rural Utilities Service
NHTSA	National Highway Traffic Safety Administration	SAMHSA	Substance Abuse and Mental Health Services Adm.
NIFA	National Institute of Food and Agriculture	SBA	Small Business Administration
NIGC	National Indian Gaming Commission	SLSDC	Saint Lawrence Seaway Development Corporation
NIH	National Institutes of Health	SSA	Social Security Administration
NIST	National Institute of Standards and Technology	TREAS	Treasury Department
NLRB	National Labor Relations Board	TSA	Transportation Security Administration
NOAA	National Oceanic and Atmospheric Administration	TTB	Alcohol and Tobacco Tax and Trade Bureau
NPS	National Park Service	USC	United States Courts
NRC	Nuclear Regulatory Commission	USCBP	U.S. Customs and Border Protection
NRCS	Natural Resources Conservation Service	USCG	Coast Guard
NSF	National Science Foundation	USCIS	U.S. Citizenship and Immigration Services
NTIA	National Telecommunications and Information Adm.	USDA	Agriculture Department
NTSB	National Transportation Safety Board	USPC	Parole Commission
OCC	Comptroller of the Currency	USTR	Trade Representative, Office of United States
OFAC	Foreign Assets Control Office	VA	Veterans Affairs Department
OFCCP	Federal Contract Compliance Programs Office	VETS	Veterans Employment and Training Service
OFPP	Federal Procurement Policy Office	WCPO	Workers Compensation Programs Office
OJJDP	Juvenile Justice and Delinquency Prevention Office	WHD	Wage and Hour Division