The Exposure Theory of Access: Why Some Firms Seek More Access to Incumbents Than Others*

February 29, 2016

Abstract

Studies of American politics consistently find little link between campaign contributions and electoral and policy outcomes, concluding that donors gain little from donating. Despite this, the donations of access-oriented interest groups continue to generate a large part of incumbents' financial advantage in U.S. legislative campaigns. We argue that we can learn directly about the motivations of interest groups, and indirectly about the possible value that they extract from incumbents, by examining differences in the degree to which they seek access. Specifically, we construct a measure of firm-level exposure to regulation using the text of over 170,000 SEC filings, and we use a variety of empirical techniques to estimate how firms' sensitivity to incumbency varies with exposure. The results indicate that firms seek more access to incumbents when they are more exposed to regulation. Exposure to the effects of policy decisions therefore appears to be an important motivator of firm contribution behavior, suggesting that firms seek access in order to influence policy, and that they benefit, or at the very least believe that they benefit, from doing so.

*Alexander Fouirnaies is a Prize Postdoctoral Research Fellow at Nuffield College, Oxford (alexander.fouirnaies@nuffield.ox.ac.uk, http://fouirnaies.com). Andrew B. Hall is an Assistant Professor in the Department of Political Science at Stanford University (andrewbhall@stanford.edu, http://www.andrewbenjaminhall.com). Authors are listed in alphabetical order and contributed equally. Hall acknowledges financial support from the Institute for Quantitative Social Science at Harvard University. The authors thank Adam Brown, Devin Caughey, Andy Eggers, Bob Erikson, Anthony Fowler, Gary King, and Jim Snyder for helpful comments and advice. Previous versions of this paper were presented at the 2013 Midwest Political Science Association Conference and the 2013 NYU-LSE Political Economy Conference. The usual disclaimer applies.

Introduction

Despite the prevailing view among American voters that there should be less money in politics,[1] empirical work on campaign contributions consistently finds that they have little or no observable strategic value. If campaign spending does not matter in general—and in particular if it does not matter for incumbents (Abramowitz 1988; Gerber 2004; Jacobson 1978; Levitt 1994)—then it is not clear that campaign contributions induce bias in the political system. Levitt (1994: 796) is especially clear on this point: "If campaign spending has little impact on election outcomes, representatives should not feel unduly influenced by PACs." In line with this view, empirical work also finds no link between donor behavior and incumbent roll-call voting. As Ansolabehere, de Figueiredo and Snyder (2003: 114) concludes, surveying this literature, "PAC contributions show relatively few effects on voting behavior."

On the other hand, incumbents enjoy a large advantage in U.S. elections (Erikson 1971; Gelman and King 1990), and they hold a large financial edge over challengers (Ansolabehere and Snyder 2000; Krasno, Green and Cowden 1994; Jacobson 2009). Roughly two-thirds of the financial advantage that incumbents possess over challengers comes from contributions from donor groups the literature identifies as "access seeking" or "access oriented" (Fouirnaies and Hall 2014). We therefore face a puzzling contradiction. Campaign spending seems to have little or no effect on either elections or roll-call voting, yet strategic donors seem to direct funds to incumbents in a systematic fashion.

We argue that we can learn about why strategic donors support incumbents by examining the variation *among* access-oriented groups in the way that they donate to candidates.[2] Why do some access groups seek more access—and in so doing, produce more of the incumbency advantage—than others? By answering this question, we will shed light on the motivations of access-oriented donors,[3] and thus, indirectly, we will learn about the value they receive, or at least believe that they receive, in return for their contributions.

[1] See for example: `http://thehill.com/blogs/ballot-box/campaign-ads/264087-poll-majority-want-corporate-money-out-of-politics`

[2] In so doing, we follow the advice of Gordon and Hafer (2005). In order "to pinpoint the place of these expenditures in American democracy...it is critical to identify the precise mechanism through which any preferential treatment might occur" (Gordon and Hafer 2005: 245)

[3] For a review of the literature on access and ideologically motivated contributors, see Barber and McCarty (2013).

Specifically, we measure the varying behavior of access-seeking firms. If strategic interest groups use access to secure favorable policy outcomes, then firms more *exposed* to policy decisions—those firms whose future profits depend most on current policy decisions—should display more access-seeking behavior. Consistent with this prediction, we show that firms exposed to more government regulation distribute their campaign contributions in a more access-seeking manner than do other firms, allocating more of their contributions based purely on incumbency status and not on the basis of ideology, district type, or other political factors. These findings are consistent in both the U.S. federal legislatures and in state legislatures, 1994–2010, and they suggest that firms benefit from seeking access to those in office when they are exposed to regulatory policy.[4]

To establish this, we rely on roughly 11 million firm-level observations on corporate contributions to U.S. legislative elections. We use the text of publicly traded firms' 10-K filings—legally required by the SEC for all publicly traded firms—to construct a scaling of firms' self-reported exposure to regulation, and we show in a series of analyses that more heavily regulated firms donate in a more access-seeking manner than do less heavily regulated firms. We employ several supplementary analyses to suggest that the link between exposure and access-seeking behavior is causal in nature, and to show that the results are robust to alternate measurement strategies.

The remainder of the paper is organized as follows. In the next section, we motivate the study. Subsequently, we lay out the empirical strategy we use to scale companies in terms of their exposure to regulation and to measure firm-level sensitivity to incumbency. Following that, we present results. Finally, we conclude.

Motivation: Why Do Firms Donate to Incumbents?

The role of "special interests" and campaign contributions in American politics is well known and much scrutinized. In the words of legal scholar Lawrence Lessig, "Americans believe that money

[4]This pattern of evidence could also be consistent with a different mechanism, in which more exposed firms support incumbents more not to gain access but to keep the their allies in office, if they feel that the incumbents with whom they have developed relationships offer better policy positions than other candidates. This alternate mechanism would not change the overall conclusion that firms benefit, or at least think they benefit, by seeking out incumbents with their contributions. However, we think this is a less likely explanation. First, we have direct evidence of firms switching between both parties in order to seek out incumbents. Unless both parties offer the same positions, the results do not seem consistent with the idea of supporting fixed allies. Second, we also investigate the results only in open-seat races, where firms cannot yet have long-term relationships with allies that they want to keep in office. We continue to find the same donation patterns, which suggests to us an access motivation. Nevertheless this is an important alternative mechanism to keep in mind.

buys results in Congress, and that business interests wield control over our legislature."[5] Despite this widespread belief, the political science literature has found a conspicuous lack of correlation between campaign contributions and political outcomes, suggesting on the whole only a murky link between the two (for a review, see: Ansolabehere, de Figueiredo and Snyder 2003; Barber and McCarty 2013).

The lack of a link between contributions and observable policy outcomes like roll-call voting may be explained by the observation that the effect of contributions on electoral outcomes, too, appears to be muted (Abramowitz 1988; Gerber 2004; Jacobson 1978; Levitt 1994). This effect seems especially small for incumbents, who already possess many of the benefits that campaign spending can provide (e.g., name recognition). If incumbents do not benefit much from campaign contributions, then they have little reason to cater to those who offer them these contributions. In turn, these results may also help explain the puzzle that interest-group donors do not appear to contribute as much money to campaigns as might be expected if they benefit from doing so. Ansolabehere, de Figueiredo and Snyder (2003) show that many PACs do not contribute up to the maximums they are allowed to, and that many firms, even large firms, do not bother to set up PACs at all. They conclude, based on these and other analyses, that "it doesn't seem accurate to view campaign contributions as a way of investing in political outcomes" (125).

In contrast to these arguments, there are other reasons to think that interest groups may benefit quite a bit from their contributions. Incumbents expend a tremendous amount of effort on fundraising (Jacobson 2009), which suggests that they at least believe contributions to be valuable. And despite the low overall amounts of money contributed, access-oriented interest groups display a marked degree of strategic sensitivity in how they allocate their contributions (Snyder 1990, 1992).[6,7] Grimmer and Powell (2013) highlights this strategic behavior. Examining U.S. House committees, they show that access-oriented donors stop contributing to incumbents who are kicked off of committees that hold jurisdiction over policy areas relevant to their business. While the total

[5] `http://republic.lessig.org/`, accessed March 27th, 2014.

[6] We focus on interest group access by means of contributions. This should not diminish the important role of lobbying in access as well (e.g., Hall and Deardorff 2006).

[7] Incumbents' need for finance and their resulting willingness to trade access to interest groups in exchange for contributions has a grounding in the theoretical literature as well. Baron (1989), Austen-Smith (1995) and Ashworth (2006), for example, present models in which candidates who need money can secure it from interest groups in exchange for promises of future favors should they win office.

size of their contributions may not be extremely large, these groups are clearly quite careful in how they allocate them to members of Congress.

A recent field experiment also supports the idea that contributions generate access. Kalla and Broockman (N.d.) establishes that legislators are more willing to hold meetings—and more willing to attend the meetings themselves—when the requesting group discloses that it is a donor. While this result does not prove the value of access, itself, it highlights a mechanism by which access can have value. Access-oriented groups gain access—in the literal form of meetings with members of Congress—by contributing.

It is possible, then, that interest groups do obtain value from their contributions, but that this value is vague and difficult to observe. The theoretical literature on access argues that "...contributors must develop a relationship of mutual trust and respect with officeholders in order to receive tangible rewards for contributions" (Snyder 1992: 17). Though they may not make exchanges of policy for contributions on a *quid pro quo* basis,[8] firms can use contributions in order to have conversations with incumbents about issues important to them. These conversations may influence policy and other activities of legislators in ways beneficial for firms, even if they do not manifest themselves in observable shifts in a variable as coarse as roll-call voting. Supporting this view, studies indicate that the stock market appears to value firms that are connected to incumbents (Gaikwad 2013; Goldman, Rocholl and So 2009).

Finally, even if we were to grant that access can be beneficial for firms, we know little about the specific factors that motivate the behavior of *individual* firms. It seems obvious that regulated industries have a vested interest in regulatory policy, but it is less clear why some firms are willing to incur the cost of seeking access while others free-ride (for a review of the literature of collective-action and lobbying see for example Barber, Pierskalla and Weschle 2014). Thus, even if access to incumbents can be helpful for influencing regulatory policy (De Figueiredo and Edwards 2007), individual firms might still choose not to contribute.

Why do some interest groups donate more strategically than others? As we show in the next section, some publicly traded corporations display far more access-seeking behavior than others. Understanding why this variation exists is important for understanding the value of access and the motivations of access-seeking groups. If, for example, some firms are simply more politically

[8]Among other reasons, the explicit exchange of contributions for roll-call votes is illegal.

engaged than others due to the idiosyncratic preferences of their management, this variation in access-seeking behavior may be of little import. If the decision to seek access seems "random" in this sense, it may suggest that firms have little to gain from access, or at least that firms disagree over the value of access. If, on the other hand, this behavior varies systematically with firm conditions, it strongly suggests that firms believe their donations to be valuable. Furthermore, if we view firms as profit-maximizing entities who do not deploy their money without careful reason, then the fact that they donate in this manner also suggests their contributions are indeed valuable. Of course, this latter conclusion is more tentative; we will not offer any direct evidence for the value of access. But our indirect evidence, via the behavior of firms, will be suggestive.

In this paper, we show that some firms perceive themselves to be more *exposed* to policy than others.[9] Some firms, for example, operate in industries in which the decisions the government makes play a large role in determining business outcomes—e.g., energy firms exposed to the policy decisions local governments make over price controls, the construction of power infrastructure, etc. Other firms, on the other hand, operate in sectors relatively independent from government policy—e.g., online retailers who operate in some ways beyond the reach of legislation. If access is valuable for influencing policy, then more exposed firms should seek more access than should less exposed firms. We find consistent evidence in both the U.S. federal and state legislatures for this "exposure theory of access."

Empirical Strategy

Using Text To Measure Firm-Level Exposure To Regulation

To measure the degree to which firms are exposed to regulation, we rely on the text of publicly traded firms' annual 10-K filings with the SEC. The 10-K filing provides overall information on the firm and the market in which it operates, including statistical information about its revenues and profits as well as discussions of the firm's outlook and performance over the past year.

Ours is, to our knowledge, the first paper that systematically uses 10-K filings to measure regulatory risks, but the idea of studying managers' perceptions using the text from 10-K files has

[9]This exposure goes in both directions. Firms with incentives to care about policy may seek out incumbents in order to obtain the access they need, but incumbents, too, may seek out firms knowing that these firms rely on the decisions the government will make.

long been exploited by scholars in finance and accounting. For example, text-based analyses of 10-K files have been used to study managers' perception of competition (Li, Lundholm and Minnis 2013; Shi and Zhang 2014), customer satisfaction (Balvers, Gaski and McDonald 2012), litigation risks (Rogers, Van Buskirk and Zechman 2011), environmental risks (Doran and Quinn 2009) and ethical responsibilities (Loughran, McDonald and Yun 2009).

An important legal aspect of the 10-K filings is that managers may be personally exposed to substantial litigation costs if they fail to disclose bad news in a timely manner (Campbell et al. 2014; Skinner 1994). As a consequence, it is a common belief in this literature that the 10-K filings accurately reflect the risks faced by the firms. This view is also supported by recent research. In a comprehensive study that examines the risks described in more than 30,000 filings from 2005-2008, Campbell et al. (2014: p. 396) conclude:

> ...we find that firms facing greater risk disclose more risk factors, and that the type of risk the firm faces determines whether it devotes a greater portion of its disclosures towards describing that risk type. That is, managers provide risk factor disclosures that meaningfully reflect the risks they face.

Although the previous literature is not devoted to regulatory risks, several studies note that government regulation is a common risk factor discussed in the 10-K files (Campbell et al. 2014; Kravet and Muslu 2013). Simple descriptive facts support this notion: Approximately 57% of 10-K files submitted by firms publicly traded on either the New York Stock Exchange, the NASDAQ or NYSE MKT mention at least one of the following word stems: 'state law', 'state leg', 'state reg', 'state agenc' or 'state gov'. For federal word stems, the pattern is almost exactly the same.[10]

To scale the firms, our operating assumption is, in line with the view in the finance and accounting literatures, that firms that discuss the regulatory environment more in their 10-K filings perceive themselves as more exposed to regulation. The resulting scaling is thus a direct measure of firms' *perceptions* toward their regulatory risk, and indirectly a measure of actual exposure, to the extent that firms' perception accords with reality. This perception, and the measure, may reflect exposure to existing regulation or to the threat of impending changes to the regulatory environment.

[10]For further details, see Table A.15 in the Appendix.

The text analysis is based on the list of 10-K files used by Loughran and McDonald (2011). These filings are available for the years 1994–2010 on the SEC's online repository, EDGAR. For each filing, we recorded the frequency of the following word stems: government; federal; congress; senat; governor; agency; court; administration; commission; legislat; polic; rule; politic; penalt; fine; law; regulat; zoning; licens; oversight; compliance; enforce; require; pursuant; protect.[11] For each document we thus have a vector of word counts, or combining these vectors together, a document-term matrix.[12] What exactly is written in these filings to indicate exposure to regulation? As an example, consider a passage from the 10-K filing of Wells Fargo in 2008—a firm-year observation coded as being one of the most exposed to regulation in our state legislative dataset. They write:

> "Banking statutes, regulations and policies are continually under review by Congress and state legislatures and federal and state regulatory agencies, and a change in them, including changes in how they are interpreted or implemented, could have a material effect on our business."[13]

The highlighted words are those included in the scaling procedure. Passages like this—which use many of the words in our regulatory vocabulary list—indicate that the firm perceives itself exposed to regulation, both to existing regulation and to the possibility of new or revised regulation in the future.

To create a simple scaling of firms reflecting their self-reported exposure to regulation by year, we extract the first principal component of the document-term matrix.[14,15] The "loadings" on the first dimension, the coefficients that reflect how much each word's frequency contributes to the scaling (or "score") each document receives, help indicate whether the scaling we extract is

[11] This word list is a combination of "regulatory words" used in the finance and accounting literature on 10-K filings(Campbell et al. 2014), and an additional list of words that we developed ourselves after reading selections from the Federal Register as well as SEC filings. The scaling, as well as the findings, are not sensitive to the exact set of words used. Previous versions of this paper also included the word "act" in the scaling; however, we later found that, because almost all 10-Ks make frequent reference to laws governing their reporting requirements (e.g., the SEC Act), this word appears very frequently and does not distinguish documents well. We therefore removed it.

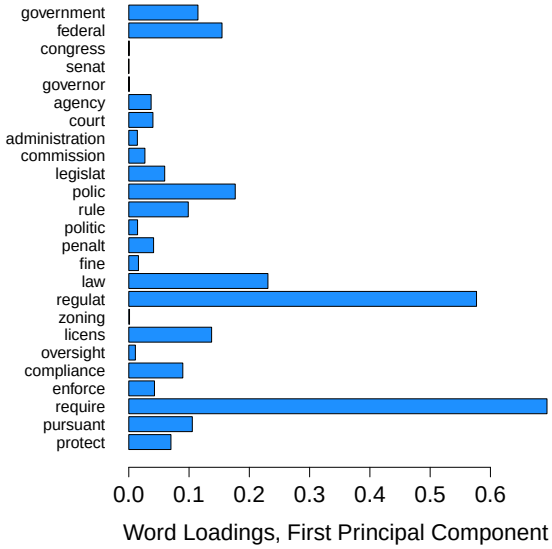
[12] We have chosen not to normalize documents by total word length because concerns over regulatory exposure might produce longer documents; however, normalizing by length in reality does not alter the scalings much if at all. We have also re-run the analysis using an alternative scaling in which we record only a binary indicator for the presence of each word, rather than the frequency. Results are highly similar.

[13] `http://www.sec.gov/Archives/edgar/data/72971/000095013408003822/f38267e10vk.htm`

[14] For an overview of principal component analysis, see for example Jolliffe (2005). For examples of its use in political science, see: Spirling (2012); Wiesehomeier and Benoit (2009).

[15] R code to produce the scaling is provided in Appendix A.4.

Figure 1 – Construction of the Exposure to Regulation Measure. Shows the coefficients on each word stem in constructing the scaling, the "loadings" in the first principal component. All regulatory word stems, including "regulat," enter positively, indicating that larger values on the scale reflect higher degrees of exposure to regulation.



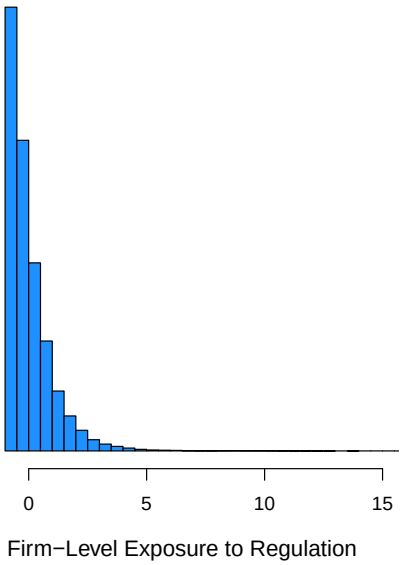
meaningful. As Figure 1 shows, the loadings are uniformly positive, suggesting that larger scores in the scale indicate more exposure to regulation.

Figure 2 presents the distribution of the exposure to regulation measure. We normalize the measure have mean 0 and standard deviation 1. The distribution has a strong skew, with a tail of positive, heavily exposed outliers. Though we include these observations in the main analyses, results are robust to their exclusion.[16]

Who are the firms most and least exposed to regulation, according to this measure? Figure 3 plots the exposure measure by industry, using two-digit SIC codes to define industries. Industries are sorted from most exposed (top of the graph) to least exposed (bottom of the graph). The most exposed industries are those related to energy—both gas and coal mining—and depository institutions, which includes a large swath of the financial industry. The least exposed industries include a variety of merchandise-focused industries as well as non-depository credit institutions, a category that includes the vast majority of so-called hedge funds, entities that are subject to very little regulation. The rankings thus seem to accord with broad notions of which industries are most

[16]Specifically, we re-run the main analysis excluding firm-years where the scaling exceed 0.4. See Appendix Table A.4.

Figure 2 – Distribution of the Exposure to Regulation Measure, 1994–2010. Publicly traded firms are scaled on the basis of word frequencies in their yearly 10-K filings with the SEC. Scaling is calculated as the first principal component of the document-word matrix. Figure plots distribution of average firm scaling across years.



regulated. Next, Figure 4 plots individual firms according to their estimated exposure. Because we cannot fit all of the firms into a single graph, we focus only on the firms which contribute to federal campaigns. To further make the plot legible, we arbitrarily emphasize the text of every fifth firm's name, as well as the most and least exposed firms. Consistent with the industry-level plot, we see large energy firms (like Exelon and PG&E) to the right of the plot, i.e., more exposed, while a variety of general service corporations, like Darden Restaurants and Avon Products, are to the left, i.e., less exposed.

In the Appendix, we carry out a number of analyses that suggest ours is a valid measure of regulatory exposure (see section A.3). We investigate variation in the measure over time and across industries, we correlate the measure with an existing measure of regulatory burden, and we attempt to map changes in the regulatory environment to changes in the scaling where possible. Since we use this exposure measure as an explanatory variable, random measurement error in the scaling will bias us away from detecting differences between the access-seeking behavior of more and less exposed firms. Nevertheless, we are careful to present results in a variety of ways to ensure that our findings are not driven by idiosyncratic features or systematic measurement error in this scaling.

Figure 3 – **Industry-Level Exposure to Regulation.** Industries at the two-digit SIC level are scaled in terms of their exposure to regulation, using the text of their 10-K filings with the SEC. Energy, mining, and financial institutions that hold deposits rank among the most regulated industries according to the measure.

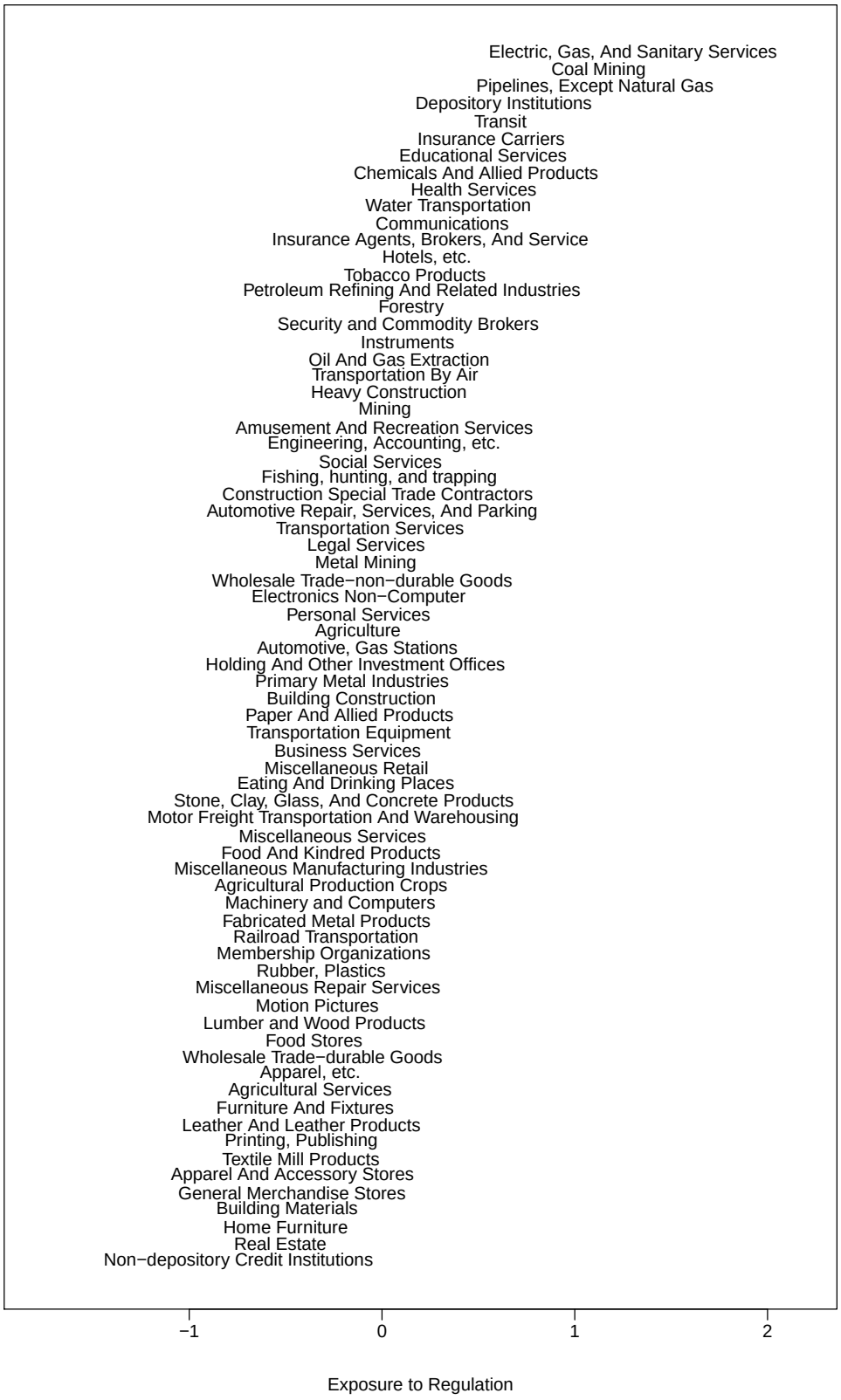


Figure 4 – **Firm-Level Exposure to Regulation.** Firms are scaled in terms of their exposure to regulation, using the text of their 10-K filings with the SEC. Due to the large number of firms, we only graph those who contribute at the federal level. To make the plot legible, only every fifth firm, as well as the max and min, are emphasized.

Measuring Access-Seeking Behavior: The Financial Incumbency Advantage

With our measure of perceived exposure to regulation in hand, we need to measure the degree to which firms seek access to incumbents. To do so, we follow the logic of Fouirnaies and Hall (2014), calculating the "financial incumbency advantage" for each firm. This advantage reflects the degree to which incumbents receive extra money from donors purely by virtue of their incumbency status, separate from partisanship, ideology, or other factors that might attract donors. Specifically, at the firm level, this advantage measures how much extra money firm j allocates to candidates purely because they are incumbents, and thus reflects how much firm j values access to those in office.

Firm-Specific Estimates of the Financial Incumbency Advantage

We begin by estimating firm-specific estimates of the financial incumbency advantage in order to see variation in the degree to which firms seek access. The purpose of this exercise is purely descriptive; by seeing the variation in the firm-level effects, we can motivate the need to identify systematic features that lead some firms to seek more access to incumbents than others.[17]

To estimate firm-level effects, we calculate the total amount of money that flows to the Democratic candidate in each election from every firm. For the federal legislatures, we obtained data from the Federal Election Commission (FEC) on contributions from corporate PACs to candidates running for federal offices.[18] For state legislatures, these contribution records are provided by the National Institute on Money in State Politics.[19] We merge these firm-election contribution records with the election data. For the U.S. House and Senate, the election data was compiled from primary sources for a series of papers by Ansolabehere et al. (2010). For state legislatures, we utilize ICPSR dataset 34297 (Klarner et al. 2013). Because the scalings are available for 1994–2010, this is the year-range for which we analyze elections and contributions.[20]

[17] Because the goal is descriptive, we will not perform any formal statistical tests on individual firm effects. Doing so would lead to inevitable issues related to multiple testing; obviously, there will be considerable variation in firm-level effects due purely to chance. The goal of the main analyses in the paper, after this section, will be to uncover systematic (that is, not due to chance) variation in these effects related to concerns about regulatory exposure.

[18] All files can be downloaded at the FEC's website: `http://www.fec.gov/finance/disclosure/ftpdet.shtml`

[19] Available at `http://www.followthemoney.org`.

[20] The availability of data varies for the state legislatures due to differences over time and across states in the laws governing what types of donors can contribute to elections. In the Appendix, we provide detailed tables that show the exact states, years, and offices for which observations enter the analysis. See Tables A.18 and A.19.

The donation records do not list donations of "0" from firms that do not donate to a given campaign. Omitting firm-election observations whenever a firm did not contribute would induce a selection bias into the analysis, since it would involve selecting on the dependent variable. To address this, we construct a list of every firm that occurs anywhere in the donation data and then insert a total of 0 for that firm in every election in which it does not contribute, producing a large and sparse dataset containing all firm-election pairs. Because the resulting state level dataset is massive, we restrict attention at the state level to donors that contribute at least a total of $10,000 over all years in the sample. The findings are thus *local* to the types of firms that donate to elections, but are not biased within this sample by selecting on the occasions on which particular firms choose to participate.[21]

To obtain a firm-specific estimate of the financial incumbency advantage, we estimate models of the form

$$\log(Dem\ Money_{ij,t+1} + 1) = \beta_0 + \beta_{1j} Dem\ Win_{it} + X_{it} + \epsilon_{ij,t+1} \quad (1)$$

where $\log(Dem\ Money_{ij,t+1}+1)$ measures the amount of money donated by firm j to the Democratic candidate in district i in the election at time $t+1$. In the Appendix, we show that results are substantively unchanged if we use $\log(Dem\ Money_{ij,t+1} + 1000)$ as the outcome variable.[22] The variable $Dem\ Win_{it}$ is an indicator for a Democratic victory in district i in election t. The coefficient on this variable, β_{1j}, thus indicates an estimate of the financial incumbency advantage for firm j. The variable X_{it} stands in for a possible vector of controls. Throughout this analysis and the rest of the paper, we present results using: no controls; controlling for district ideology (using the presidential vote share), candidate quality (using previous office-holding experience as proposed by Jacobson (2009)) and open seats; using year and district dummies to perform a difference-in-

[21] This approach is important for removing selection bias, but it also deflates estimates. Since there are many elections in which any firm j does not participate—either because the company does not operate in the area in which the election is held, or because the office up for election is deemed unimportant for that firm's business, or for any other reason—there are a large number of treated and control cases with outcomes of 0 for each firm, thus moving the effect towards zero. In principle we could attempt to address for this downward bias by identifying a set of elections *ex ante* as "potential" races for a given firm, and only use those elections in estimating the effect for that firm. However, because it is difficult to select a principled procedure by which to identify these elections, we prefer the conservative approach in which we include all races. Thus effects should be considered in terms of relative values, and not in absolute sizes. Given that firms contribute large amounts of money, in total, to elections, their relative behavior is likely to matter substantively.

[22] In Table A.6, we also re-estimate the results using per-capita contributions. Results are highly similar.

differences; and finally, including the Democratic candidate's vote-share winning margin in various specifications in order to perform a regression discontinuity (RD) analysis.

This final approach mirrors that in Fouirnaies and Hall (2014). The RD addresses potential biases from unobserved differences across districts by focusing on close elections in which the winning party is "as-if" randomly assigned. By comparing how firms contribute to the Democratic candidate in the next election after a bare Democratic victory[23] to how they contribute to the Democratic candidate in the next election after a bare Democratic loss, we can estimate the causal effect of incumbency on firm-level campaign contributions under weak assumptions.[24] In the Appendix, we validate these assumptions by showing that treated and control units are balanced at the discontinuity in terms of the lagged outcome variables (Eggers et al. 2015).[25] While the RD has the benefit of addressing potential issues of bias, it is quite local to the set of races in which it can be implemented—namely, close races. These will occur in more competitive districts and may be contexts in which contribution behavior is different. By comparing the RD results to those from the pooled and diff-in-diff analyses, we will ensure that our conclusions are not overly narrow.

To start, we apply this technique to all 362 firms in the state legislative dataset and all 664 firms in the federal election dataset in order to descriptively characterize the variation across firms. Figures 5, 6, and 7 plot the firm-specific financial incumbency advantage estimates along with 95% confidence intervals from robust standard errors clustered by election, using equation 1 with no controls.

[23] We focus arbitrarily on Democratic candidates. Of course we could examine the opposite treatment for the same set of elections, where we consider Republican victories and defeats. Results are substantively identical in this setup.

[24] For an in-depth treatment of these assumptions, see Lee (2008) and Imbens and Lemieux (2008). For concerns about these assumptions in the U.S. House, see Caughey and Sekhon (2011) as well as Grimmer et al. (2012) and Snyder (2005). However, for evidence that the assumptions are generally plausible and are likely to hold even in the U.S. House, see Eggers et al. (2015). Finally, for detailed validity tests for the RD in the context of state legislative elections and contributions, see Fouirnaies and Hall (2014).

[25] Due to its reliance on close elections, the RD estimate is inherently *local*. This limits the conclusions we can draw; we cannot directly assess how firms seek access to incumbents in safe districts. However, Hainmueller, Hall and Snyder (N.d.) show that the RD incumbency advantage estimate is surprisingly externally valid for U.S. statewide elections. In addition, there is simply no way to measure access-seeking behavior where there is little variation in incumbency status. Simply tabulating the contributions of firms in these places will mix access-seeking behavior with other underlying factors that make for safe districts and interest group contributions, like the partisanship of the district, the quality of the incumbents there and so forth. Regardless, access in competitive districts is likely to be important in and of itself. Competitive districts receive the lion's share of all campaign spending (this fact is easily verified using data on elections and campaign spending available from the FEC; see for example Snyder (1992)) and incumbents in competitive districts are likely to be the ones exerting the most effort to raise funds.

As the plots show, there is substantial variation in the financial incumbency advantage across firms. Consider Figure 5, showing the firm-level effects for state legislative elections. A number of large positive outliers are apparent; AT&T and Verizon are among the largest. Other firms appear far less sensitive to incumbency; Yahoo!, for example, has an effect close to zero. Similar patterns are shown for federal elections in Figure 6. AT&T and Verizon are again large positive outliers, as is, for example, Comcast.[26]

What explains this variation? The remainder of the paper explores one systematic factor underlying this heterogeneity in firm-level sensitivity to incumbency.

Examining Variation in the Effect Across Exposure to Regulation

To investigate what factors influence the decision to seek access, we compare these firm-level effects across firm characteristics. In particular, we can examine how the effect varies with firm-level exposure to regulation, as well as with industry-level regulatory constraints.

Before estimating these relationship formally, we investigate them graphically in Figure 8. The figure presents the correlation between firm-level exposure to regulation and firm-level contributions made to incumbents vs. non-incumbents. Because there are far too many observations to graph comfortably, the plots instead present binned averages, where each point represents an average calculated within an equal sample-size bin of the exposure to regulation variable. In the left panel, we see a positive association between exposure and the incumbency "premium"—i.e., the difference in log total firm contributions made to incumbents vs. to challengers. The same pattern is present in the right panel for the U.S. House and Senate. Although this graphical evidence is only speculative, it accurately foreshadows the formal analyses we turn to next.

Formally, we estimate models of the form

$$\begin{aligned} \log(Dem\ Money_{ij,t+1} + 1) &= \beta_0 + \beta_1 Dem\ Win_{it} + \beta_2 Exposure_{jt} \\ &+ \beta_3 Dem\ Win_{it} \times Exposure_{jt} + X_{it} + \epsilon_{ij,t+1} \end{aligned} \tag{2}$$

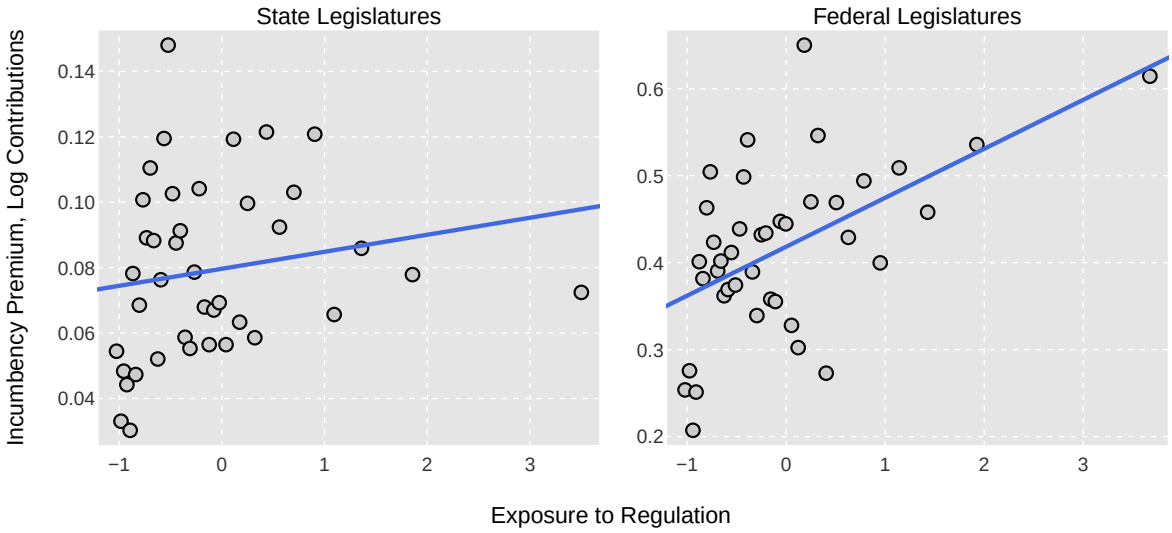
[26] These very large effects primarily reflect that the firms in question almost exclusively donate to incumbents. For example, on average Comcast donated approximately 2,879 dollars to incumbents but only 65 dollars to challengers.

Figure 5 – Firm-Level Impact of Incumbency on Contributions, U.S. State Legislatures, 1994–2010. Plots firm-specific estimates of the effect of Democratic incumbency on subsequent logged contributions, as in equation 1.

Figure 6 – Firm-Level Impact of Incumbency on Contributions, U.S. House and Senate Firms A–L, 1994–2010. Plots firm-specific estimates of the effect of Democratic incumbency on subsequent logged contributions, as in equation 1.

Figure 7 – Firm-Level Impact of Incumbency on Contributions, U.S. House and Senate Firms L–Z, 1994–2010. Plots firm-specific estimates of the effect of Democratic incumbency on subsequent logged contributions, as in equation 1.

Figure 8 – Exposure to Regulation and Firm Contributions to Incumbents and Non-Incumbents. Regulated firms are more sensitive to incumbency.



Note: Points represent averages in equal-sample-sized bins of the exposure to regulation variable. Lines are simple OLS predictions from a regression fitted to the binned points.

where $Exposure_{jt}$ measures the exposure to regulation of firm j in year t and all other variables are defined as before. The coefficient β_3 measures our main quantity of interest—the difference in the financial incumbency advantage estimate across levels of firm exposure.[27]

In the simplest case with no controls (when X_{it} is empty), the analysis calculates the average difference in contributions each firm makes to incumbents and non-incumbents, respectively, and looks at how this relationship varies across levels of exposure. This is both the simplest approach and the most powerful, in the statistical sense, since it uses the entire dataset. To ensure that it is not biased by unobserved heterogeneity across districts and candidates, we also carry out the techniques discussed above, in which we either rely on controlling for district characteristics, or on a regression discontinuity design which ensures causal identification but at the cost of focusing the analysis on a particular set of close elections. In practice, we find extremely consistent results across all specification choices, as we show now in the next section.

Because *Exposure* is inevitably measured with error, performing inference for regressions of this type is difficult—a ubiquitous problem in any research that uses a measure as a right-hand side

[27] Because there are many firm observations for each "treatment" (election), we cluster all standard errors by election. In addition, in the Appendix we present results using two-way clustering by election and firm. Some specifications become noisier with two-way clustering; however, our main specification, as well as the RD specification, remain highly statistically significant.

variable. Ideally, we would perform a non-parametric bootstrap where we sample from the text files with replacement, calculating the exposure scaling each time and then re-estimating the regression, thereby obtaining a full measure of our uncertainty over the regression coefficients that includes both the sampling variation in the regression and in the scaling measure itself. Unfortunately, the scale of our data makes this impossible; by our calculations, such a bootstrap would take over a year to complete. However, the very fact that our dataset is so large provides some reassurance that failing to incorporate this additional source of variation is not overly problematic; because we have so much text, our scaling is likely to be highly statistically precise. The bigger concern, of course, is not that the measure is noisy due to sampling variation in word frequencies but rather due to a loose connection to the underlying concept of regulatory exposure. We do as much as we can to address this deeper concern in the Appendix, where we validate the measure.

Results: Regulatory Exposure Predicts Access-Seeking Contributions

More Regulated Firms Seek More Access

In Table 1 we investigate access-seeking behavior in the U.S. House and Senate, using the specification discussed above. In the first column, we present the simple pooled OLS specification in which we examine average differences in the amount of contributions firms give to incumbent and non-incumbent candidates across firm-level exposure to regulation. In the second column, we control for district-level presidential vote share, candidate quality and open seats (and interactions with the exposure measure) to account for differences across districts. In the third column, we add district and year fixed effects, making the analysis a difference-in-differences design in which changes in incumbency status are used to estimate the access-seeking behavior of firms. Finally, in the fourth column we use a regression discontinuity design.

The quantity of interest is estimated in the second row, "Dem Win $\times$ Exposure." Since the exposure variable is scaled to have mean 0 and standard deviation 1, this interaction term represents the difference in the effect for a one standard-deviation increase in exposure to regulation.

Table 1 – Federal Legislatures: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.418	0.485	0.412	0.647
	(0.019)	(0.022)	(0.026)	(0.129)
Dem Win × Exposure	0.057	0.040	0.025	0.078
	(0.008)	(0.016)	(0.006)	(0.025)
Exposure	-0.003	-0.081	-0.010	-0.001
	(0.005)	(0.026)	(0.004)	(0.007)
Constant	0.081	0.339	0.002	0.085
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	753,928	711,894	753,928	126,390
Number of Elections	1,553	1,467	1,553	280

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Presidential vote share and the interaction with Exposure, Democratic incumbency and the interaction with Exposure, Democratic candidate quality and the interaction with Exposure.

As the table shows, there is a marked increase in the sensitivity to incumbency for more exposed firms. Consider the first column. A Democratic victory is estimated to increase log Democratic donations in the subsequent race from the firm with the average level of exposure—measured by the estimate in the first row, when the interaction is zero—by 0.418 log points. For a firm with a level of exposure one standard deviation above the mean, though, the effect is 0.475 log points $(0.418 + 0.057 = 0.475)$.[28] Though these numbers are not immediately interpretable—because the sparsity of firm contributions deflates all the estimates—proportionally they are large. The effect for a firm with exposure one standard deviation above average is almost 14% larger than the effect at the mean.

The positive association between exposure and access-seeking contribution behavior is robust across the various empirical specifications we employ. The first two columns use a much larger set

[28] Note that we do not need to consider the coefficient on the main effect of *Exposure* in calculating these marginal effects. The RD effect at $Exposure = 1$, for example, is the RD difference in contributions for firms at the maximum level of exposure when the Democrat wins and when the Democrat loses. Since in both of these cases $Exposure = 1$, the coefficient on *Exposure* differences out, and likewise for computing the RD effect at $Exposure = 0$. Thus this effect also differences out when then comparing these differences to see how the effect changes across levels of exposure.

Table 2 – State Legislatures: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.080	0.079	0.054	0.056
	(0.001)	(0.002)	(0.001)	(0.005)
Dem Win × Exposure	0.006	0.005	0.004	0.003
	(0.001)	(0.001)	(0.000)	(0.001)
Exposure	-0.001	-0.000	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.004	0.004	0.001	0.013
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	8,238,340	8,192,487	8,238,340	1,183,089
Number of Elections	29,835	29,673	29,835	4,316

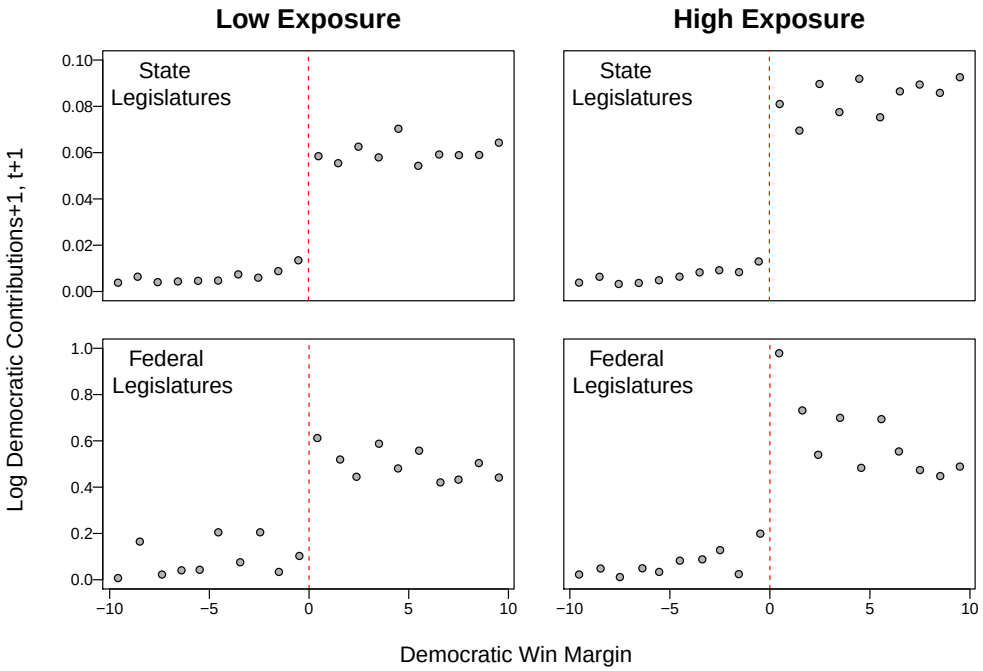
Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Democratic incumbency and the interaction with Exposure.

of elections, but are vulnerable to bias from the unobserved heterogeneity across districts. The diff-in-diff in the third column and, to a larger extent, the RD in the fourth column, address this source of bias but at the cost of narrowing the focus of the analysis (by leveraging districts that switch parties, in the diff-in-diff, and by focusing on close elections, in the case of the RD). That the results are consistent across these approaches suggests that neither the omitted variable bias nor the locality of the estimates are serious problems.

We uncover the same pattern of evidence in state legislatures, too. Table 2 presents the same analyses for the state legislative data. Again, we see a large and positive coefficient for the interaction term in the second row, indicating that firms more exposed to regulation contribute more to incumbents in state legislatures. These results, too, are consistent across empirical strategies.

It is also instructive to view these results graphically. We present RD plots to give a graphical sense of the effect in Figure 9. For state legislatures and the federal legislatures, respectively, we constructed plots comparing the Democratic vote-share winning margin and subsequent average donations to the Democratic candidate for two sets of firms: those in the lowest quartile in terms of their exposure to regulation, and those in the highest quartile. As the plots show, the "jump" at

Figure 9 – Impact of Incumbency on Firm Donations to Democrats for Firms with Low and High Exposure to Regulation, 1994–2010. Compares the RD "jump" in the least-exposed firms to that in the most-exposed firms. The jump is larger for the most-exposed firms. Points are averages in 1% bins of the running variable. "Low" and "high" exposure firms are those in the lowest and highest quartile of the exposure measure, respectively.



the discontinuity—measuring the sensitivity of firms to incumbency—is larger for the more exposed firms than for the less exposed firms.

Consider the results for state legislatures in Figure 9. In the upper left panel, Democratic incumbency appears to cause approximately a 0.04 point jump in log contributions to the Democratic candidate in the subsequent electoral cycle from firms with low exposure to regulation. In the upper right panel, we see that this same jump is approximately 0.06 points for high exposure firms—significantly larger than for low exposure firms. The results are similar for the federal legislatures in the second row. As the latter plot shows, the results have more noise for the federal legislatures because there are far fewer elections at the federal level than in all state legislatures. Nonetheless, the jump clearly grows for high exposure firms relative to low exposure firms.

In the previous section, we documented substantial variation in the degree to which different "access-seeking" firms are sensitive to incumbency. This variance suggests that some firms seek more access than others. In this analysis, we have identified one systematic component of this variation. Firms that spend more time discussing their exposure to regulation in their SEC filings

display significantly more sensitivity to incumbency than firms that spend less time discussing this exposure.

One possibility, however, is that firms more exposed to regulation are also firms who contribute to campaigns in a different way systematically. For example, exposure to regulation might motivate employees to contribute more to their firms' PACs, giving exposed firms larger political budgets to work with. Thus the observed contribution behavior might have more to do with employee sentiment than with a desire for access *per se*. To address this and similar possibilities, we have re-estimated the same equations but with a control for total contributions made over the PAC's entire lifetime within the sample. We continue to find the same link between exposure and access-seeking behavior even when we make comparisons only among firms with similar PAC budgets (see the Appendix for the results).

Another potential issue is about the supply vs. demand of contributions. Firms might seek out politicians in order to obtain access, but politicians, too, might seek out firms for contributions. This latter mechanism might be especially prevalent once legislators have gained connections and grown their network in Washington DC. To test for it, we re-estimate the federal legislature results using only open-seat races–races after which new incumbents must raise money before they have this extensive network in place.[29] We continue to find the same pattern of results, suggesting that the contributions we observe are not driven by the growth of the these networks in Washington DC.

Thus far, we have seen associational evidence that firms who perceive themselves to be more exposed to regulation are more aggressive in seeking access to incumbents through their contribution behavior. In turn, this suggests that access can convey benefits associated with regulatory policy. Though we have just ruled out two other explanations for this observed pattern, we turn now to a more thorough analysis of alternative explanations. After rejecting these possibilities, we conclude that exposure to regulation induces more access-seeking behavior.

Further Evidence: More Regulated Industries Seek More Access

Does exposure to regulation *cause* firms to seek more access? Though suggestive, one concern with the estimates above is that the scaling method depends on self-reported exposure to regulation.

[29]These results are presented in the Appendix.

Other factors that lead some firms to write more about exposure in their 10-Ks might influence their sensitivity to incumbency, changing the interpretation of the large interactive effects we have uncovered. For example, a firm with more politically engaged upper management may donate more strategically to incumbents but may also be more likely to write about political issues in their 10K filings. Alternatively, firms with larger operating budgets might also write more about regulatory exposure and simultaneously be able to seek more access to incumbents.

To address concerns like these, we use the 10-K filings of firms that never contribute to any elections in our sample to account for the idiosyncratic writing styles of those firms that do enter the sample. Specifically, we construct an average text-based exposure scaling for each industry using only the text of firms that never contribute, and then we use those to instrument for the scalings of the firms in our sample. For these purposes, we define each firm's industry as the firm's Standard Industrial Code (SIC), and we define non-contributing firms to be firms that *never* contribute to any election in any year in our sample.

We instrument for *Exposure* and *Dem Win* × *Exposure* in equation 2 using the exposure scaling of non-contributing firms. Formally, we have

$$\begin{aligned} Exposure_{jpt} &= \alpha_0 + \alpha_1 Exposure_{-j,pt} + \eta_{ipt} && (3) \\ \log(Dem\ Money_{ijp,t+1} + 1) &= \beta_0 + \beta_1 Dem\ Win_{it} + \beta_2 Exposure_{jpt} && (4) \\ &+ \beta_3 Dem\ Win_{it} \times Exposure_{jpt} + f(V_{it}) + \epsilon_{ijp,t+1} \end{aligned}$$

where $Exposure_{-j,pt}$ represents the average exposure of non-contributing firms in industry p in year t.

If an industry is indeed heavily regulated, we would expect that all firms—including the non-contributors—would write about the regulatory issues in their 10-K filings. In other words, we expect that the average exposure to regulation for the non-contributing firms strongly predicts the exposure of the donating firms. As Table A.5 in the Appendix shows, the first stage for the instrument is positive and extremely strong, with F statistics well over 1,000.

Table 3 – Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation: 2SLS Estimates. Firm-level exposure to regulation is instrumented for using the average exposure of non-contributing firms in the same industry. Exposure is again shown to increase access-seeking behavior.

	State Legislatures		**Federal Legislatures**	
	Log Dem Contributions, $t+1$			
Dem Win	0.086	0.060	0.419	0.648
	(0.001)	(0.006)	(0.019)	(0.129)
Dem Win × Exposure	0.018	0.015	0.102	0.143
	(0.001)	(0.003)	(0.017)	(0.049)
Exposure	-0.001	0.001	-0.004	-0.008
	(0.000)	(0.001)	(0.010)	(0.015)
Constant	0.005	0.013	0.081	0.086
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	3,685,336	535,021	747,933	124,905
Number of Elections	28,645	4,187	1,553	280

Robust standard errors clustered by election in parentheses.

Table 3 presents the results, estimated by 2SLS. As the table shows, the results are consistent using this alternate strategy.[30] More exposed firms are again seen in the second row to seek more access than less exposed firms, even when we account for the possibility that access-seeking firms stress regulatory words in their 10-K reports for reasons other than their actual exposure to access. In fact, across all specifications, the estimated effects are larger in magnitude than those estimated without the IV.

To address the issue of self-reporting another way, and to show the robustness of the findings to an alternate measurement approach, we re-estimate equation 2 using the measure of industry-level exposure to regulation from Al-Ubaydli and McLaughlin (2013). Unlike the firm-level scaling we construct from SEC filings, this scaling relies only on the text of the *Code of Federal Regulations*, measuring regulatory constraints based on the frequency of constraint-relevant words in the *CFR* by industry.[31] Again, we scale this measure to run from 0 to 1 so that the interaction presents the difference in the effect for the least and most exposed firm.

Table 4 presents the results, which are consistent with those from before. Again, the second row presents the quantity of interest, the interaction of the treatment indicator with the industry-level

[30] Sample sizes differ from the previous analyses because the instrument has missing values for a small number of industries in which there are too few observations on non-contributing firms.

[31] We merge this data with ours by two-digit SIC codes.

Table 4 – Effect of Incumbency on Subsequent Donations Across Levels of Industry-Level Exposure to Regulation. Firms in industries more exposed to regulation are more sensitive to incumbency.

	State Legislatures		**Federal Legislatures**	
	Log Dem Contributions, $t+1$			
Dem Win	0.074	0.054	0.274	0.489
	(0.001)	(0.006)	(0.019)	(0.158)
Dem Win × Exposure	0.031	0.019	0.309	0.419
	(0.002)	(0.005)	(0.028)	(0.070)
Exposure	0.002	0.006	0.034	0.065
	(0.000)	(0.001)	(0.020)	(0.030)
Constant	0.004	0.012	0.062	0.099
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	4,899,907	695,336	407,571	58,786
Number of Elections	26,580	3,789	1,086	156

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation.

exposure to regulation. Firms in more regulated industries are much more sensitive to incumbency than are firms in less regulated industries.

Results Using Only Within-Firm Variation

As the previous subsection showed, the results are not driven by firm-specific reporting practices. However, it could still be the case that unobserved differences across industries contribute to the results. Perhaps more exposed industries attract executives who enjoy dabbling in political campaigns, for example. The observed contribution behavior could simply reflect the preferences of executives in exposed industries rather than a causal effect of exposure on access-seeking behavior.

To account for possibilities like this, we re-estimate equation 2 with the addition of firm and firm-treatment fixed effects:

$$\begin{aligned} \log(\textit{Dem Money}_{ij,t+1}+1) &= \alpha_j + \delta_j \textit{Dem Win}_{it} + \beta_2 \textit{Exposure}_{jt} \\ &+ \beta_3 \textit{Dem Win}_{it} \times \textit{Exposure}_{jt} + X_{it} + \epsilon_{ij,t+1}, \end{aligned} \tag{5}$$

Table 5 – Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation. Changes in firm-level exposure to regulation over time increase access-seeking contribution behavior.

	State Legislatures		**Federal Legislatures**	
	Log Dem Contributions, $t+1$			
Dem Win × Exposure	0.009	0.005	0.122	0.151
	(0.001)	(0.002)	(0.014)	(0.043)
Exposure	-0.001	-0.001	-0.009	-0.010
	(0.000)	(0.001)	(0.008)	(0.013)
Bandwidth	–	5	–	5
Firm Fixed Effects	Yes	Yes	Yes	Yes
Firm-Treatment Fixed Effects	Yes	Yes	Yes	Yes
Specification	OLS	Local Linear	OLS	Local Linear
N	8,238,340	1,183,089	753,928	126,390
Number of Elections	29,835	4,316	1,553	280

Robust standard errors clustered by election in parentheses. The main effect (Dem Win) is not reported because Firm-Treatment effects are included in all models.

where α_j are firm-fixed effects that account for level differences between firms; $\delta_j Dem\ Win_{it}$ are firm-treatment fixed effects that account for the time-invariant effect of incumbency for each firm; all other variables are the same as in equation 2.

The addition of these variables means that the effect is only identified usingwithin-firm changes in regulatory exposure and sensitivity to incumbency, relative to changes common to all firms and industries. In other words, this approach is equivalent to estimating the average incumbency sensitivity for each firm (the effects reported in Figures 5, 6 and 7) and observing whether changes in regulatory exposure *within* firms are, on average, associated with higher sensitivity to incumbency status.

Table 5 presents the results. The overall pattern is the same as in Tables 1 and 2: the coefficient on the interaction term is positive and highly statistically significant. In fact, the results are even stronger when we rely on within-firm variation. When a firm changes from low to high regulatory exposure, it becomes more inclined to target incumbents with campaign contributions relative to challengers.

Without randomly assigning firm-level exposure to regulation, we can never be sure whether exposure causes access-seeking behavior or not. Nevertheless, the correlations we have offered are telling. Firms that devote more time in their 10-K filings to discussing regulation seek more access via campaign contributions. This effect is not driven by idiosyncratic features of these firms'

reports—when we instrument for firm-level exposure using the reports of other, non-contributing firms in the same industry, we continue to find the same link between exposure and access-seeking behavior. Likewise, when we use an alternate industry-level measure of exposure that is not based on firm self-reporting, we again find this same link. Finally, firms increase their access-seeking behavior in times when they are more exposed and decrease it in times when they are less exposed. In sum, the evidence shows that the types of firms that face more exposure to regulation seek more access to incumbents. What is more, taken together, the findings strongly suggest that exposure itself increases access-seeking behavior.

Discussion and Conclusion

In this paper we have shown that access-seeking firms vary in the degree to which they seek access to incumbents through their contribution behavior, and we have demonstrated that firms which perceive themselves to be more exposed to regulation seek more access than firms with lower perceived levels of exposure. These results are consistent across U.S. federal and state legislatures, 1994–2010.

Whether, and to what extent, donors can influence the political process has remained an open question in political science. Direct attempts to correlate donation behavior with electoral or policy outcomes have found mostly null results, leading scholars to ask: "why is there so little money in politics?" (Ansolabehere, de Figueiredo and Snyder 2003). In this paper, we have taken an alternate approach. Rather than search for difficult-to-observe policy outcomes, we have focused on the variation in the behavior of access-oriented groups stemming from differences in their economic conditions. By doing so, we have uncovered an important dimension of access-seeking behavior. The fact that firms seek more access when they are more exposed to regulation strongly suggests that, despite previous evidence to the contrary, firms believe they are able to extract value from incumbents through their contributions.

Though one of the most visible political activities firms can undertake, campaign contributions are likely to represent only a small portion of the money and effort firms devote to politics (e.g., Drutman and Hopkins 2013). Indeed, the roughly 650 publicly traded firms in our dataset combined to contribute approximately 30 million dollars to U.S. House, Senate, and state legislative cam-

paigns in 2010—by no means a modest amount, but small relative to their revenues (Ansolabehere, de Figueiredo and Snyder 2003). Nevertheless, the patterns of corporate contribution activity are important.

First, the money these corporations contribute—even if a small fraction of their budgets—may play a direct role in influencing electoral outcomes.[32] Even with a conservative estimate for the average return of money on incumbent vote share, the 30 million dollars the publicly traded firms in our dataset contributed in 2010 is likely to have shifted quite a few votes. Second, and perhaps more importantly, the pattern in which the contributions are allocated highlight the strategic interests of these firms. That we can uncover a highly regular logic to the manner in which they contribute strongly suggests that they extract value from incumbents. In addition, the contribution activity of these firms acts as a proxy for a broader set of less observable behaviors. Firms who seek more access through contributions may be more likely to lobby incumbents and support campaigns in other ways. The fact that firms exposed to more regulation seek more access through contributions thus suggests that these same firms likely seek more access through these other channels, too.

The results are also important for understanding the motivations and behaviors of interest groups. The literature has long understood that some types of interest groups are "access–seeking." In this paper, we have demonstrated that there is significant variation in motivations among this category of interest groups. By identifying exposure to regulation as an important predictor of access-seeking behavior, our results point to a precise location for the link between firms and policymakers.

The pattern of evidence we uncover could distort the incentives of incumbents in several ways. First, the disproportionate access that heavily regulated firms seek out might lead incumbents to alter policies in ways contrary to the interests of most voters if, as is likely, heavily regulated firms differ from other people in their regulatory policy preferences.[33] This helps explain the difficulty of regulation; the very firms subject to regulation are those most likely to influence the political process through their contribution activity. Although scholars have long theorized about how firms try to influence regulatory policy (e.g., Peltzman 1976; Stigler 1971), we have very limited

[32]For a discussion of the possible effects of money on electoral outcomes, see: Erikson and Palfrey (2000); Gerber (1998, 2004); Green and Krasno (1988); Jacobson (1978, 1990).

[33]Of course, there may also be instances where this interaction results in better policy, if for example regulated firms have useful information about policy that voters and incumbents do not.

empirical evidence on the specific strategies that firms employ and how these strategies shape regulatory policy. This paper offers evidence for one way this strategy unfolds. Second, this access-seeking behavior could have a knock-on effect on the way incumbents craft policy. If the threat of regulatory exposure is sufficient to make firms seek access, then incumbents in need of campaign contributions have a clear incentive to create the potential for regulatory exposure, even in times and places where voters are not calling for it.

Concerns over "special-interest politics" have been ever-present in American politics, but perhaps never more so than today. Who are these "special interests"? And what do they look for when they contribute? In this paper we have offered support for an exposure theory of access: firms contribute to incumbents more when they are exposed to the decisions the government makes. The conspicuous correlation we uncover between exposure and access-seeking behavior suggests, via the revealed preference of firms, that exposed firms gain from access.

References

Abramowitz, Alan I. 1988. "Explaining Senate Election Outcomes." *The American Political Science Review* 82(2):385–403.

Al-Ubaydli, Omar and Patrick A. McLaughlin. 2013. "RegData: A Numerical Database on Industry-Specific Regulations for All U.S. Industries and Federal Regulations, 1997-2010." George Mason University Mercatus Center Working Paper No. 12-20. http://dx.doi.org/10.2139/ssrn.2099814.

Ansolabehere, Stephen and James M. Snyder, Jr. 2000. Money and Office: The Sources of the Incumbency Advantage in Congressional Campaign Finance. In *Continuity and Change in House Elections*. Stanford, CA: Stanford University Press pp. 65–87.

Ansolabehere, Stephen, John M. de Figueiredo and James M. Snyder. 2003. "Why is There So Little Money in US Politics?" *The Journal of Economic Perspectives* 17(1):105–130.

Ansolabehere, Stephen, John M. Hansen, Shigeo Hirano and James M. Snyder, Jr. 2010. "More Democracy: The Direct Primary and Competition in US Elections." *Studies in American Political Development* 24(2):190–205.

Ashworth, Scott. 2006. "Campaign Finance and Voter Welfare with Entrenched Incumbents." *American Political Science Review* 100(1):55–68.

Austen-Smith, David. 1995. "Campaign Contributions and Access." *American Political Science Review* 89:566–581.

Balvers, Ronald J, John F Gaski and Bill McDonald. 2012. "Financial Disclosure and Customer Satisfaction: Do Companies Talking the Talk Actually Walk the Walk?" *Journal of Business Ethics* pp. 1–17.

Barber, Benjamin, Jan Pierskalla and Simon Weschle. 2014. "Lobbying and the Collective Action Problem: Comparative Evidence from Enterprise Surveys." *Business and Politics* 16(2):221–246.

Barber, Michael and Nolan McCarty. 2013. Causes and Consequences of Polarization. In *Negotiating Agreement in Politics*, ed. Jane Mansbridge and Cathie Jo Martin. pp. 19–53.

Baron, David P. 1989. "Service-induced Campaign Contributions and the Electoral Equilibrium." *The Quarterly Journal of Economics* 104(1):45–72.

Campbell, John L, Hsinchun Chen, Dan S Dhaliwal, Hsin-min Lu and Logan B Steele. 2014. "The information content of mandatory risk factor disclosures in corporate filings." *Review of Accounting Studies* 19(1):396–455.

Caughey, Devin and Jasjeet S. Sekhon. 2011. "Elections and the Regression Discontinuity Design: Lessons from Close U.S. House Races, 1942–2008." *Political Analysis* 19(4):385–408.

De Figueiredo, Rui J.P. and Geoff Edwards. 2007. "Does Private Money Buy Public Policy? Campaign Contributions and Regulatory Outcomes in Telecommunications." *Journal of Economics & Management Strategy* 16(3):547–576.

Doran, Kevin L and Elias Leake Quinn. 2009. "Climate change risk disclosure: a sector by sector analysis of SEC 10-K filings from 1995-2008." *North Carolina Journal of International Law and Commercial Regulation* 34.

Drutman, Lee and Daniel J. Hopkins. 2013. "The Inside View: Using the Enron E-mail Archive to Understand Corporate Political Attention." *Legislative Studies Quarterly* 38(1):5–30.

Eggers, Andrew, Anthony Fowler, Jens Hainmueller, Andrew B. Hall and James M. Snyder, Jr. 2015. "On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: Evidence From Over 40,000 Close Races." *American Journal of Political Science* 59(1):259–274.

Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3(3):395–405.

Erikson, Robert S. and Thomas R. Palfrey. 2000. "Equilibria in Campaign Spending Games: Theory and Data." *American Political Science Review* 94(3):595–609.

Fouirnaies, Alexander and Andrew B. Hall. 2014. "The Financial Incumbency Advantage: Causes and Consequences." *Journal of Politics* 76(3):711–724.

Gaikwad, Nikhar. 2013. "Presidential Prospects, Political Support, and Stock Market Performance." *Quarterly Journal of Political Science* 8(4):451–464.

Gelman, Andrew and Gary King. 1990. "Estimating Incumbency Advantage Without Bias." *American Journal of Political Science* 34(4):1142–1164.

Gerber, Alan. 1998. "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables." *American Political Science Review* 92(2).

Gerber, Alan S. 2004. "Does Campaign Spending Work?" *American Behavioral Scientist* 47(5):541–574.

Goldman, Eitan, Jrg Rocholl and Jongil So. 2009. "Do Politically Connected Boards Affect Firm Value?" *Review of Financial Studies* 22(6):2331–2360.

Gordon, Sanford C. and Catherine Hafer. 2005. "Flexing Muscle: Corporate Political Expenditures as Signals to the Bureaucracy." *American Political Science Review* 99(2):245–261.

Green, Donald Philip and Jonathan S. Krasno. 1988. "Salvation for the Spendthrift Incumbent: Reestimating the Effect of Campaign Spending in House Elections." *American Journal of Political Science* 72:469 – 491.

Grimmer, Justin, Eitan Hersh, Brian Feinstein and Daniel Carpenter. 2012. "Are Close Elections Random?" Presented at the 2012 Annual Meeting of the Midwest Political Science Association. http://stanford.edu/ jgrimmer/CEF.pdf.

Grimmer, Justin and Eleanor Neff Powell. 2013. "Money in Exile: Campaign Contributions and Committee Access." Presented at the 2013 Annual Meeting of the Midwest Political Science Association. http://stanford.edu/ jgrimmer/money.pdf.

Hainmueller, Jens, Andrew B. Hall and James M. Snyder, Jr. N.d. "Assessing the External Validity of Election RD Estimates: An Investigation of the Incumbency Advantage." *Journal of Politics*. Forthcoming.

Hall, Richard L. and Alan V. Deardorff. 2006. "Lobbying as Legislative Subsidy." *American Political Science Review* 100(1):69–84.

Imbens, Guido W. and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2):615–635.

Jacobson, Gary C. 1978. "The Effects of Campaign Spending in Congressional Elections." *American Political Science Review* 72(2):469–491.

Jacobson, Gary C. 1990. "The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments." *American Journal of Political Science* 34(2):334–362.

Jacobson, Gary C. 2009. *The Politics of Congressional Elections*. Pearson/Longman.

Jolliffe, Ian. 2005. *Principal Component Analysis*. Wiley Online Library.

Kalla, Joshua L. and David E. Broockman. N.d. "Congressional Officials Grant Access Due To Campaign Contributions: A Randomized Field Experiment." *American Journal of Political Science*. Forthcoming.

Klarner, Carl, William Berry, Thomas Carsey, Malcolm Jewell, Richard Niemi, Lynda Powell and James Snyder. 2013. "State Legislative Election Returns (1967-2010).".
URL: *http://dx.doi.org/10.3886/ICPSR34297.v1*

Krasno, Jonathan S., Donald Philip Green and Jonathan A. Cowden. 1994. "The Dynamics of Campaign Fundraising in House Elections." *Journal of Politics* 56(2):459–474.

Kravet, Todd and Volkan Muslu. 2013. "Textual risk disclosures and investors? risk perceptions." *Review of Accounting Studies* 18(4):1088–1122.

Lee, Davis S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics* 142(2):675–697.

Levitt, Steven D. 1994. "Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House." *The Journal of Political Economy* 102(4):777 – 798.

Li, Feng, Russell Lundholm and Michael Minnis. 2013. "A Measure of Competition Based on 10-K Filings." *Journal of Accounting Research* 51(2):399–436.

Loughran, Tim and Bill McDonald. 2011. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." *The Journal of Finance* 66(1):35–65.

Loughran, Tim, Bill McDonald and Hayong Yun. 2009. "A wolf in sheep?s clothing: The use of ethics-related terms in 10-K reports." *Journal of Business Ethics* 89(1):39–49.

Peltzman, Sam. 1976. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19:211–248.

Rogers, Jonathan L, Andrew Van Buskirk and Sarah LC Zechman. 2011. "Disclosure tone and shareholder litigation." *The Accounting Review* 86(6):2155–2183.

Shi, Guifeng and Li Zhang. 2014. "Product market competition and unemployment: a county-level analysis." *Applied Economics Letters* 21(16):1134–1138.

Skinner, Douglas J. 1994. "Why firms voluntarily disclose bad news." *Journal of accounting research* pp. 38–60.

Snyder, Jason. 2005. "Detecting Manipulation in US House Elections." Unpublished Manuscript, University of California, Los Angeles. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.335.6505&rep=rep1&type=pdf`.

Snyder, Jr., James M. 1990. "Campaign Contributions as Investments: The U.S. House of Representatives, 1980-1986." *Journal of Political Economy* 98(6):1195 – 1986.

Snyder, Jr., James M. 1992. "Long-term Investing in Politicians; or, Give Early, Give Often." *Journal of Law & Economics* 35(1):15–43.

Spirling, Arthur. 2012. "US Treaty Making with American Indians: Institutional Change and Relative Power, 1784–1911." *American Journal of Political Science* 56(1):84–97.

Stigler, George. 1971. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2:3–21.

Wiesehomeier, Nina and Kenneth Benoit. 2009. "Presidents, Parties, and Policy Competition." *The Journal of Politics* 71(4):1435–1447.

Online Appendix

Contents

A.1 Additional Statistical Results and Robustness Checks

A.1.1 Balance Tests for RD

Table A.1 presents balance tests for the RD design using the same bandwidth and specification as that in the paper. As the table shows, no evidence for RD imbalance is found.

Table A.1 – Balance Tests: Effect of Incumbency on Lagged Donations Across Levels of Industry-Level Exposure to Regulation.

	State Legislatures		**U.S. House**	
	Log Dem Contributions, $t-1$			
Dem Win	0.019	0.010	0.008	0.068
	(0.014)	(0.010)	(0.105)	(0.069)
Bandwidth	2.5	5	2.5	5
Specification	Local Linear	Local Linear	Local Linear	Local Linear
N	300,879	610,859	122,879	234,560

Robust standard errors clustered by election in parentheses.

A.1.2 RD Estimates Across Bandwidths and Specifications

Figures A.1 and A.2 present RD estimates of the quantity of interest—β_3, the coefficient on the interaction term measuring how much larger the effect of incumbency is for the most exposed firm relative to the least exposed firm—for U.S. state legislatures and the U.S. House and Senate according to equation 2 at all possible bandwidths from 1 to 20 and four possible specifications: local linear or a two, three, or four degree polynomial. As the graphs show, the same substantive conclusion holds at every bandwidth or specification. Figures A.3 and A.4 present RD estimates using the local linear specification at all possible bandwidths along with 95% confidence intervals from robust standard errors clustered by election. Again, we see that the same substantive conclusion holds across bandwidths.

Figure A.1 – RD Estimates Across Bandwidths and Specifications.

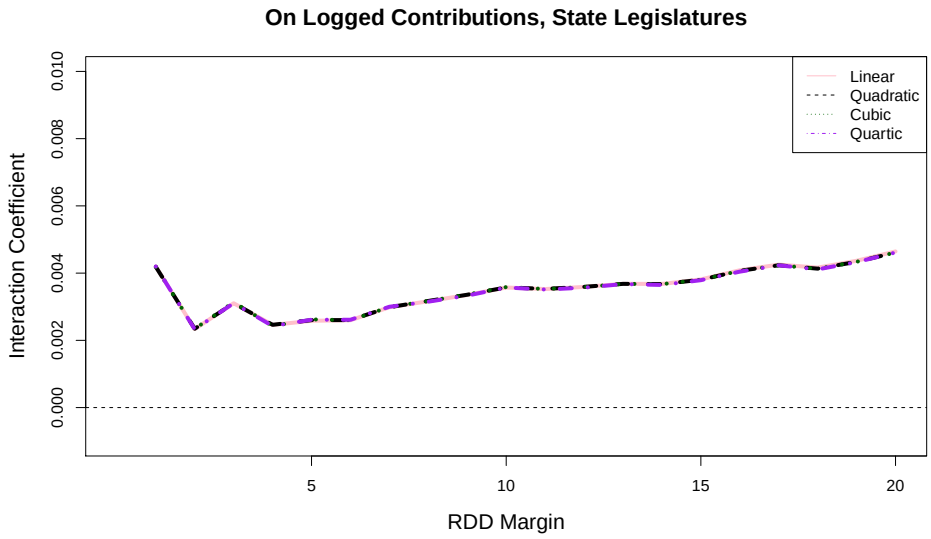


Figure A.2 – RD Estimates Across Bandwidths and Specifications.

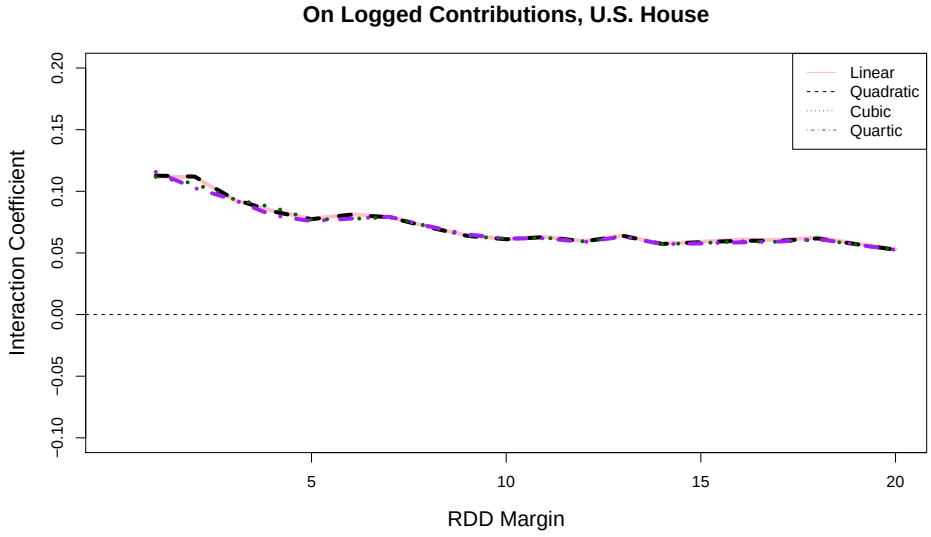


Figure A.3 – Local Linear RD Estimates Across Bandwidths

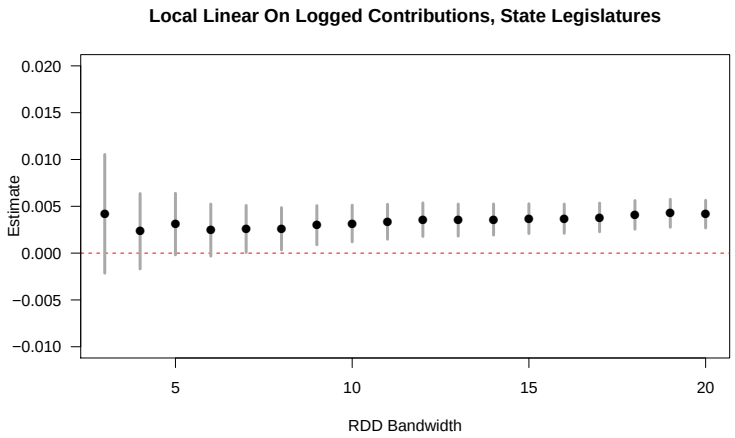
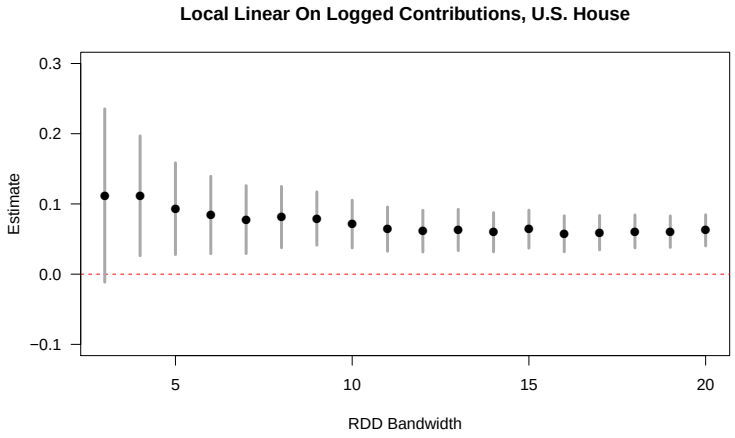


Figure A.4 – Local Linear RD Estimates Across Bandwidths



A.1.3 Additional Statistical Results

A.1.3.1 Two-Way Clustering

Table A.2 presents the federal estimates like in Table 1 but with two-way clustered standard errors, clustered by both firm and election. These standard errors are produced using the `cgmreg.ado` program in Stata. Unfortunately, the state legislative data is too large for cgmreg.ado to handle. That said, the results for the federal legislatures are those where we might worry more about the standard errors since there are fewer observations. As the table shows, the standard errors in both the main specification (column 1) and the RD (column 4) remain highly statistically significant; in the middle two columns the interaction coefficients are no longer significant.

Table A.2 – Federal Legislatures: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation, Two-Way Clustering. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.418	0.485	0.412	0.647
	(0.033)	(0.040)	(0.037)	(0.133)
Dem Win × Exposure	0.057	0.040	0.025	0.078
	(0.022)	(0.029)	(0.024)	(0.033)
Exposure	-0.003	-0.081	-0.010	-0.001
	(0.006)	(0.033)	(0.007)	(0.008)
Constant	0.081	0.339	0.302	0.085
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	753,928	711,894	753,928	126,390
Number of Elections	.	.	.	.

Robust standard errors two-way clustered by election and by firm in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Presidential vote share and the interaction with Exposure, Democratic incumbency and the interaction with Exposure, Democratic candidate quality and the interaction with Exposure.

A.1.3.2 Robustness to Log Specification

Table A.3 presents results using $\log(Dem\ Money + 1000)$ instead of $\log(Dem\ Money + 1)$ as the outcome variable. The results are robust to this variation—note that the magnitude of the estimates changes because of the changing interpretation of the coefficients.

Table A.3 – Robustness Check Using Log Money + 1000. Firms more exposed to regulation are more sensitive to incumbency.

	State Legislatures		**U.S. House**	
	Dem Contributions Per 100,000 State Residents, $t+1$			
Dem Win	0.0039	0.0036	0.1251	0.1101
	(0.0008)	(0.0005)	(0.0368)	(0.0239)
Dem Win × Exposure	0.0002	0.0002	0.0196	0.0153
	(0.0001)	(0.0001)	(0.0067)	(0.0045)
Exposure	-0.0000	-0.0000	0.0012	-0.0005
	(0.0000)	(0.0000)	(0.0014)	(0.0013)
Constant	6.9086	6.9085	6.9283	6.9201
Bandwidth	2.5	5	2.5	5
Specification	Local Linear	Local Linear	Local Linear	Local Linear
N	587,559	1183089	68,798	126,402
Number of Elections	2,128	4,316	152	280

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation.

A.1.3.3 Robustness to Outliers

Table A.4 presents results excluding the positive outlier firms with exposure scalings above 0.4 (see Figure 2). Estimates are robust to this variation.

Table A.4 – Outliers Removed. Firm-years with scalings above 0.4 are excluded. Firms more exposed to regulation are more sensitive to incumbency.

	State Legislatures		**U.S. House**	
	Log Dem Contributions, $t+1$			
Dem Win	0.061	0.057	0.738	0.660
	(0.008)	(0.006)	(0.193)	(0.128)
Dem Win × Exposure	0.009	0.009	0.169	0.131
	(0.005)	(0.003)	(0.075)	(0.052)
Exposure	0.000	0.000	-0.005	-0.037
	(0.001)	(0.001)	(0.031)	(0.023)
Bandwidth	2.5	5	2.5	5
Specification	Local Linear	Local Linear	Local Linear	Local Linear
N	465,966	941,906	54,645	101,167
Number of Elections	2,128	4,316	152	280

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation.

A.1.3.4 First-Stage Effect of Industry Instrument

Table A.5 shows that there is a very strong first stage between the instrument—exposure for non-contributing firms in the same industry—and the exposure measure.

Table A.5 – First-Stage Estimates: Instrumenting for Exposure Using Exposure of Non-Contributing Firms in Same Industry. There is a very strong first stage for the instrument.

	State Legislatures		**Federal Legislatures**	
	Firm-Level Exposure			
Exposure Instrument	0.505	0.504	0.809	0.792
	(0.000)	(0.001)	(0.002)	(0.004)
First Stage F-test	2.6e+05	3.8e+04	5.9e+04	1.2e+04
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	3685336	535,021	747,933	124,905
Number of Elections	28,645	4,187	1,553	280

Robust standard errors clustered by election in parentheses.

A.1.3.5 Per-Capita Donations Instead of Logs

Table A.6 presents the main results using firm donations per 100,000 residents in the state in which the donation occurs, to help with the interpretation problem resulting from using log dollars plus one.

Table A.6 – Per-Capita Donations instead of logs. Firms more exposed to regulation are more sensitive to incumbency.

	State Legislatures		**U.S. House**	
	Dem Contributions Per 100,000 State Residents, $t+1$			
Dem Win	0.073	0.078	22.498	14.742
	(0.012)	(0.009)	(10.764)	(6.535)
Dem Win × Exposure	0.002	0.002	1.270	1.177
	(0.003)	(0.002)	(1.562)	(0.969)
Exposure	-0.002	-0.001	-0.030	-0.135
	(0.001)	(0.001)	(0.082)	(0.118)
Constant	0.021	0.020	0.419	0.224
Bandwidth	2.5	5	2.5	5
Specification	Local Linear	Local Linear	Local Linear	Local Linear
N	587,559	1183089	68,810	126,422
Number of Elections	2,128	4,316	152	280

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation.

A.1.3.6 Results For Open-Seat Vs. Incumbent-Contested Races

Table A.7 and A.8 show how the estimated effect varies across open seats and seats held by an incumbent at time at the federal and state level, respectively. At both the state and federal level, the results suggest that the effect is very similar across open-seats and seats held by an incumbent (at time t).

Table A.7 – Federal Legislatures: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation.

	Open Seat, t		**Dem. Incumbent, t**	
	Log Dem Contributions, $t+1$			
Dem Win	0.359	0.515	0.454	0.588
	(0.037)	(0.102)	(0.015)	(0.100)
Dem Win × Exposure	0.046	0.051	0.055	0.042
	(0.024)	(0.043)	(0.010)	(0.031)
Exposure	0.007	-0.001	-0.005	-0.001
	(0.011)	(0.008)	(0.006)	(0.009)
Constant	0.045	0.031	0.034	0.043
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	73,450	29,343	544,747	33,821
Number of Elections	158	68	1,120	89

Robust standard errors clustered by election in parentheses.

Table A.8 – State Legislatures Open Seats: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation.

	Open Seat, t		**Dem. Incumbent, t**	
	Log Dem Contributions, $t+1$			
Dem Win	0.085	0.075	0.068	0.053
	(0.003)	(0.010)	(0.002)	(0.008)
Dem Win × Exposure	0.007	0.005	0.002	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
Exposure	-0.001	-0.000	0.003	0.004
	(0.000)	(0.000)	(0.001)	(0.001)
Constant	0.004	0.009	0.016	0.013
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	1754310	426,266	3,254,989	340,080
Number of Elections	6,547	1,503	11,857	1,452

Robust standard errors clustered by election in parentheses.

A.1.3.7 Results Using Total Donations

Tables A.9 and A.10 re-estimate the main results with the addition of a control variable measuring the total amount of all contributions each firm makes over the entire dataset. Again, results are unchanged by this specification.

Table A.9 – Federal Legislatures: Controling for total firm donations. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.425	0.499	0.411	0.648
	(0.019)	(0.022)	(0.026)	(0.129)
Dem Win × Exposure	0.050	0.025	0.020	0.076
	(0.009)	(0.016)	(0.007)	(0.025)
Exposure	0.024	-0.066	0.012	0.019
	(0.005)	(0.026)	(0.004)	(0.008)
Total Firm Donations	0.591	0.590	0.593	0.462
	(0.591)	(0.590)	(0.593)	(0.462)
Constant	0.067	0.328	-0.073	0.074
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	753,928	711,894	753,928	126,390
Number of Elections	1,553	1,467	1,553	280

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include presidential vote share as a control variable.

Table A.10 – State Legislatures: Controling for total firm donations. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.080	0.079	0.054	0.056
	(0.001)	(0.002)	(0.001)	(0.005)
Dem Win × Exposure	0.006	0.006	0.004	0.003
	(0.001)	(0.001)	(0.000)	(0.001)
Exposure	-0.001	-0.000	-0.002	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Total Firm Donations	0.079	0.079	0.079	0.074
	(0.079)	(0.079)	(0.079)	(0.074)
Constant	0.003	0.003	-0.004	0.012
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	8238340	8192487	8,238,340	1,183,089
Number of Elections	29,835	29,673	29,835	4,316

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include seniority as a control variable.

A.1.3.8 Results Controlling for Firm Characteristics

An important concern with the results in the paper is that exposure to regulation—or, simply, the decision to discuss regulation in 10K filings—might be correlated with other firm characteristics. For example, larger firms might be more concerned about regulation, because they are larger, and might also seek out incumbents more with their contributions, not because they care about regulatory policy but because, due to their size, they have other reasons to care more about policy. In this subsection, we investigate this possibility by directly controlling for the number of employees and the market value of each firm. To do so, we gathered data from Compustat's database of financial, statistical and market information on active and inactive companies via `www.compustat.com`. These variables are merged in using CIK and year.

Tables A.11 and A.12 re-estimate the main results with the addition of two control variables that measure firm characteristics.

Table A.11 – Federal Legislatures: Sensitivity to Firm Characteristics.
Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	-4.112	-4.336	-4.099	-4.495
	(0.174)	(0.129)	(0.192)	(0.400)
Dem Win × Exposure	0.067	0.039	0.047	0.070
	(0.009)	(0.016)	(0.008)	(0.026)
Exposure	0.012	-0.085	0.016	0.016
	(0.005)	(0.028)	(0.006)	(0.007)
Constant	-0.423	0.180	-0.524	-0.469
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	515,735	485,672	515,735	74,228
Number of Elections	1,086	1,022	1,086	156

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Presidential vote share and the interaction with Exposure, Democratic incumbency and the interaction with Exposure, Democratic candidate quality and the interaction with Exposure. All models include log(employees) and log(market value) as well as their interactions with the treatment variable.

Table A.12 – State Legislatures: Sensitivity to Firm Characteristics. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	-0.730	-0.730	-0.758	-0.613
	(0.012)	(0.011)	(0.012)	(0.028)
Dem Win × Exposure	0.017	0.017	0.016	0.012
	(0.001)	(0.001)	(0.001)	(0.002)
Exposure	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.031	-0.029	-0.019	-0.055
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	6,466,319	6,433,126	6,466,319	914,362
Number of Elections	26,542	26,410	26,542	3,782

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Democratic incumbency and the interaction with Exposure. All models include log(employees) and log(market value) as well as their interactions with the treatment variable.

A.1.3.9 Results Omitting Defense Firms

Another concern is that firms who are particularly reliant on government business might (a) talk about government matters in their regulatory filings more, and (b) might contribute to incumbents more. In some ways, this link would still suggest the value of access, but it would alter our ideas over what access means. "Access" is less interesting if it only relates to firms already reliant on the government for contracts, revenue, etc, than if it relates also to firms whose business is largely private but who desire influence over government policies that affect private business.

Ideally we would address this possibility by controlling for the proportion of a firm's revenue that comes from the government. Unfortunately we do not have access to such a variable. What we can do, however, is re-estimate the results excluding defense firms—a set of firms we know to be particularly reliant on government business. To do so, we take advantage of the fact that the state legislative contributions from FollowTheMoney record each firm's sector, identifying Defense-related firms directly. Accordingly, Table A.13 presents these estimates at the state legislative level. As we see, we see the same pattern of results even among non-defense firms.

Table A.13 – State Legislatures: Effect of Incumbency on Subsequent Donations Across Levels of Donor Firm Exposure to Regulation, Excluding Defense Firms. Firms more exposed to regulation are more sensitive to incumbency.

	Log Dem Contributions, $t+1$			
Dem Win	0.081	0.080	0.055	0.057
	(0.001)	(0.002)	(0.001)	(0.005)
Dem Win × Exposure	0.006	0.005	0.004	0.002
	(0.001)	(0.001)	(0.000)	(0.001)
Exposure	-0.001	-0.000	-0.001	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.004	0.004	0.001	0.013
Specification	OLS	OLS	Diff-in-Diff	Local Linear
Controls		✓		
District and Year FE			✓	
Bandwidth	–	–	–	5
N	8,123,306	8,078,087	8,123,306	1,166,656
Number of Elections	29,835	29,673	29,835	4,316

Robust standard errors clustered by election in parentheses. Exposure normalized by subtracting mean and dividing by standard deviation. In the second model, we include the following controls: Democratic incumbency and the interaction with Exposure.

A.1.3.10 Industry Clustered Standard Errors

Table A.14 corresponds to Table 3 in the paper, but reports robust standard errors clustered at the 4-digit SIC industry code (the industry definition used to define the instrument).

Table A.14 – 2SLS Estimates with Industry Clustered Standard Errors. Firm-level exposure to regulation is instrumented for using the average exposure of non-contributing firms in the same industry. Exposure is again shown to increase access-seeking behavior.

	State Legislatures		**Federal Legislatures**	
	Log Dem Contributions, $t+1$			
Dem Win	0.086	0.060	0.419	0.648
	(0.018)	(0.014)	(0.033)	(0.048)
Dem Win × Exposure	0.018	0.015	0.102	0.143
	(0.009)	(0.008)	(0.038)	(0.051)
Exposure	-0.001	0.001	-0.004	-0.008
	(0.001)	(0.002)	(0.006)	(0.009)
Constant	0.005	0.013	0.081	0.086
Bandwidth	–	5	–	5
Specification	OLS	Local Linear	OLS	Local Linear
N	3,685,336	535,021	747,933	124,905
Number of Elections	62	62	257	257

Robust standard errors clustered by industries in parentheses.

A.2 Detailed Description of Data

A.2.1 Campaign Contributions

At the federal level, the data on campaign contributions is obtained from the Federal Election Commission (FEC). We keep all observations that pertain to donations from PACs associated with corporations, and based on firm names we map FEC IDs to the Central Index Key (CIK) used by the Securities and Exchange Commission. SEC use the CIK id to uniquely identify corporations traded on U.S. stock exchanges.

At the state level, the data on campaign contributions is obtained from `Followthemoney.org`. We keep all observations pertaining to firms that in total donate more than $10,000 from 1994-2010. Based on firm name, we map each firm to the unique CIK Identifier.

A.2.2 Securities and Exchange Commission data

Each year all firms that are publicly traded on U.S. stock exchanges submit a comprehensive summary of the company's financial performance in the so-called 10-K file. The 10-K files contains qualitative and quantitative information on the individual firm and the market it operates in. By law the management is required to discuss all major threats and risks that can affect the business. The 10-K filings are publicly available (back to the mid-1990s) via the SEC's website `http://www.sec.gov/edgar.shtml`.

Table A.15 – Share of 10-K files Discussing State and Federal Regulation

Industry	% 10-K files containing State Regulatory Words	% 10-K files containing Federal Regulatory Words	N
Mining, Agriculture, Construction	70.12	69.43	3467
Manufacturing	41.20	41.37	20742
Utilities, Transport, Trade	61.87	61.99	10098
Services, Financial Industry	68.81	68.13	19645
All Industries	56.98	56.78	53952

State regulatory words refer to 'state law', 'state leg', 'state reg', 'state agenc', 'state gov'.
For the federal regulatory words, 'state' is replace with 'federal'

The 10-K file contains four main parts:

Part 1: Description of Business (ITEM 1), Risk Factors (ITEM 1A), Unresolved Staff Comments (ITEM 1B), Description of Properties (ITEM 2), Legal Proceedings (ITEM 3), Mine Safety Disclosures (ITEM 4)

Part 2: Market for Registrant's Common Equity, Related Stockholder Matters and Issuer Purchases of Equity Securities (ITEM 5), Selected Financial Data (ITEM 6), Management's Discussion and Analysis of Financial Condition and Results of Operations (ITEM 7), Quantitative and Qualitative Disclosures About Market Risk (ITEM 7A), Financial Statements and Supplementary Data (ITEM 8), Changes in and Disagreements With Accountants on Accounting and Financial Disclosure (ITEM 9), Controls and Procedures (ITEM 9A), Other Information (ITEM 9B)

Part 3: Directors, Executive Officers and Corporate Governance (ITEM 10), Executive Compensation (ITEM 11), Security Ownership of Certain Beneficial Owners and Management and Related Stockholder Matters (ITEM 12), Certain Relationships and Related Transactions, and Director Independence (ITEM 13), Principal Accounting Fees and Services (ITEM 14)

Part 4: Exhibits, Financial Statement Schedules Signatures (ITEM 15)

Table A.16 – 10-K Filing Firms and Donors by Year

Year	Total # Firms	Pct. of Firms Donating	# Donating Firms (*both* Federal and State)	# Donating Firms (*only* Federal)	# Donating Firms (*only* State Elections)
1994	1925	15.273	143	114	37
1996	6353	6.879	190	199	48
1998	9831	5.391	225	249	56
2000	9531	5.886	235	276	50
2002	8717	6.849	252	298	47
2004	8385	7.287	260	302	49
2006	8619	7.263	269	312	45
2008	8711	7.301	277	319	40
2010	8831	6.998	271	315	32

The sample includes firms submitting 10-K files to SEC.

Table A.17 – Number of Firms in Sample

	Federal Donors	State Donors
# Firms	1553	278853
# Firms Donating more than $10,000	1197	9038
# Publicly-traded Donor Firms	664	363

At the state level, the final sample is restricted to firms donating more than $100,000 from 1994-2010.

We collected 10-K files for firms traded on New York Stock Exchange, NASDAQ and AMEX from 1994-2014, and we exclude firms from the sample where the main business address is located outside the country. In total, this gives us a sample of 4,222 unique publicly-traded firms. Table A.17 shows that firms discuss both state and federal regulation in the 10-K files.

Table A.16 shows the number of 10-K filing firms in our sample and the share of these firms that contribute to either federal or state legislative campaigns. The table illustrates that publicly available 10-K files increase over time, primarily because the filing process was digitized. Approximately, 20-30% of the publicly-traded firms donate in either state or federal legislative elections. Approximately 56% of the 10-K files mention key regulatory words.

Table A.17 shows how the number of firms in the sample is affected by various restrictions on the sample.

Figure A.5 shows the industrial composition of 10-K filing firms.

Figure A.6 shows the number of firms by the state of their main business address.[34] Although many firms, for obvious reasons, have head quarters in more populated states (CA, NY, TX), publicly-traded firms are distributed across all states. The share of publicly traded firms that in total donate more than $10,000 is illustrated by the light-grey area in the pie charts.

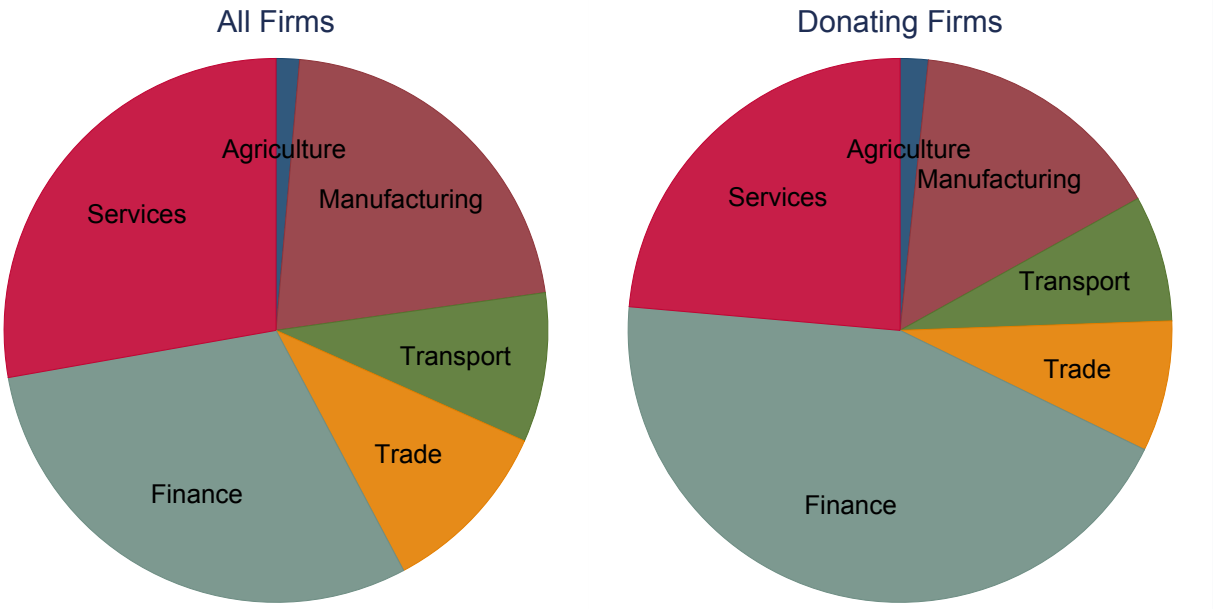
A.2.3 Donations in close elections

Tables A.18 and A.19 display the number of observations that enter the RD sample within the 5% bandwidth for both the U.S. state legislatures analysis and the federal legislatures analysis. For the state legislatures, some missingness stems from the fact that

1. `Followthemoney.org` started collecting data in the late 1990s for some states;

[34]Note that for most firms the state of their main business operations is *not* the same as the state of incorporation, since most firms are incorporated in Delaware.

Figure A.5 – 10-K Filing Firms by Industry



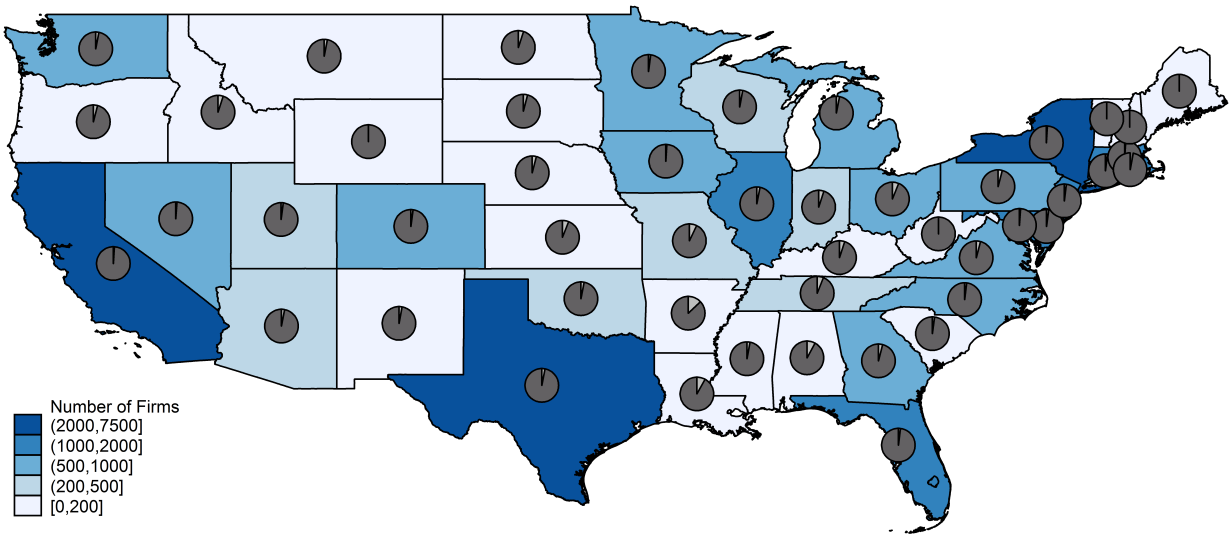
NOTE: The pie chart on the left show the industrial composition of *all* 10-K filing firms 1994-2010, and the pie chart on the right shows the composition of *donating* firms.

2. we exclude chambers where legislators are elected in multi-member districts;

3. some states forbid contributions from corporations.

In general, most missingnes comes from times and places with relatively few close elections.

Figure A.6 – Firms in Sample by State



NOTE: The light-grey area in the pie chart illustrates the share of publicly-traded firms in a given state that donate to either federal or state legislative elections.

Table A.18 – Observations in State Legislative Data Set, by State and Office. Each cell provides the total number of data points in the dataset used for analysis with the 5% RD Bandwidth.

State	# Upper House	# Lower House	Min Year	Max Year
AK	0	7435	1994	2010
AL	1780	2808	1998	2010
AR	15	3965	2000	2010
AZ	1706	0	1996	2010
CA	1107	6205	1998	2010
CO	3882	12907	1996	2010
CT	2902	6643	1996	2010
DE	719	3826	2000	2010
FL	1246	11198	1998	2010
GA	4864	10920	1994	2010
HI	864	6334	1998	2010
IA	4298	14707	1998	2010
ID	3133	0	1994	2010
IL	2774	9498	1996	2010
IN	1912	13958	1994	2010
KS	2807	13013	1996	2010
KY	3189	11501	1994	2010
MD	1537	0	1998	2010
ME	2406	11126	1996	2010
MI	3083	13879	1996	2010
MN	5811	0	1996	2010
MO	2163	19153	1996	2010
MS	1115	1422	1999	2007
MT	2316	12578	1994	2010
NC	8181	14081	1996	2010
ND	1844	0	1998	2010
NH	6846	0	1996	2010
NJ	2133	0	1997	2007
NM	1439	9980	1994	2010
NV	0	6643	1994	2010
NY	4489	8226	1998	2010
OH	1503	11834	1996	2010
OK	1442	6799	2000	2010
OR	3148	11553	1994	2010
PA	1399	10901	1998	2010
RI	2371	1559	1994	2010
SC	806	5889	1996	2010
SD	4886	0	2000	2010
TN	1522	7996	1996	2010
TX	405	10625	1998	2010
UT	2585	8709	1996	2010
VA	845	5541	1999	2009
WA	6197	0	1994	2010
WI	2274	12103	1998	2010
WY	867	7412	1994	2010

Table A.19 – Observations in U.S. House Data Set, by State. Each cell provides the total number of data points in the dataset used for analysis with the 5% RD Bandwidth.

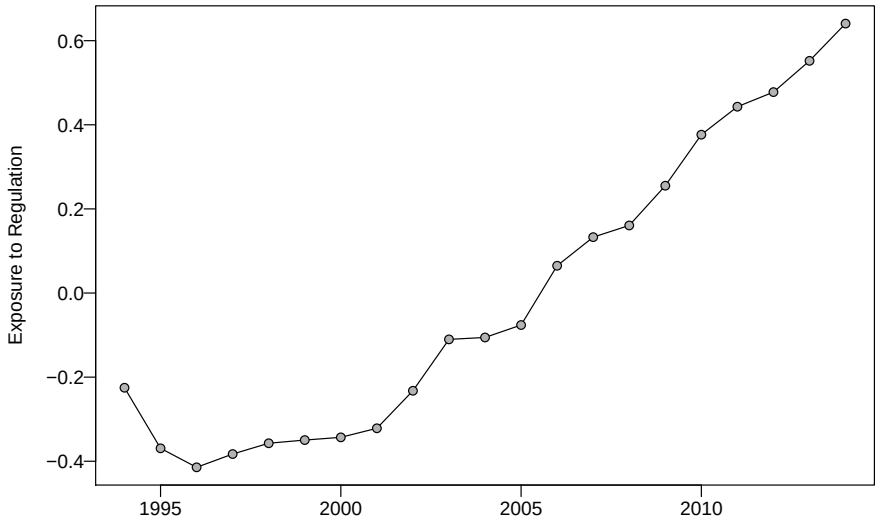
State	# Obs	Min Year	Max Year
AL	2227	1994	2008
AR	1964	1994	2002
AZ	2696	1996	2008
CA	8295	1994	2006
CO	3263	1998	2006
CT	2885	1994	2006
FL	6120	1994	2008
GA	3881	1994	2008
HI	388	1996	1996
IA	2935	1994	2006
ID	986	1996	2008
IL	3861	1994	2008
IN	8090	1994	2006
KS	1188	1994	2002
KY	4394	1994	2008
LA	1196	2008	2008
MA	776	1996	1996
MD	1690	2002	2008
ME	1654	1994	2008
MI	4494	1994	2008
MN	1868	1994	2006
MO	1286	1994	2004
MS	473	1998	1998
MT	582	2006	2006
NC	4321	1994	2008
NE	1108	1994	2008
NH	1196	2008	2008
NJ	2659	1994	2008
NM	1613	1998	2006
NV	3350	1994	2008
NY	6191	1994	2008
OH	5335	1994	2008
OK	898	1994	1996
OR	3848	1994	2008
PA	9787	1994	2008
RI	835	1994	2006
SC	388	1996	1996
SD	934	1996	2002
TN	1529	1994	2002
TX	3591	1994	2008
UT	1019	1998	2002
VA	1346	1994	2006
WA	5071	1994	2006
WI	3412	1994	2008

A.3 Measure Description and Validation

A.3.1 Variation in Exposure Over Time

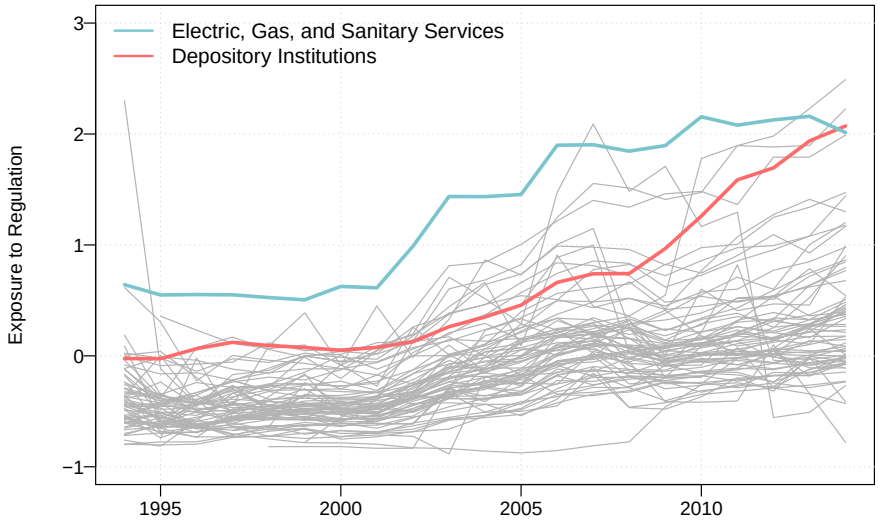
To investigate the measure, we first plot the average exposure to regulation, across all firms in the dataset, for each year. As Figure A.7 shows, there has been a steady increase in the frequency of firms discussing regulatory exposure in their 10-K filings over time.

Figure A.7 – Average exposure to regulation over time.



We can investigate variation in the exposure scaling over time, also, in a first attempt to validate the measure. Figure A.8 shows the average exposure to regulation by industry over time, for every industry in the dataset. Two industries are highlighted in the plot: Electric, Gas, and Sanitary Services (sic code 49) and Depository Institutions (sic code 60). First, we see that energy companies (the main component of the Electric, Gas, and Sanitary Services industry grouping) are consistently scaled as the most exposed to regulation—which we regard as consistent with anecdotal evidence and political chatter about the regulation of this group of companies. Second, we also see that there has been a pronounced uptick in the frequency with which Depository Institutions write about regulatory exposure in their filings. This is consistent with the view that financial regulation has increased in intensity since the financial crisis.

Figure A.8 – Industry-by-industry scalings over time. There is a general upward trend in the regulatory exposure measure. Two particularly highly-scaled industries are shown. The Electric, Gas, and Sanitary Services category (sic code 49) consistently ranks extremely high in discussing regulatory issues. Historically, Depository Institutions exhibited lower levels of exposure, but have risen sharply in the aftermath of the financial crisis.

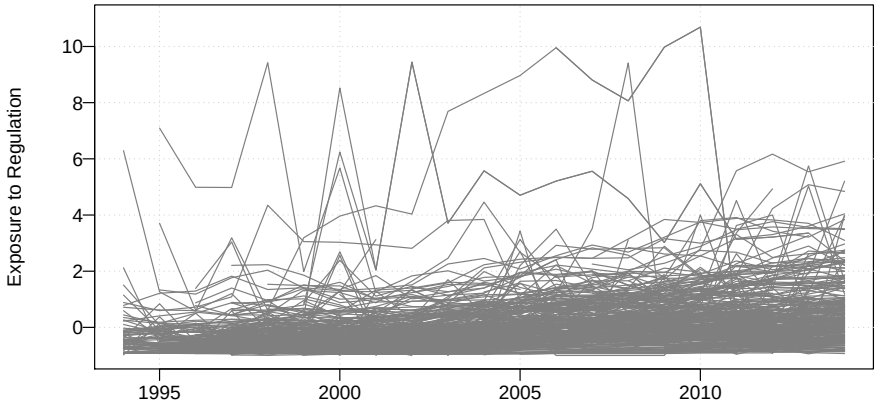


A.3.2 Within-Firm Vs. Across-Firm Variation in Exposure

There is considerable within- and between-firm variation in the exposure measure. The between-firm standard deviation for the standardized measure is 0.84, while the within-firm standard deviation is 0.53. This helps explain why we are able to obtain significant empirical leverage even when we include firm fixed effects in our regression.

To help visualize the variation, Figure A.9 plots the over-time exposure for 500 randomly sampled firms in the dataset for which we have at least 4 years of data. Several things are notable in the plot. First, a few outlier firms are noticeable that are both (a) much higher in the measure, producing a large between-firm variation, and (b) noisy over time, producing a large within-firm variation (note that we re-estimate the results, omitting outliers, in a previous Appendix section). The larger mass of firms, while more tightly clustered, exhibits considerable fluctuations over time. We continue to see the same general increase in exposure over time that the previous figures showed, as well.

Figure A.9 – Across and Within-Firm Variation.



A.3.3 Correlation with Previous Measure of Regulation

Here, we attempt to validate our measure by comparing it to an existing measure of the intensity of regulation based on the Federal Register (Al-Ubaydli and McLaughlin 2013). There are several obstacles in making this comparison. First, this alternative measure is only available at the industry level, where industries are defined by two-digit NAICS codes. The mapping between NAICS codes and SIC codes is complicated, and for many industries there are many SIC codes that correspond to a single two-digit NAICS code. As a result, we can only make the comparison after (a) merging in the alternative scaling, which drops a significant number of observations, and (b) aggregating our scaling to the two-digit NAICS industry code level as best as possible.

Second, this alternative measure does not target the same concept as our measure. While we measure regulatory exposure—focusing on how concerned firms are about the regulatory environment, either because of existing regulation or the possibility of future regulation—the CFR-based approach measures the volume of existing regulation. This is more akin to regulatory burden rather than exposure. In particular, we might suspect that, while the CFR-based measure accumulates over time, the exposure measure is instead "stationary," that is to say, not necessarily increasing or decreasing over time since it is forward as well as backward looking. Despite this conceptual difference, comparing the two may still prove useful since both concern the regulatory process broadly speaking.

Figure A.10 – Correlation of Scaling to Previous Measure of Industry-Level Regulation. Clear within-industry correlations are present. Graph is best viewed in color to differentiate industries.

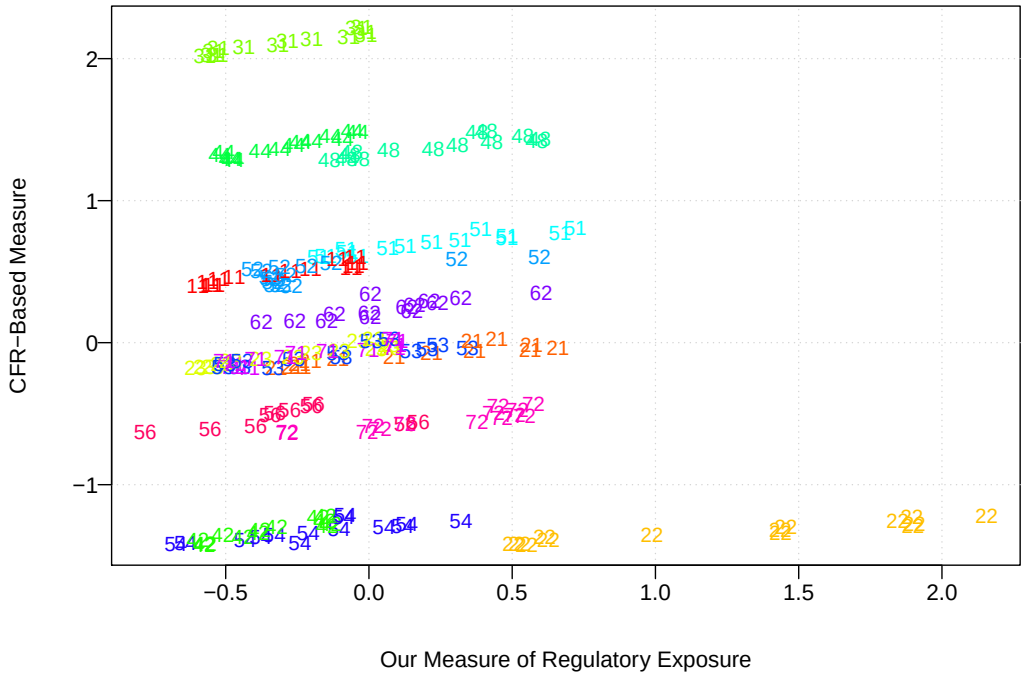


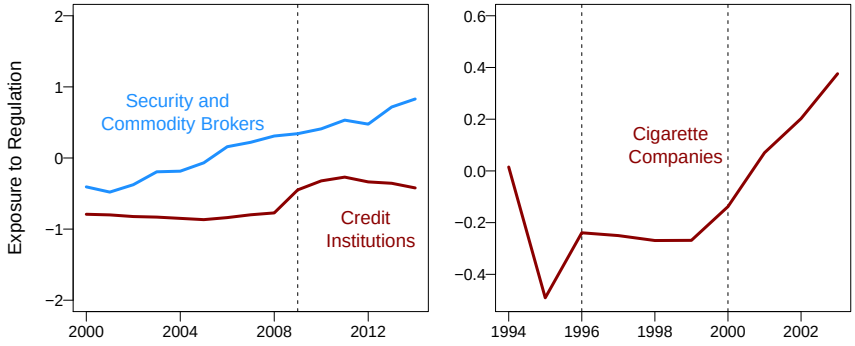
Figure A.10 presents the resulting comparison. Each point in the graph represents an industry-year observation. The vertical axis represents the CFR-based scaling, and the horizontal axis represents our text-based measure. Points are labeled according to the two-digit NAICS code of the corresponding industry, and industries are given unique colors. While it is clear that there is no overall, cross-industry correlation, there appears to be a strong within-industry correlation—that is to say, over time within industries the two measures rise and fall together. This is to be expected for the reason discussed above. The CFR-based measure is cumulative. The level of regulation within each industry for the CFR-baed measure is therefore indicative of the industry's regulatory history, but not necessarily of its future exposure. *Changes* in the CFR-based measure, on the other hand, indicate present-day regulatory changes, which we would expect our exposure measure to pick up as well. That is why when we make within-industry comparisons—which "difference out" the pre-existing levels of regulatory burden, we find positive slopes.

We also confirm this within-industry correlation more formally. Specifically, we estimate a regression of the CFR-based measure on our regulatory measure (again at the industry-year level),

and we include industry fixed effects in the regression. We find a strong, highly statistically significant association between the two. In particular, the estimated coefficient on our measure of exposure is 0.18—indicating that a one standard-deviation increase in exposure maps to a 0.18 standard deviation increase in the CFR-based measure—with a t-statistic of 20.7.

A.3.4 Investigating Policy Changes and the Exposure Measure

Figure A.11 – Validating the Exposure to Regulation Measure. Financial firms are scaled as more exposed to regulation in the build-up to, and aftermath of, the Dodd-Frank financial regulation legislation in 2009. Cigarette companies are scaled as much more exposed to regulation after the regulatory measures and legal battles beginning in 1996 and resolving in the Supreme Court in 2000.



It is difficult to validate a new measure since there is little to compare it to. However, Figure A.11 attempts two ways to check for whether the measure appears to pick up fundamental changes in regulatory environments. In the first panel, we plot the average exposure to regulation over time for two two-digit SIC-code industries: 61 and 62, Security and Commodity Brokers and Non-depository Credit Institutions. The vertical line in the plot indicates the implementation of the Dodd-Frank Bill (HR 4173), a large-scale overhaul of financial regulation. As the plot shows, there is a steady increase in the exposure to regulation measure for these affected industries, both in the year preceding final passage of the bill and in the years after its passage. Though of course this cannot prove the validity of the measure, it is a "sanity check" which the measure passes.

The second panel in Figure A.11 presents another test of this form. The plot shows the exposure to regulation measure for the two-digit SIC industry 21, Cigarettes and Tobacco Products. The vertical line in this plot indicates the beginning of a major change in tobacco regulatory policy, starting with the FDA's assertion in 1996 of its authority over tobacco products and culminating

with the Supreme Court case *FDA v. Brown & Williamson Tobacco Corporation.* Though this court case overturned the FDA's authority to regulate tobacco companies, it ushered in an era of significant uncertainty for tobacco companies. Indeed, the next nine years were to feature repeated efforts by the U.S. government to regulate tobacco products more stringently, building to the 2009 passage of the "Family Smoking Prevention and Tobacco Control Act."[35] Again, the scaling seems to pick up this increase in exposure for cigarette companies.

These are only crude validation methods, but they suggest the scaling is detecting actual exposure to regulation.

[35]For information on this era in tobacco regulation, see for example `http://www.cancer.org/cancer/news/expertvoices/post/2012/10/29/the-fda-and-tobacco-regulation-three-years-later.aspx`.

A.4 Code to Produce Scalings

```
1
2  ##########################################
3  #### Code to Produce Exposure Scaling #####
4  ##########################################
5
6  ### change to your directory if needed
7  setwd("~/Dropbox/incumbency_campaign_finance/Donors/10k")
8
9  ### install these if you don't have them
10 library(foreign)
11 library(stringr)
12
13 ### fn that will be used to count words
14 ### uses str_count from the stringr package
15 ### return word counts
16 countWords <- function(text, words) {
17   # only want to search for matches that start at beginning of word
18   patterns <- paste("\\b", words, sep="")
19   return (
20     sapply(patterns, str_count, string=text)
21   )
22 }
23
24 # set of words we are going to count to create index
25 words <- c(
26              "protect", "pursuant", "require", "enforce",
27              "compliance", "oversight", "licens", "zoning",
28              "regulat", "law", "fine", "penalt", "politic",
29              "rule", "polic", "legislat", "commission",
30              "administration", "court", "agency", "governor",
31              "senat", "congress", "federal", "government"
32          )
```

```
33
34
35  # directory where files are held
36  dir <- "/Users/andyhall/Documents/SEC_stripped"
37
38  # read in master data
39  tot.data <- read.csv("10k_urls.csv")
40  # first, grab just the variables we need
41  total_counts <- tot.data[,c("url", "n_words")]
42  # next, split the urls to get just the filenames
43  splitted <- strsplit(as.character(as.vector(total_counts$url)), "/")
44  # we use just the file name, which is the 8th element of each splitted vector
45  file.names <- sapply(splitted, function(x) x[8])
46  # put the merge name with the word counts so we can merge in
47  files.with.counts <- cbind(file.names, total_counts$n_words)
48
49  # initialize matrix for counts
50  # will be attached to master data frame at the end
51  dtm <- matrix(nrow=nrow(files.with.counts), ncol=length(words))
52  colnames(dtm) <- words
53
54
55  # loop through the 10k files; open files and generate word counts; create a DTM
56  for (i in 1:nrow(files.with.counts)) {
57    print(i)
58    text <- readLines(paste(dir, "/", files.with.counts[i, 1], sep=""), encoding="latin1")
59    # text comes in as a list; concatenate into one character vector
60    text <- paste(text, collapse=" ")
61    if (nchar(text) > 0) {
62      # pass to countWords to create word counts for dtm
63      dtm[i,] <- countWords(text=text, words=words)
64    }
65
```

```
}

# save off raw word count data
final <- cbind(files.with.counts, dtm)
write.csv(file="raw_output.csv", final)

### produce scaling
# extract principal components from term-doc matrix
dtm2 <- dtm
dtm2[is.infinite(dtm2)] <- NA
dtm2[is.na(dtm2)] <- 0
pc.exposure <- princomp(dtm2)

# check loadings
pc.exposure$loadings

# extract loadings first principal component
# in current setup, get almost all negative loadings, so flip polarity
pc.loading <- -pc.exposure$loadings[,1]

# graph loadings
pdf(file="../loadings.pdf")
par(mar=c(5, 8, 5, 5))
barplot(pc.loading, horiz=T, las=1, xlab="Word Loadings, First Principal Component",
    col="dodgerblue", cex.axis=1.4, cex.lab=1.4)
dev.off()

# document scores = our scaling
scaling <- pc.exposure$scores[,1]

# as before, need to flip polarity
# also standardize to have mean 0 and sd 1 for interpretability
scaling <- -(scaling-mean(scaling))/sd(scaling)
```

```
98
99
100 scaling.output <- cbind(tot.data, scaling)
101
102 write.csv(file="scaling_output.csv", scaling.output)
```