The Value of Congressional Committees to Firms: Measuring Regulatory Exposure and Targeted Donations*

Aaron Rudkin† In Song Kim‡

August 24, 2026

Abstract

Why is there so little money in U.S. politics? The small sums firms give relative to the stakes suggest donations are political "consumption" rather than strategic investment. We argue that resolving this puzzle requires measuring not how much firms give, but how well they target. Existing studies match each industry to one or a few committees the research deems "relevant", underestimating the extent of strategic giving by overlooking a firm's full interests. Linking 1.6 million lobbying reports to 9,345 firms, their PACs, and donations, we trace the bills lobbied by industry peers to the committees that consider them. The result is a data-driven, continuous measure of every committee's relative importance to *any* firm, computable at the subcommittee level and at any industry aggregation. We find that firm interests span many committees, that conventional measures miss much of this exposure, and that corporate PACs target high-exposure legislators more aggressively than previously estimated.

*The authors acknowledge financial support from the National Science Foundation (SES-1725235, SES-2017315), the Russell Sage Foundation (#1908-17912). Any opinions expressed are those of the authors alone and should not be construed as representing the opinions of the Russell Sage Foundation. We thank Suyeol Yun for development of the original version of our disambiguation algorithm; Becca Sealy for assistance with legislator metadata variables; Tyler Philp for assistance maintaining the LobbyView database and identifying cases for in-text examples; and the contributions of many researchers throughout the last decade to LobbyView.

†Postdoctoral Associate, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: rudkin@mit.edu

‡Associate Professor, Department of Political Science, Massachusetts Institute of Technology, Cambridge, MA, 02139. Email: insong@mit.edu

1 Introduction

Publicly traded U.S. firms have a market capitalization of some $68 trillion (World Bank, 2026), and their fortunes turn on an enormous array of policy choices, including tax, trade, antitrust, federal procurement, labor and environmental regulations. Against that backdrop, the sums they channel through corporate PACs to influence these outcomes remain surprisingly small, in the order of $1.5 billion since 1999. That disproportion has long framed the "Tullock Paradox" in the study of money in politics: if donations are a serious instrument of influence, why don't firms spend more (Tullock, 1972; Ansolabehere, de Figueiredo and Snyder, 2003; Milyo, Primo and Groseclose, 2000)?

We suggest that the puzzle has captured the literature and unduly focused attention on topline spending, while comparatively less attention has been paid to how firms target. A dominant strand of research treats corporate giving as more like consumption or civic participation than strategic investment; the amounts are too small, the returns too uncertain, and the most credible empirical results on direct firm payoffs are small or null (Fowler, Garro and Spenkuch, 2020; Ansolabehere, Snyder and Ueda, 2004). A parallel literature, however, insists that corporate PAC money is strategic, targeting key incumbents and their institutional power (Bonica, 2013; Barber, 2016; Gordon, Hafer and Landa, 2007; Fouirnaies and Hall, 2014).

We argue that progress requires shifting attention from the question that has largely defined the field, *whether firms target*, to one that it has set aside: *how well firms target*. For firms that donate, even scarce dollars can achieve objectives when they are concentrated on the legislators who govern the firm's interests. The distinction matters because the two questions demand very different measurement strategies. Establishing that firms target at all requires only identifying one committee a firm plausibly cares about and showing that donations follow, and a large literature has done exactly this, finding consistently that money flows to the committee members who hold jurisdiction over a firm's interests (Powell and Grimmer, 2016; Barber, Canes-Wrone and Thrower, 2017; Fouirnaies and Hall, 2018; Romer and Snyder, 1994). Assessing *how well* firms target, by contrast, requires knowing the full set of committees that govern a firm and their relative importance, which is precisely what existing studies do not recover.

The challenge is that firms do not declare their interests, so researchers must infer which committees matter. The standard method, as Romer and Snyder (1994) note, requires researchers to "postulate a priori the committees that are relevant to a particular PAC or set of PACs... [i]n many cases, however, the choices are not [obvious]." In practice, studies map each firm or industry to a single committee, or at most a small number, using keyword-based coding rules: an agribusiness firm, for example, is assigned to the House and Senate Agriculture Committees, even though it clearly has interests before the tax-writing committees (Ways and Means and Finance) on taxes and tariffs, the Appropriations Committees that control federal spending, and the Energy and Commerce Committee on food safety and environmental regulation. The trouble runs deeper than any particular coding choices: such approaches rely on researcher discretion, define firm interests too narrowly, and, by treating committees as simply "relevant" or "irrelevant," cannot rank or weight them by importance. As the agribusiness example suggests, a focus on functionally responsible committees misses the long shadow cast by broadly powerful ones, and vice versa. Firms are exposed to many committees at once, and existing measures are simply not built to capture that breadth.

We overcome these limitations by letting firms reveal their own interests. The core intuition is simple: when a firm lobbies a bill, it signals a stake in the policy that bill governs, and every bill is assigned to the committees that hold jurisdiction over it. By tracing the full set of bills a firm lobbies to the committees that consider them, we recover a direct, behavioral record of which committees matter to the firm and how much, without any researcher having to guess. This is precisely the breadth that keyword approaches discard: a single firm may lobby hundreds of bills in a given Congress, with stakes that potentially reach across the 36 permanent standing committees (16 in the Senate and 20 in the House) and the roughly 180 subcommittees through which Congress legislates. It is the accumulation of these stakes, not any one "primary" committee, that defines a firm's regulatory exposure.

Executing this at scale is a formidable empirical undertaking. We develop a novel record-linkage technique to connect over 1.6 million U.S. federal Lobbying Disclosure Act reports (Kim, 2018) to 9,346 publicly traded firms, link each report to the specific bills lobbied, and map every bill to its committees of jurisdiction. Because some firms do not lobby and others lobby only minimally, we nest firms within their broader industries to borrow strength from similarly situated peers, and we measure lobbying in a temporal pre-period so that our exposure measure cannot be a downstream consequence of the donations we seek to explain. The resulting measure is continuous, ranks committees by their relative importance to any firm, and can be aggregated or disaggregated to any industry grouping, trading off smoothness against granularity. Its payoff is immediate: the median industry is exposed to 29 committees across both chambers (or 13.19 adjusting for committee importance), and the single most important committee captures less than a third of an industry's total exposure, exactly the breadth that keyword-based measures are built to ignore.

Armed with this measure, we ask whether firms are more likely to donate to legislators who sit on the committees that govern them. They are, by a margin that is both statistically and substantively large, and considerably larger than prior work has recognized. Against a baseline donation rate of roughly one percent, moving a legislator from no exposure to a high but realistic level of exposure raises the probability of receiving a firm's donation by an amount comparable to, and often exceeding, the effect of sharing the firm's home state. The relationship strengthens when we restrict attention to firms that donate at all, holds across industries and across every Congress in our data, and survives a battery of robustness checks, placebo tests, alternative aggregations, and a formal sensitivity analysis. We further corroborate it with a firm-committee-congress design that exploits quasi-exogenous turnover in committee membership, isolating variation that individual firms cannot control. Throughout, the pattern is the same: firms concentrate their scarce dollars on the legislators best positioned to act on their interests. The puzzle of why firms give "so little" begins to dissolve once we see how precisely they give it.

Our study makes three distinct contributions to the literature on money in politics, congressional committees, and corporate political strategy. First, we extend a substantial literature that examines whether political contributions strategically target committee members who have jurisdiction over donors' interests (Esterling, 2007; Hojnacki and Kimball, 2001). These studies range from broad accounts of PACs following committee assignments (Grier and Munger, 1991) to closer analyses of particular industries (Kroszner and Stratmann, 1998) or individual committees (de Figueiredo and Silverman, 2006). The most credibly identified recent work confirms the pattern: exogenous exile from a committee depresses firm PAC donations (Powell and Grimmer, 2016), interest groups target industry-relevant committees in state legislatures (Fouirnaies and Hall, 2018), and individuals direct contributions to committee members aligned with their occupational interests (Barber, Canes-Wrone and Thrower, 2017). The consistent finding is that money follows committees (Romer and Snyder, 1994). We do not dispute this conclusion; we show that it has been substantially understated. Because prior work captures only a fraction of the committees a firm cares about, it necessarily understates how aggressively firms target. Measuring exposure across the full array of committees that consider a firm's lobbied legislation, we find that corporate PACs concentrate their donations on high-exposure legislators far more precisely than previously recognized. The apparent modesty of corporate spending reflects not disengagement but surgical precision.

Second, we contribute to the literature on congressional committees and their influence over policy and political actors (Shepsle and Weingast, 1987; Krehbiel, 1992; Hall and Wayman, 1990; Berry and Fowler, 2018). This work has long recognized that committees carve out durable, overlapping jurisdictions (King, 1994) and that members compete for seats precisely because some confer greater influence than others (Groseclose and Stewart, 1998). Yet empirical studies of firm behavior have reduced this rich institutional structure to a binary, treating each committee as either relevant or irrelevant to a given firm. That framing cannot capture the simultaneous exposure of firms to many committees of varying importance, nor distinguish the broad reach of power committees from the narrower but deeper pull of functionally responsible ones. Our continuous, data-driven measure of the relative significance of every committee to every firm overcomes this limitation. By ranking and weighting committees by their importance to any firm or industry at any point in time, it enables far more precise estimation of how the committee system shapes corporate political activity, and it opens new questions about how shifts in committee jurisdiction, membership, and power feed back into firm behavior.

Third, our findings speak to a broader literature on corporate political strategy that has concentrated on explaining the *scale* of political spending: total PAC expenditure, rates of PAC formation, and aggregate corporate political activity (Bonica, 2013; Barber, 2016; Hillman, Keim and Schuler, 2004; Mellahi et al., 2016). We do not claim to resolve the longstanding puzzle of aggregate spending. Instead, we shift analytical attention to how strategically firms *allocate* what they give. Among firms that give, the precision with which they distribute donations across legislators is high, systematic, and tightly tied to each firm's regulatory exposure. This reframing bears directly on the long-standing puzzle of why there is so little money in U.S. politics (Ansolabehere, de Figueiredo and Snyder, 2003). Firms need not spend large sums if they can place small ones with great accuracy. Understanding the strategic logic of corporate political investment thus requires moving beyond the question of scale to the question of distribution, and it is there that the returns to political activity come into view.

The rest of the paper proceeds as follows. Section 2 introduces our data, record-linkage algorithm, and exposure weight calculations, and elaborates on the limitations of existing measurement approaches. We then turn to two empirical strategies for estimating the effect of committee exposure on donation behavior in Section 3. We first present our firm-legislator-congress level analysis, which exploits pre-period exposure weights to address endogeneity concerns. Next, we bolster these identification claims by exploiting quasi-exogenous variation in committee membership across congresses, yielding a firm-committee-congress level analysis. Section 4 concludes by discussing the broader implications and limitations of our findings and suggesting directions for future research.

2 Measuring Firm-Committee Exposure

This section constructs our measure of firm-committee exposure in two steps and then illustrates it. We first connect the universe of publicly traded firms to their lobbying and donations through a novel record-linkage technique, and we then trace the bills firms lobby to the committees that consider them, yielding a continuous measure of every committee's relative importance to every firm. Having built the measure, we illustrate what it captures with a single firm, the pharmaceutical company Amgen Inc., whose exposure weights and donation behavior foreshadow the central finding of the paper and motivate the formal analysis in Section 3.

2.1 From Firms to Political Activity

We begin with the universe of firms. For the purposes of this research, we consider publicly-traded firms in the S&P Compustat database. An observation in Compustat is a firm-year. We exclude listings for ETFs, pass-through trusts, and firms that do not trade on North American exchanges.[1] This results in a set of 9,345 firms. We further de-duplicate firms that are merged or acquired and subset to just those firms active in the period we study, resulting in a list of 6,562 entities.

Our objective is to connect these entities to their political activity (lobbying and donations). Doing so is non-trivial because no standardized identifier exists to connect disparate sources, and firm names and representations change over time. We develop a novel and sophisticated disambiguation method, described in Appendix A.1, which resolves and connects records with extremely high accuracy by making use of search engine response patterns. In brief, we use web search results to determine which domain names are associated with a given representation of a firm's name; when domains are shared, firms are deemed connected. Once firms have been disambiguated, we provide unique identifiers that allow researchers to connect with datasets by firm name, including attaching lobbying reports, PAC names, and canonical business datasets such as S&P Compustat and BvD/Moody's Orbis.[2]

Using the identifiers we provide, we connect firms to LobbyView, our database of over 1.6 million federal LD-2 Lobbying Disclosure Reports, each of which represents a quarterly or semi-annual record of lobbying by a registered lobbying firm on behalf of a client (Kim, 2018). We proceed by

[1] In practice, we explicitly restrict our sample to firms active in 2006 or later (due to challenges in the record-linkage procedure, described in Appendix A.1) and implicitly restrict our estimation sample to firms active in the estimation period. Within the estimation period we observe no temporal heterogeneity and so we expect that, in the presence of better data, we would find the same results further back.

[2] Our method additionally works with interest groups, unions, private companies, industrial organizations, and other firm and entity types not explored in this paper.

connecting the lobbying data in two directions: (1) connecting the lobbying report to individual bills lobbied by the firm, and from there to the congressional committees that considered these bills and the legislators serving on those committees; (2) connecting the firm to its corporate PAC(s), and from there to legislators via donations. We depict a stylized version of this process in Figure 1.

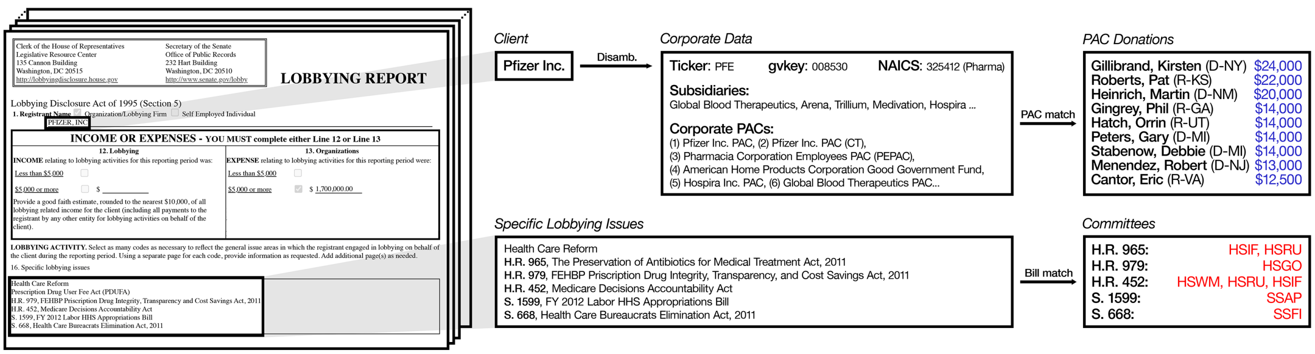

Figure 1: *Extracting Lobby Report-Committee Linkages:* This figure depicts a stylized version of our record linkage process. Client/registrant names are used to connect lobbying reports to individual firms using our novel disambiguation method. S&P Compustat provides industry, revenue, and other business data. We connect firm PACs to FEC data to observe donations associated with the firm to legislators. Simultaneously, lobbying activity text is used to extract specific bills that were lobbied. We connect these to congressional records to extract the committees associated with the bills. The illustrated example depicts Pfizer Inc.'s Q4 2011 lobbying report.

From our dataset, 3,121 of the firms listed have lobbied (i.e., appear as lobbying clients), of whom 1,726 lobbied a specific congressional bill. 1,078 have corporate PACs (of whom 1,001 make at least one donation to a legislator).

2.2 Constructing Committee Exposure Weights

Now that the data is connected, we construct our measure of interest, `firm-legislator committee exposure` ($D_{ij[f],t}$), where i indexes a legislator, t a session of Congress, and $j[f]$ denotes the nesting of firm f within industry j (we develop this notation further in Section 3). The measure captures the extent, scaled from zero to one, to which a legislator's committee portfolio overlaps with a firm's committees of interest, which we discover by observing the committees whose bills are lobbied by firms in the same industry. We therefore begin by constructing industry-committee exposure λ_{jm}.

Industry-committee exposure λ_{jm}: We construct `industry-committee exposure` in three steps. First, for each firm we list every bill it lobbied during an earlier reference period and connect each bill to the committees that considered it.[3] We assign bills to committees rather than subcommittees, though our results are robust to building weights at the subcommittee level instead (Appendix A.9.3). Second, we assign each committee a firm-level weight equal to the fraction of the firm's lobbied bill-committee pairs that the committee accounts for. Third, we average these weights across all firms in the industry, where firms are grouped by their 3-digit NAICS sub-sector. The result, λ_{jm}, is the share of industry j's lobbying activity that flows through committee m.

We deliberately measure lobbying in a historical window that precedes the donations we later study. This timing is central to our analysis: it ensures that the lobbying used to build our exposure measure cannot itself be a response to those donations, much as a shift-share design relies on pre-determined shares. The point is not incidental, since lobbying and donations are closely intertwined—firms use donations to prime connections for future lobbying (Kim, Stuckatz and Freiheyt, 2026)—and a window years earlier sidesteps this entanglement, isolating a firm's enduring regulatory interests.[4] We develop how this timing supports identification when we introduce our model in Section 3. We further provide an in-depth analysis of average committee exposure across industries in Appendix A.2 and discuss the NAICS classification used to nest firms within industries in Appendix A.4.

Constructing weights at the industry level rather than the firm level offers three advantages. First, it attenuates the endogeneity between a firm's own lobbying and its own donations, since no firm can control the lobbying behavior of all its peers. Second, it smooths the exposure mapping, which carries both measurement and conceptual benefits: we treat lobbying as a *proxy* for a firm's regulatory jurisdiction—the committees that govern the goods and services it produces—rather than as a full expression of the particularistic interests behind any single lobbying effort. This interpretation rests on a deliberately weak assumption. We require only that firms in the same industry face the same basket of committees because they produce similar goods and services; we do *not* require that they share *positions* on the underlying bills, and they often do not.[5] Two firms on opposite sides of a tariff bill, for instance, still share an interest in the bill itself, and hence in the committee that considers it. Third, the industry-level approach lets us recover exposure even

[3]We use the 105^{th}–111^{th} Congresses as our reference period, which provides the optimum trade-off between capturing a significant degree of lobbying while retaining the longest possible panel of donations to study afterward. Exact weights are sensitive to this choice, but the results reported throughout the paper are unaltered by perturbations in the choice of Congresses used to build the weights; see Appendix A.9.5.

[4]We differ from Kim, Stuckatz and Freiheyt (2026) in that we use lobbying solely to measure pre-existing patterns of industry regulatory exposure; we explore the extent of strategic donations by corporate PACs and make no claim about the temporal ordering of lobbying and donation. We are also not the first to use lobbying as a proxy for firm interests, though our measurement itself is novel (Crosson, Furnas and Lorenz, 2020).

[5]See, e.g., Rogowski (1987); Kim and Osgood (2019); Kim et al. (2019); Osgood (2018); Madeira (2016); Plouffe (2017).

for firms that rarely or never lobby.

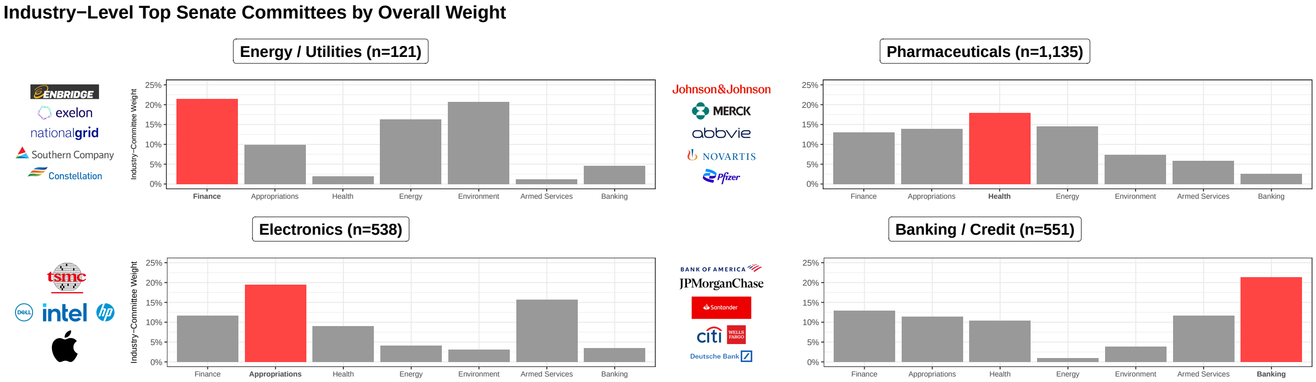

Figure 2: *Industry-Committee Exposure Weights:* This figure demonstrates our ability to map industries to the congressional committees they are exposed to. Industries are identified by their three-digit NAICS industry code. The committees listed here are seven of the most important (highest exposure) Senate standing committees for these industries. The most important committee per industry is colored in red, and histogram heights show the relative weights. The industries chosen show considerable variation in their exposure weights.

Figure 2 depicts a stylized version of our resulting exposure weights. A key takeaway is that industries differ fairly systematically in their exposure to committees, despite the concentration of power and productivity in key committees. For instance, exposure weights of Pharmaceutical/Chemical Manufacturers and Banking/Credit firms correlate at only $\rho = 0.4$.[6]

The diversity of interests cautions against approaches that presume either firm interests or that committee power is concentrated narrowly. Across industries, the top committee averages 25% of the industry's committee exposure in the House (36% in the Senate), and even the top 3 committees capture an average of only 50% (66%) of the exposure. "Power committees" identified in the literature (Groseclose and Stewart, 1998; Stewart and Groseclose, 1999) make up an average of 38% of industry-committee exposure, which, although substantial, also supports the need for our broader approach.

Because NAICS industries are hierarchically clustered, we can rely on the same method of exposure weight construction to investigate how exposure varies *within* industries (across narrower sub-segments of industry classification). We observe that narrower sub-industries often have highly different patterns of exposure to committees than their parent industries. In Figure 3, we show a stylized tree of broad (NAICS-2) industry categories, our standard sub-sector (NAICS-3) categories,

[6] We perform an analysis of the facial validity of this measure in Appendix A.3 by examining the top weighted committees for various industries and conclude that the measure reflects the expected industry interests, considering both the heterogeneity of firms within NAICS sub-sector grouping and our knowledge of responsibilities of committees and their subcommittees.

and narrower industry (NAICS-4) categories, highlighting the flexibility of the measure.

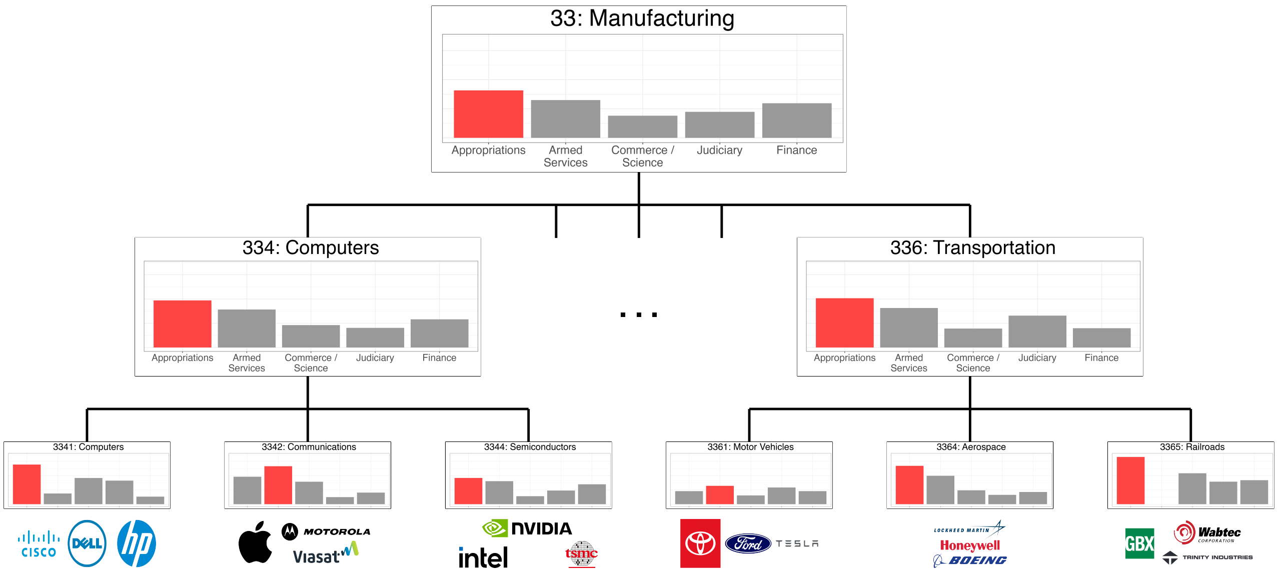

Figure 3: *Hierarchical Clustering of Exposure Weights:* This figure shows heterogeneity in exposure weighting between parent, child, and grand-child industrial groupings. The parent histogram represents the top five committees across the broader NAICS 33 (Manufacturing) sector. The children depicted are two of the many sub-sector categories nested under NAICS 33. The grandchildren are six of the narrow industries. Narrower industrial classifications trade smoothness and data density for precise estimations of firm interests.

For our primary analyses, we choose the 3-digit NAICS sub-sector code, but our results are robust to this choice. In Appendix A.9.2 we show our primary analyses using NAICS-2 and NAICS-4 weights (and in principle our method does not require NAICS at all and would work with any means of clustering firms with their peers).

Firm-Legislator Committee Exposure ($D_{ij[f],t}$): The final step in constructing our data is to move from the level of firm-committee to the level of firm-congress-legislator. In a given congress, for a given firm, each legislator is assigned a weight from 0 (no committee overlap with the firm's industry) to 1 (legislator serves on all committees lobbied by the industry) by summing the industry-committee exposure weights of each committee the legislator serves on.[7]

Thus, our unit of observation is a firm-legislator-congress (although treatment clusters at the

[7]Practical constraints discussed in Appendix A.5 limit the effective range of values from 0 to approximately 0.21 in cross-chamber weights: we discuss how to adjust effect sizes in light of this below.

industry-legislator-congress level). Our full panel consists of $n = 16,040,562$ observations (6,562 firms × 6 congresses × approximately 535 legislators per Congress, less inactive firm-congress combinations) and, as noted above, captures $30.8 billion in lobbying activity and $1.5 billion in donations.[8] Despite the large topline number of observations, donations are quite rare (mean across-firm probability of donating to a randomly selected legislator is 5%, mean probability of an individual firm-legislator-congress to contain a donation is 1.2%), and the rarity of donations at baseline helps drive a striking high-level pattern: the average probability of donating to a legislator who is serving on the firm's top committee is 9.6%, and the probability of donating to a legislator who serves on both the firm's top and second committee is 15.3%.

2.3 From Exposure to Donations: An Illustration

Having constructed our exposure measure, we now illustrate what it captures by following a single firm through the data. Amgen Inc. is a large publicly traded pharmaceutical company whose products treat migraines, arthritis, osteoporosis, complications of HIV/AIDS, cancer, and other conditions. It is a prolific donor that also actively lobbies: our data record over 1,400 donations to more than 500 recipients since the 112$^{\text{th}}$ Congress, totaling over $6.2 million, alongside more than $300 million spent lobbying 2,840 bills.[9] By virtue of its classification in NAICS industry 325 (Pharmaceutical / Chemical Manufacturing), Amgen is exposed to a range of committees, including the House Energy and Commerce Committee (10.1% weight), the Senate Committee on Health (7.3%), and House Appropriations (6.2%).

The descriptive patterns we document here, all drawn from the measure built above, foreshadow the central result that our empirical analyses in Section 3 will confirm: firms direct their donations toward the legislators who govern their interests. Members of Congress in the top quartile of Amgen's committee exposure (weights > 0.1014) are 90% more likely to receive a donation than those in the bottom quartile (weights < 0.0489), 60.1 p.p. versus 32 p.p. in real terms, and the average donation to a top-quartile member is 49% larger: $5,064 versus $3,408. Individual legislators make the pattern concrete. In the 113$^{\text{th}}$ Congress, Sen. Ed Markey ($D_{ij[f],t} = 0.180$) received $40,000 from Amgen, while Sen. Tom Udall ($D_{ij[f],t} = 0.042$) received $2,000 and Sen. Claire

[8] Due to resignations, deaths, and special elections, each Congress has between 541 and 555 legislators, of which a maximum of 535 are active at a given time. Because treatment λ is assigned at the industry level, the *effective* treatment variation is coarser; we report each model's number of observations without correcting for this, but cluster standard errors appropriately so that inference is valid. Additionally, our number of observations varies significantly by specification, as firm-level revenue data is only sporadically available via S&P Compustat. Please refer to the number of observations listed in the results tables for relevant counts. We address missingness in Appendix A.12 and do not find that strategies for addressing missingness alter results.

[9] Amgen is nonetheless broadly representative of large firms rather than exceptional: it ranks 27th among firms by number of unique recipients, 22nd by total donation amount, and 4th by the size of its largest donation.

McCaskill ($D_{ij[f],t} = 0.049$) received nothing at all.

The relationship holds not only across legislators but within them, as a legislator's exposure to Amgen rises and falls with committee service. Sen. Markey received no donations from Amgen in the 117$^{\text{th}}$ ($D_{ij[f],t} = 0.053$) or 112$^{\text{th}}$ ($D_{ij[f],t} = 0.127$) Congresses, despite his large 113$^{\text{th}}$-Congress donation. Sen. Roger Wicker, similarly, received his largest donation of $9,000 in the 115$^{\text{th}}$ Congress ($D_{ij[f],t} = 0.136$), just $1,000 in the 114$^{\text{th}}$ (0.074), and nothing in the 113$^{\text{th}}$ (0.072) or 116$^{\text{th}}$ (0.068).

Taken together, legislators substantially exposed to Amgen received over $1.81 million more from the firm during our sample period than unexposed legislators. Amgen, however, is just one of 98 Pharmaceutical / Chemical Manufacturing firms in our data. Aggregating across the industry, members in the top quartile of exposure received $22 million more from the industry's corporate PACs than those in the bottom quartile, amounting to millions of dollars per Congress. Although the pattern is not perfectly monotonic, it sets the tone for the analysis that follows: legislators with high committee exposure are far more likely to receive donations, and far larger ones, than legislators with low exposure. We now turn to the empirical analysis that establishes this relationship more systematically.

3 Estimating the Effect of Committee Exposure on Donations

The descriptive patterns above suggest that firms concentrate their donations on the legislators who govern their interests, but they cannot rule out the alternative explanations a formal design must address. We now estimate the effect of committee exposure on donations through two complementary strategies. The first operates at the firm-legislator-congress level, using exposure weights built in a deep historical window to break the contemporaneous link between a firm's lobbying and its giving. The second moves to the firm-committee-congress level and exploits quasi-exogenous turnover in committee membership, isolating variation that no individual firm can control. The two designs rest on different assumptions and different sources of variation, yet reach the same conclusion: firms target high-exposure legislators far more aggressively than prior work has recognized.

3.1 A Firm-Legislator-Congress Design

Our primary design takes the firm-legislator-congress as the unit of analysis and asks whether a firm is more likely to donate to a given legislator as that legislator's committee portfolio comes to overlap more heavily with the firm's regulatory interests. The central inferential challenge is that committee exposure is not randomly assigned: the same forces that draw a firm's interest to a set

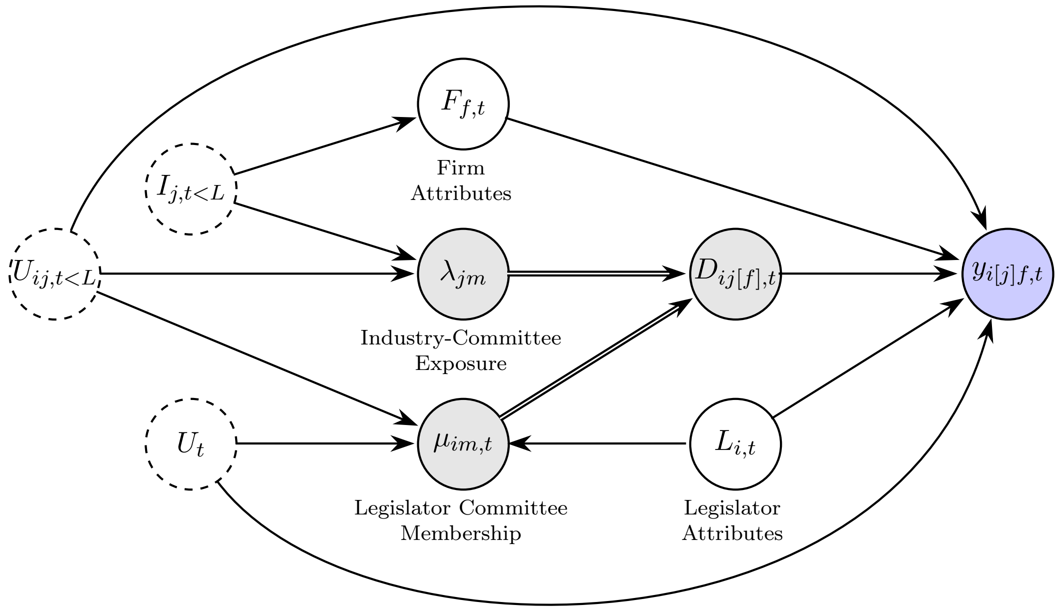

Figure 4: *Data-Generating Process DAG:* This figure depicts a directed acyclic graph describing the major variables at play in our analysis. Our identification relies on leveraging exogenous industry-level pre-period lobbying activity to determine committee exposure, and adjusting for appropriate confounders. i denotes legislator, j industry, f firm, $j[f]$ the nesting of firm in industry, m committee, t time, and $t < L$ an indicator of a lagged period. Double edges indicate that treatment $D_{ij[f],t}$ is a deterministic interaction of λ_{mj} and $\mu_{im,t}$.

of committees may also shape which legislators it chooses to fund, raising the prospect that any observed correlation reflects confounding rather than strategic targeting. To reason transparently about these threats, we represent the data-generating process as a directed acyclic graph (DAG), which makes explicit the variables that determine both committee exposure and donation behavior and, in turn, identifies the confounders we must adjust for to recover a credible estimate. The DAG also clarifies why the timing of our exposure measure matters: because the weights are built from lobbying in a deep historical window, several of the most worrying back-door paths are blocked by construction. We present the DAG in Figure 4 and then translate it into an estimating equation.

This, in turn, supplies the workhorse model for this analysis: a linear probability model with 2-way fixed effects (TWFE[10]) using the following specification:

[10]TWFE estimators have been criticized because they recover a weighted average of many implicit comparisons, with some comparisons receiving negative weights. This property can yield biased results (see e.g., Imai and Kim (2019); Goodman-Bacon (2021); De Chaisemartin and d'Haultfoeuille (2023); Chiu et al. (2026)). Most modern heterogeneous treatment effect estimators have identifying assumptions that require never-treated units, and many have implementations that require binary treatments. We estimate our primary models in Appendix A.8.5 using the estimator from De Chaisemartin and d'Haultfoeuille (2026), which identifies treatment based on "switchers" and supports data with our structure; although the estimand differs somewhat, the results are positive, significant, and comparable to our primary results.

$$y_{ij[f],t} = \tau \times D_{ij[f],t} + \beta \times \mathbf{X} + \gamma_t + \delta_{j[f]} + \epsilon_{if,t}$$

where the unit of analysis is a firm-legislator-congress. Our dependent variable, $y_{i[j]f,t}$ encodes whether firm f has donated to congressperson i during congress t as a binary 0/1 variable. A secondary model considers the logged dollar amount of donations instead. $D_{ij[f],t}$ is the firm-legislator committee exposure weight (bounded between 0 and 1); the subscript $j[f]$ refers to the Industry j of the firm f (coded by NAICS 3-digit industry code).

Weights are constructed as specified above. $D_{ij[f],t} = \sum_m \mu_{im,t} \times \lambda_{j[f]m}$ where m indexes a committee, $\mu_{im,t} = \mathbf{1}\{i \in m \,\text{at}\, t\}$ is a membership function that takes the value 1 if legislator i is a member of committee m in congress t, and $\lambda_{j[f]m}$ is the industry-committee exposure weight (indexed by industry and committee).

$\mathbf{X}$ is a matrix of potentially confounding variables (and β their corresponding coefficients). The directed acyclic graph informs our choice of confounders. Since any confounder must correlate with both donation behavior and one or both of the inputs into the exposure weights, the controls consist of legislator attributes $L_{i,t}$, firm attributes $F_{f,t}$, and dyadic attributes $z_{if,t}$. Legislator attributes include gender; state; chamber; party; party majority status (Cox and Magar, 1999); dummies if the legislator is a committee chair or a member of party leadership (Fouirnaies, 2018); past electoral experience (Jacobson and Kernell, 1983); length of service (Kroszner and Stratmann, 2005); electoral margin of victory (Fouirnaies and Hall, 2014; Fenno, 1973); legislator 1$^{\text{st}}$ dimension NOMINATE (Lewis et al., 2019) score entered as linear and quadratic terms). Firm attributes include logged revenue, logged capital expenditures, logged employee headcount, and value added productivity per employee[11]). Dyadic attributes include legislator-firm state match, and in some specifications prior observed donations and legislator-firm region match. We include firm attributes because, as the DAG suggests, unobserved pre-period industry level confounders may affect present firm donation behavior, but only through the firm. The included control variables are described in greater detail in Appendix A.6.

γ_t is a per-Congress fixed effect corresponding to the U_t time-specific confounding in the DAG. $\delta_{j[f]}$ is a per-industry fixed effect (in alternative specifications, it is replaced by δ_f, firm fixed effects), which further restricts the backdoor path between the unobserved industry confounders, industry-committee exposure, and donation behavior. $\epsilon_{if,t}$ is an error term, clustered by industry (or by firm in alternative specifications). Finally, τ is the treatment effect of interest.

[11] $VAPL = (Rev - CapEx)/(||Employees||)$

Our identification strategy relies on the fact that our exposure weights are constructed in a deep pre-treatment period using industry-wide behavior. This strategy mitigates the greatest risk of contemporaneous endogeneity between firm lobbying and firm donation behavior. The inclusion of other similarly situated firms as a measurement proxy further attenuates back-door confounding through firm-specific characteristics. The fixed-effect structure guarantees that we analyze within-Congress, within-industry variation (and in alternative specifications, within-legislator or within-firm; see Appendix A.8.4).

In order to jeopardize our identification, a confounder needs to, in some sense, cause both industry-wide lobbying behavior prior to 2011 (or whichever reference pre period a specification uses) or legislator committee assignment in the reference congress *and* firm-level donation behavior in the reference congress; vary within a congress-industry (or congress-firm, depending on specifications); vary across time to survive the time fixed effects; and not be captured or proxied by the observable characteristics of legislators or firms.

As a bounding exercise, we perform a formal sensitivity analysis in the **Robustness of Results** subsection below. Any confounder capable of attenuating our results to zero would have to be substantially more powerful than the most powerful confounders identified and controlled-for; moreover, even the most likely confounders demonstrate little confounding, precisely because of the estimation strategy.

Our results are presented in Table 1. The first three models run the above workhorse specification on different subsets of the data: Model 1 represents our full panel. Model 2 subsets our panel to first-term legislators only. We present Model 2 to reduce the risk that our $L_{i,t}$ (legislator attributes) fails to close the backdoor path between legislator-committee membership and donation outcomes due to nuances of a legislator's career development. Model 3 subsets our panel to only those firms that have at least one donation in our data ("active" firms). There are many firms whose $y_{i[j]f,t} = 0$ for all observations. This may occur because a firm is categorically unwilling to donate (or does not donate using corporate PACs) or else because a firm has not yet donated but could under some set of circumstances. In our primary model, the inclusion of these firms mechanically attenuates the effect (in a manner analogous to "never-takers" versus "compliers" in intent-to-treat terminology); in Model 3, excluding them mechanically inflates the effect (because some excluded firms *could* be coerced to donate and simply have not been by the mechanism we propose). Model 4 estimates an alternative dependent variable, logged dollar amount, subsetting to just those observations that make donations (and thus estimating the *intensive* margin of the treatment effect).[12]

Dependent Variables:	Donated?			Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.067***	0.052***	0.386***	0.793***
	(0.012)	(0.013)	(0.049)	(0.133)
Same State?	0.030***	0.033***	0.148***	0.330***
	(0.006)	(0.007)	(0.018)	(0.051)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	10,068,788	1,831,230	1,838,416	137,957

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table 1: *Effect of Committee Exposure on Donations:* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level including legislator and firm controls and two-way industry and congress fixed effects. Model 2 repeats the estimation for first-term congresspersons. Model 3 repeats the estimation for firms that have at least one donation. Model 4 uses a logged dollar amount dependent variable and estimates on the subsample of rows with donations recorded. All results are large and statistically and substantively significant.

In our primary specification, Model 1, the coefficient of interest is 0.067, meaning that a one unit change in the committee weight variable — a legislator going from having zero committee exposure to being a member of each committee that a company's industry lobbies — produces a 6.7 percentage point increase in the likelihood of donating to a legislator, ceteris paribus. We contextualize the size of this movement in an effect plot, Figure 5, where we show the marginal effect of moving from 0 exposure weight through each tercile of exposure weight. As the distribution of weights is highly skewed towards zero, much of the movement in predicted probability occurs above the median weight.

Returning to our model results, Model 2 yields comparable results to Model 1. In Model 3,

[12] In cases where firms donate multiple times to the same legislator, we aggregate to the firm-legislator-congress level and sum the donations, including in the Amgen example above. We also run this model on the full data in Appendix A.8.2, using Poisson regression to avoid distorting results per Chen and Roth (2024); Mullahy and Norton (2024), though conflating the intensive and extensive margins makes interpretation elusive. The effect is positive, significant, and robust.

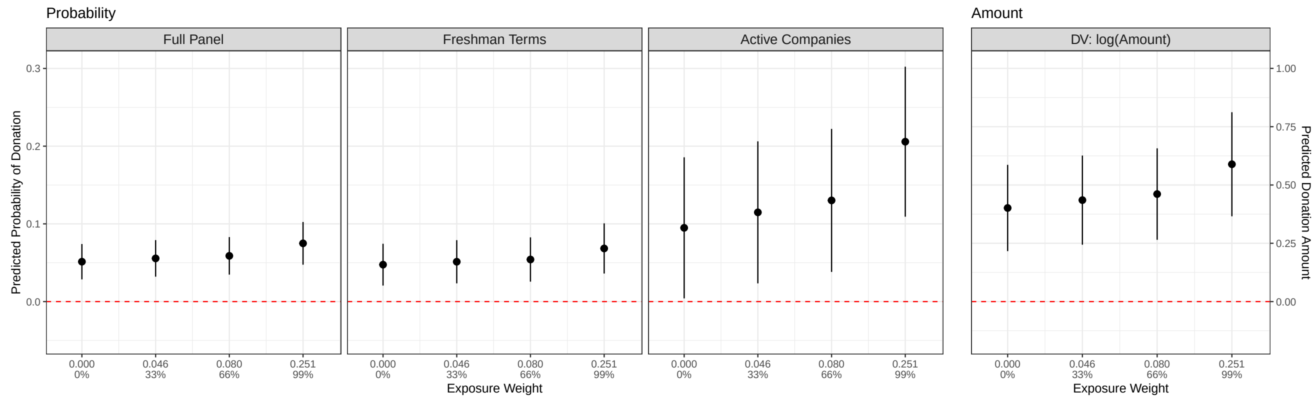

Figure 5: *Marginal Effect of Treatment:* This figure depicts the implied effect of moving from 0 committee-exposure weight through each tercile of the variable (33%, 66%, 99%), to a maximum of weight 0.253 on the predicted probability of donation, holding all other variables at their mean or mode. In comparison to the baseline predicted probability of donation, our effect sizes are positive, large, and significant. From left to right, panels report: results of our full panel; our panel restricted to Members of Congress in their first terms; politically active companies online; and our full panel using donation amount as a dependent variable.

which considers only politically active companies, a feasible effect size is 8.5 percentage points compared with a baseline donation rate of 7.4%. The restricted sample in Model 4, which considers variation in donation size, implies a 121% increase in donation size ($e^{0.793} - 1 = 1.21$), conditional on donating.

Restricting Effect Size: The regression coefficient reflects the treatment effect of moving from zero exposure weight to full exposure weight; this is an infeasibly large treatment effect (because legislators are restricted in the number and sort of committees they can serve on simultaneously), so we propose shrinking the estimate to reflect feasibility. We assess the range of the treatment variable in Appendix A.5 and conclude that a change of 0.21 units is a reasonable benchmark for a large but realistic increase in exposure weight. Thus, a legislator moving from no exposure to a high but reasonable exposure has a 1.5 percentage point increase (i.e., 0.067×0.219) in the probability of receiving a donation from a firm. The size of this effect is best understood in light of the very low overall probability of donation: only about 1.2% of observations (firm-legislator-congress level) record a donation. We thus consider the effect to be quite large indeed.[13]

Industry-Level Heterogeneity: We report the heterogeneity of our calculated effects by

[13] It is also useful to compare the effect size to a strong confounder: whether the firm and legislator are located in the same state, which also proxies for much of the idiosyncratic homophily between firms and legislators. In all cases, the committee effect is larger in raw terms than the same-state effect and is not appreciably smaller when reduced to the feasible effect.

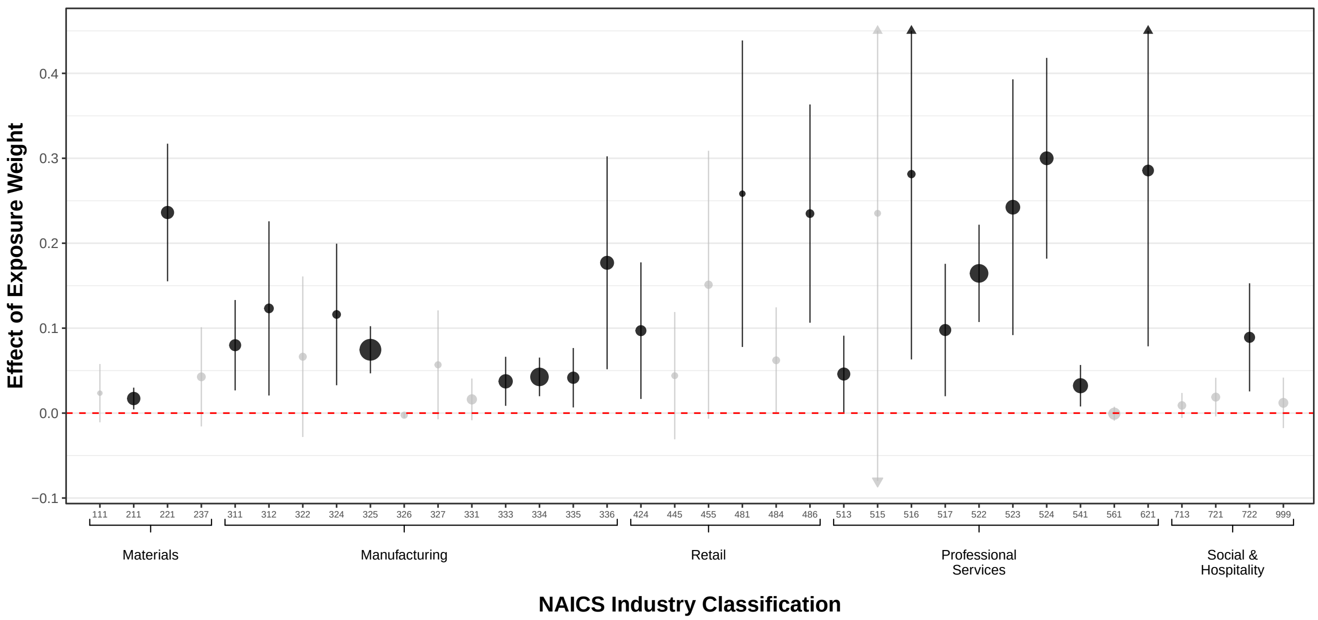

Figure 6: *Effect of Committee Exposure by Industry*: This figure depicts our main effect disaggregated by industry. Industries are ordered by NAICS code. Points are scaled by the square root size of the industry in number of firms. Industries with individually significant effects are rendered in blue, and industries with non-significant effects in gray and lower opacity. All industries have positive coefficients, and a majority are significant. Industries with wide outlier confidence intervals are truncated on the y axis for visual clarity. Effects in the professional services sector are larger than those in the manufacturing sector.

industry, both as a measure of robustness across our data and to comment on which industries are more "strategic" in their donation behavior (i.e., have a higher effect of exposure on donation behavior). In Figure 6, we show the results of our model disaggregated by industry (excluding industries with very few firms for visual clarity and clustering by firm). In brief, all but two industries have a positive effect size coefficient (two have nearly zero negative coefficients). Smaller industries are, in some cases, not individually significant due to power considerations. Professional service industries (law, IT, finance, real estate) have larger average effect sizes than manufacturing. There is no observable relationship between industry size and effect size ($\rho = -0.022$).

Placebo Test: One immediate threat to our inference occurs if the effect is produced by a small number of powerful committees that have high exposure across many industries and attract donations for reasons that are completely non-industry specific. If this were true, the results would mean only that powerful legislators attract more donations, and nothing about heterogeneous exposure or firm-/industry-specific strategy. Above, we noted that inter-industry weight correlation was only moderate and that "power committees" in the political science literature account for a minority of all committee exposure. However, we also directly test the potential of this threat by running a placebo analysis in which industry exposure weights are randomly assigned to other industries. The resulting placebo mean, although non-zero, is greatly attenuated from the observed effect. The two-sided permutation p-value for the observed effect is $p < 0.001$ (i.e., the observed effect is more extreme than all observed placebo iterations). We present the placebo distribution graphically in Appendix A.7.

Other Robustnesss Results: We investigate restricting our treatment variable to act more like conventional measures of industry interest (Appendix A.9). In Appendix A.9.1, we restrict the number of committees permitted to contribute to the exposure weights. Restricting committee exposure weights to just the firm's top committee allows us to replicate the coding strategy in Barber, Canes-Wrone and Thrower (2017), albeit with a data-driven assignment of industries to committees and using more granular industry classifications. We find that restricting the number of committees to ten, five, or one still produces a positive and significant effect, but one that is significantly lower in magnitude than our approach. We also run an analysis in which we consider only those observations in which a legislator serves on a firm's primary committee of interest (and thus all variation across legislators or periods is identified solely by their membership in *additional* committees to which the firm is exposed): effects remain large and generally significant despite reduced power (Appendix A.8.3). These results combine to indicate that prior work vastly understates the extent to which strategic committee membership, understood properly, drives firm donation behavior. Our effect also persists when considering exposure weights at the subcommittee level (Appendix A.9.3) or measuring weights per-chamber instead of spanning across chambers (Appendix A.9.4).

Nor are our results driven by the choice to use a fixed pre-period in identifying industry interest. In Appendix A.10, we use a variety of rolling pre-periods. Whether we use a fixed pre-period (105^{th}–111^{th}), a pre-period based on a fixed delay from the reference congress (and whether this delay is short or long in relative terms), or a pre-period where newer congresses take advantage of more pre-period data (105^{th}–$(\text{n-1})^{\text{th}}$), the result is unchanged. This bolsters our defense against inferential threats from idiosyncratic time-varying pre-period shocks and confirms that, to the extent firm interests are not static, they still donate strategically in light of those changing interests.

We also test excluding smaller industries where individual firms may be "price-makers" because they make up large portions of the exposure weight construction (Appendix A.11), using interpolation or multiple imputation to resolve data missingness (Appendix A.12), changing fixed effect structures (Appendix A.8.4), running a Poisson regression on the combined intensive-extensive margin donation amount effect (Appendix A.8.2), employing a modern heterogeneous treatment estimator in Appendix A.8.5, as well as employing non-linear (logistic) functional forms (Appendix A.13, A.13.2). Across all specifications, our findings are robust and unchanged.

Sensitivity Analysis: Although we have made efforts to defend against the above listed threats, there may still be unobserved confounding that threatens our findings. To assess the risk of this, we perform a sensitivity analysis using the method in Cinelli and Hazlett (2020). Here, we specify benchmark confounders (firm-legislator same-state; legislator is a committee chair; legislator is a party leader; and quadratic 1^{st} dimension NOMINATE score) which are reasonably expected to be strong confounders. We report the extent to which an omitted confounder would have to correlate with treatment, outcome, or both in order to attenuate our observed effect to be non-significant. We find two things: first, the benchmark confounders are not very strong confounders. Party leadership correlates weakly with both treatment and outcome; NOMINATE and legislator-firm state accord correlate strongly with outcome but weakly with treatment; and committee chair status correlates strongly with treatment but weakly with outcome. This result validates our measurement and identification strategy. Second, an unobserved confounder even ten or twenty times larger than these benchmark confounders on either dimension would not attenuate our effect.

3.2 Leveraging Committee Turnover

Our primary analysis relies on the firm-legislator dyad, which may be vulnerable to the threat of unobserved firm-legislator-level confounders. Here, we introduce a firm-committee-congress level of analysis. We use industry-committee exposure weights as our primary treatment variable but interact the treatment with a source of quasi-exogenous variation: committee turnover. Committee turnover can come from several sources, including electoral replacement, exile (removal, typically due to a change in partisan control of the chamber), retirement, or death. Each committee change, in turn, reshuffles other committees in increasingly chaotic ways. Thus, even if a firm has the capability to influence the election or appointment of one member to one committee, overall turnover, aggregated to the committee level and across all committees, is far outside the firm's remit. Nevertheless, turnover in a committee of interest potentially jostles the committee equilibrium by shifting the location of pivotal voters or the "gridlock interval" (Krehbiel, 2010; Tsebelis, 2002). Thus, the treatment in this analysis involves the interaction of a firm- (industry-) specific pre-period share with a committee-specific reference-period shift in membership and thus functions analogously to a shift-share-style reduced form. The exact specification is, then:

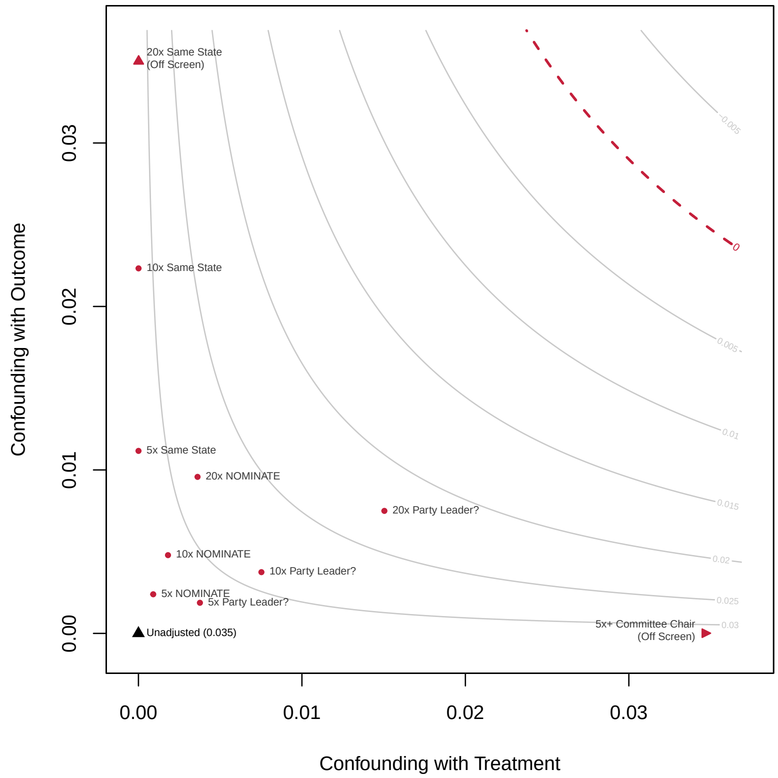

Figure 7: *Sensitivity Analysis*: This figure shows the results of our sensitivity analysis. The unadjusted estimate is shown in the bottom left corner. Other points show how the estimate would attenuate in the presence of a confounder n times "stronger" than our benchmark confounders. The contours show how the effect would attenuate and the red dashed contour denotes the attenuation of the effect to zero. The benchmark confounders are fairly weak and load primarily on one axis (confounding with either treatment or outcome, but not both), indicating that it is unlikely that omitted confounders would break the observed relationship.

$$y_{fm,t} = \tau \times (\lambda_{j[f]m} \times s_{mt}) + \beta \times \mathbf{X} + \eta_m + \gamma_t + \delta_{j[f]} + \epsilon_{fm,t}$$

Our dependent variable, $y_{fm,t}$, encodes whether firm f has donated to someone on committee m during congress t as a binary $0/1$ variable. Note that the donation need not be to the legislator impacted by the turnover. A secondary model considers the logged dollar amount of donations instead. $\lambda_{j[f]m}$ is the industry-committee exposure weight (bounded between 0 and 1); the subscript $j[f]$ refers to the Industry j of the company f (coded by the NAICS 3-digit industry code). s_{mt} is the extent of turnover in the committee in congress t, measured as a count (the number of new members on the committee) or proportion (committees vary in size). $\mathbf{X}$ in these specifications includes all firm-level confounders described above but excludes the legislator and dyadic confounders.

We present the primary results in Table 2. Models 1 and 2 show the effect on the probability of donations: Model 1 uses a proportion-based turnover calculation, while Model 2 uses a count-based turnover calculation. Models 3 and 4 show the effect on the (logged dollar) amount of donations, again using proportion-based and count-based turnover calculations.

Dependent Variables:	Donated to committee		Amount (log 1+p)	
	Donated		Amount (log)	
Model:	(1)	(2)	(3)	(4)
Variables				
Treatment (proportion)	0.301***		4.10***	
	(0.087)		(0.841)	
Treatment (count)		0.009***		0.119***
		(0.002)		(0.022)
Firm Characteristics	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Committee	Yes	Yes	Yes	Yes
Fit statistics				
Observations	617,112	617,112	616,867	616,867

Clustered (id_firm_ committee) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 2: *Firm-Committee "Shift-Share" Reduced Form (Main):* This model presents our primary regression using firm-committee level data, where firm exposure to the committee is interacted with exogenous turnover on the committee. Treatment variables can use either turnover count or turnover proportion. Standard errors are clustered by firm-committee.

The results are again positive and significant. The feasible coefficient sizes of these results are more difficult to interpret because of their interactive nature. Keeping with our prior analysis, we report the 2.5% and 97.5% quantiles of the treatment variables: the treatments interacted with count-based turnover measures have a feasible range between 0 and 1.28 (implying a 1.15 percentage point increase in the probability of donation), while the treatments interacted with proportion-based turnover measures have a feasible range between 0 and 0.027 (implying a 0.8 p.p. increase in the probability of donation). Again, given low base rates of donation, these feasible ranges yield quantitatively large increases in donation propensity. As with our primary results, our results here are robust against a variety of alternative specifications, including alternative fixed effect structures (Appendix A.14.1), disaggregated by session of Congress (Appendix A.14.2), and when estimated using a non-linear (logistic) link function (Appendix A.14.3).

4 Concluding Remarks

We introduce a novel measure of firm-level regulatory exposure that relies on three key intuitions: (1) that a firm's regulatory exposure is captured by the committees that have jurisdiction over the bills a firm lobbies on; (2) that for low- or no-activity political firms, this exposure can still be measured by the activity of peer firms in the same industry; and (3) that firm interests are much broader than a single committee. The measure we propose, based on the NAICS hierarchical industry classification system, is highly flexible and allows us to capture firm interests at any level of granularity, trading off specificity and support. Our method extends naturally to non-NAICS measurement of firm clustering. This measure overcomes limitations of prior work, which relied on research-driven assignments of firms to single (or small numbers of) committees via researcher discretion or keywords. Our results show that although historical practice produced positive effect sizes, it understated the magnitude of the effect. Firms are typically exposed to many committees, and primary committees of interest make up no more than one quarter to one third of a firm's interests. We calculate exposure in a held-out pre-period to guarantee exogenous variation in treatment. We find that firms are highly strategic in their donation behavior. The effect of committee membership on donations is very large relative to the baseline, robust across specifications and samples, and proves larger than benchmark confounders. Across sectors, professional services firms have higher coefficients, which likely reflect the dynamic regulatory environment for service industries.

Although we expand the literature's measure from one-firm-one-committee to one-firm—many-committees, we are still limited by situating firms in individual industrial categories (though our industrial categories are larger in number than past work). Our exposure weights are coarsened by this choice. Large firms today routinely produce a variety of goods or supply a variety of services, and so NAICS codes are necessarily a simplifying assumption of complex firm behavior. The nestedness of the industry construction mitigates this problem (to the extent that a firm is part of two or more industries, the industries are likely to be "cousins" rather than "strangers"), but an ideal future measure should be able to map firms to industries (or peer firms) in a more flexible way. Finally, our current analysis is based solely on firm donations via corporate PACs, excluding the contributions of "personal" donations by executive leadership (Fremeth, Richter and Schaufele, 2013; Babenko, Fedaseyeu and Zhang, 2020; Stuckatz, 2022*b*,*a*; Bonica, 2016), and so may, if anything, continue to understate the magnitude of the relationship by missing out on even more strategic donation behavior.

The effects we observe are large in magnitude whether we consider the entire dataset, or just firms that donate at all. But one remaining question is unanswered: why do so many firms choose not to donate at all, even when similarly situated to donor firms? Future work is needed to situate the never-donor firms and do more to understand why they abstain.

Our findings unearth a much richer ecosystem of firm strategic behavior and so help resolve an apparent paradox in the money-in-politics literature. We provide a road-map for complex record-linkage problems and demonstrate the value of harnessing data-driven measurement processes to capture large-scale behavior in political science. But our findings also have implications for legislators and voters alike. For legislators, committee participation provides an extremely valuable channel through which to accrue material support for re-election. For voters, the full magnitude of firm strategic donations may heighten concerns about the disproportionate influence of firms on political outcomes in the United States (Schlozman, Verba and Brady, 2012; Gilens and Page, 2014).

References

Ansolabehere, Stephen, James M Snyder and Michiko Ueda. 2004. "Did firms profit from soft money?" *Election Law Journal: Rules, Politics, and Policy* 3(2):193–198.

Ansolabehere, Stephen, John M de Figueiredo and James M Snyder. 2003. "Why is there so little money in U.S. politics?" *Journal of Economic Perspectives* 17(1):105–130.

Arora, Abhishek and Melissa Dell. 2023. "LinkTransformer: A Unified Package for Record Linkage with Transformer Language Models.".

Babenko, Ilona, Viktar Fedaseyeu and Song Zhang. 2020. "Do CEOs affect employees' political choices?" *The Review of Financial Studies* 33(4):1781–1817.

Barber, Michael J. 2016. "Donation motivations: Testing theories of access and ideology." *Political Research Quarterly* 69(1):148–159.

Barber, Michael J, Brandice Canes-Wrone and Sharece Thrower. 2017. "Ideologically sophisticated donors: Which candidates do individual contributors finance?" *American Journal of Political Science* 61(2):271–288.

Berry, Christopher R and Anthony Fowler. 2018. "Congressional committees, legislative influence, and the hegemony of chairs." *Journal of Public Economics* 158:1–11.

Bonica, Adam. 2013. "Ideology and interests in the political marketplace." *American Journal of Political Science* 57(2):294–311.

Bonica, Adam. 2016. "Avenues of influence: On the political expenditures of corporations and their directors and executives." *Business and Politics* 18(4):367–394.

Chen, Jiafeng and Jonathan Roth. 2024. “Logs with zeros? Some problems and solutions.” *The Quarterly Journal of Economics* 139(2):891–936.

Chiu, Albert, Xingchen Lan, Ziyi Liu and Yiqing Xu. 2026. “Causal panel analysis under parallel trends: lessons from a large reanalysis study.” *American Political Science Review* 120(1):245–266.

Cinelli, Carlos and Chad Hazlett. 2020. “Making sense of sensitivity: Extending omitted variable bias.” *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1):39–67.

Cox, Gary W and Eric Magar. 1999. “How much is majority status in the U.S. Congress worth?” *American Political Science Review* 93(2):299–309.

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

Dal Bó, Ernesto and Frederico Finan. 2018. “Progress and perspectives in the study of political selection.” *Annual Review of Economics* 10(1):541–575.

De Boor, Carl. 1978. *A practical guide to splines.* Vol. 27 springer New York.

De Chaisemartin, Clément and Xavier d’Haultfoeuille. 2023. “Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey.” *The Econometrics Journal* 26(3):C1–C30.

De Chaisemartin, Clément and Xavier d’Haultfoeuille. 2026. “Difference-in-differences estimators of intertemporal treatment effects.” *Review of Economics and Statistics* pp. 1–18.

de Figueiredo, John M and Brian S Silverman. 2006. “Academic earmarks and the returns to lobbying.” *The Journal of Law and Economics* 49(2):597–625.

Esterling, Kevin M. 2007. “Buying expertise: Campaign contributions and attention to policy analysis in congressional committees.” *American Political Science Review* 101(1):93–109.

Fenno, Richard F, Jr. 1973. *Congressmen in Committees.* University of California Institute of Governmental Studies.

Fernández-Val, Iván and Martin Weidner. 2016. “Individual and time effects in nonlinear panel models with large N, T.” *Journal of Econometrics* 192(1):291–312.

Fouirnaies, Alexander. 2018. “When are agenda setters valuable?” *American Journal of Political Science* 62(1):176–191.

Fouirnaies, Alexander and Andrew B Hall. 2014. “The financial incumbency advantage: Causes and consequences.” *The Journal of Politics* 76(3):711–724.

Fouirnaies, Alexander and Andrew B Hall. 2018. “How do interest groups seek access to committees?” *American Journal of Political Science* 62(1):132–147.

Fowler, Anthony, Haritz Garro and Jörg L Spenkuch. 2020. “Quid pro quo? Corporate returns to campaign contributions.” *The Journal of Politics* 82(3):844–858.

Fremeth, Adam, Brian Kelleher Richter and Brandon Schaufele. 2013. “Campaign contributions over CEOs’ careers.” *American Economic Journal: Applied Economics* 5(3):170–188.

Gilens, Martin and Benjamin I Page. 2014. "Testing theories of American politics: Elites, interest groups, and average citizens." *Perspectives on Politics* 12(3):564–581.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225(2):254–277.

Gordon, Sanford C, Catherine Hafer and Dimitri Landa. 2007. "Consumption or investment? On motivations for political giving." *The Journal of Politics* 69(4):1057–1072.

Grier, Kevin B and Michael C Munger. 1991. "Committee assignments, constituent preferences, and campaign contributions." *Economic Inquiry* 29(1):24–43.

Groseclose, Tim and Charles III Stewart. 1998. "The value of committee seats in the House, 1947-91." *American Journal of Political Science* pp. 453–474.

Hall, Richard L and Frank W Wayman. 1990. "Buying time: Moneyed interests and the mobilization of bias in congressional committees." *American Political Science Review* 84(3):797–820.

Heckman, James J. 1978. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process and Some Monte Carlo Evidence." University of Chicago, Social Systems Research Institute, Report 78-10.

Hillman, Amy J, Gerald D Keim and Douglas Schuler. 2004. "Corporate political activity: A review and research agenda." *Journal of Management* 30(6):837–857.

Hojnacki, Marie and David C Kimball. 2001. "PAC contributions and lobbying contacts in congressional committees." *Political Research Quarterly* 54(1):161–180.

Imai, Kosuke and In Song Kim. 2019. "When should we use unit fixed effects regression models for causal inference with longitudinal data?" *American Journal of Political Science* 63(2):467–490.

Jacobson, Gary C and Samuel Kernell. 1983. "Strategy and choice in congressional elections.".

Kim, In Song. 2018. "LobbyView: Firm-level Lobbying & Congressional Bills Database." Working Paper.
URL: *http://web.mit.edu/insong/www/pdf/lobbyview.pdf*

Kim, In Song, Helen V Milner, Thomas Bernauer, Iain Osgood, Gabriele Spilker and Dustin Tingley. 2019. "Firms and Global Value Chains: Identifying Firms' Multidimensional Trade Preferences." *International Studies Quarterly* 63(1):153–167.

Kim, In Song and Iain Osgood. 2019. "Firms in trade and trade politics." *Annual Review of Political Science* 22:399–417.

Kim, In Song, Jan Stuckatz and Lukas Wolters Freiheyt. 2026. "Systemic and sequential links between campaign donations and lobbying." *The Journal of Politics* 88(1):110–124.

King, David C. 1994. "The nature of congressional committee jurisdictions." *American Political Science Review* 88(1):48–62.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing incomplete political science data: An alternative algorithm for multiple imputation." *American Political Science Review* 95(1):49–69.

Krehbiel, Keith. 1992. *Information and Legislative Organization*. University of Michigan Press.

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

Kroszner, Randall S and Thomas Stratmann. 1998. "Interest-group competition and the organization of Congress: Theory and evidence from financial services' political action committees." *American Economic Review* 88(5):1163–1187.

Kroszner, Randall S and Thomas Stratmann. 2005. "Corporate campaign contributions, repeat giving, and the rewards to legislator reputation." *The Journal of Law and Economics* 48(1):41–71.

Lewis, Jeffrey B, Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin and Luke Sonnet. 2019. "Voteview: Congressional roll-call votes database.".
URL: *https://voteview.com*

Madeira, Mary Anne. 2016. "New trade, new politics: Intra-industry trade and domestic political coalitions." *Review of International Political Economy* 23(4):677–711.

Mellahi, Kamel, Jędrzej George Frynas, Pei Sun and Donald Siegel. 2016. "A review of the nonmarket strategy literature: Toward a multi-theoretical integration." *Journal of Management* 42(1):143–173.

Milyo, Jeffrey, David Primo and Timothy Groseclose. 2000. "Corporate PAC campaign contributions in perspective." *Business and Politics* 2(1):75–88.

Mullahy, John and Edward C Norton. 2024. "Why transform Y? The pitfalls of transformed regressions with a mass at zero." *Oxford Bulletin of Economics and Statistics* 86(2):417–447.

Osgood, Iain. 2018. "Globalizing the supply chain: Firm and industrial support for US trade agreements." *International Organization* 72(2):455–484.

Pepinsky, Thomas B. 2018. "A note on listwise deletion versus multiple imputation." *Political Analysis* 26(4):480–488.

Plouffe, Michael. 2017. "Firm heterogeneity and trade-policy stances: Evidence from a survey of Japanese producers." *Business and Politics* 19(1):1–40.

Powell, Eleanor Neff and Justin Grimmer. 2016. "Money in exile: Campaign contributions and committee access." *The Journal of Politics* 78(4):974–988.

Rogowski, Ronald. 1987. "Political cleavages and changing exposure to trade." *American Political Science Review* 81(4):1121–1137.

Romer, Thomas and James M Snyder. 1994. "An empirical investigation of the dynamics of PAC contributions." *American Journal of Political Science* 38(3):745–769.

Schlozman, Kay Lehman, Sidney Verba and Henry E Brady. 2012. *The Unheavenly Chorus: Unequal Political Voice and the Broken Promise of American Democracy*. Princeton University Press.

Shepsle, Kenneth A and Barry R Weingast. 1987. "The institutional foundations of committee power." *American Political Science Review* 81(1):85–104.

Stewart, Charles III and Tim Groseclose. 1999. "The value of committee seats in the United States Senate, 1947-91." *American Journal of Political Science* pp. 963–973.

Stuckatz, Jan. 2022*a*. "How the workplace affects employee political contributions." *American Political Science Review* 116(1):54–69.

Stuckatz, Jan. 2022*b*. "Political alignment between firms and employees in the United States: Evidence from a new dataset." *Political Science Research and Methods* 10(1):215–225.

Tsebelis, George. 2002. *Veto players: How political institutions work*. Princeton University Press.

Tullock, Gordon. 1972. "The Purchase of Politicians." *Western Economic Journal* 10:354–355.

World Bank. 2026. "Market Capitalization of Listed Domestic Companies (Current US$) — United States." https://data360.worldbank.org/en/indicator/WB_WDI_CM_MKT_LCAP_CD?country=USA&view=trend. World Development Indicators, sourced from the World Federation of Exchanges database. Accessed 2026-06-15.

5 Appendix

A.1 Firm Name Disambiguation Algorithm

In this appendix, we discuss how we link records and disambiguate firms. Our source data is filled with identifier fields, which are text-based and only semi-structured. For example, while registrants (lobbying firms) have canonical names and identifying numbers, their clients (corporate firms, individuals, NGOs, etc.) do not. This means that lobbying done on behalf of a firm needs to choose a representation of the firm name. For example, a firm working on Apple's behalf might enter "Apple", "Apple Inc.", "Apple Computer Inc", "Apple, Inc", or any number of other firm identifiers. With over 115,000 unique client-registrant pairings, there is significant variance in how names are reported. In order for researchers to know something about the clients who lobby, it is necessary to resolve these name representations to canonical identifiers.

In the present paper, the scope of this problem is that we need to connect *client names* from lobbying disclosures to *S&P Compustat* firm names and *corporate PAC* names; however, researchers interested in connecting lobbying or donation datasets in other ways must constantly reckon with this sort of semi-structured text matching.

Historically, resolving names meant specifying a set of rules for canonicalizing (disambiguating) the names, for instance, removing ", Inc.", ", Ltd.", or similar suffixes from firm names, and then ultimately using string distance measures to ascertain whether the remaining text is similar enough to be the same firm. For firms such as Apple, this may work, but suppose the entity in question is Meta, and the text representations are "Facebook", "Instagram", "Meta Inc.", and "Oculus VR". Here, any string distance methods will be insufficient to identify the underlying firm. Given the scope of our data, it is also impractical to assign identifiers using human coders.

We acknowledge alternative network-based or string distance-based approaches to grouping and disambiguating entities (Arora and Dell, 2023), but we depart from those traditions. We instead propose and use a novel unsupervised learning approach to disambiguate corporate entities using web search results. In effect, we run a web search for the entity's name and attach the first website returned by the search to the entity. When two entities share the same website, they are the same entity. The intuition is simple and reflects a basic alignment of incentives: corporations have a financial interest in ensuring their websites appear high in the search results for any variation of their name (including across name changes, mergers, and acquisitions, etc.); and search engines have a financial interest in ensuring users find what they are looking for when they submit a query, even if the query is somewhat mis-specified. In effect, even if Apple and Google are competitors in other domains, they are incentive compatible when it comes to ensuring that searches for "Apple Computer", "Apple Store", and "iTunes" send users to roughly the same place. We isolate the domain (or, in some cases, subdomain) associated with the URL of the first response for the query and store it. The domain itself is of no real interest to us, except that it allows the linking of unrelated search queries to common result sets. Our method is agnostic with respect to the underlying search engine used, and so for operational reasons, we do not specify the one we used here.

This method will yield good results under a single major condition: coherence. In effect, as long as the textual representation is coherent and readable enough that a search engine can infer the user's intent, and as long as the entity the user is searching for has a web presence, the method will work. And it is self-improving: to the extent that search engines update their data and algorithms over time (for instance, as they incorporate LLM/transformer-based neural networks into their internal query resolution architecture), results will improve over time. At the time of writing, all 500 S&P 500 companies as well as the top 1,000 non-S&P 500 publicly traded companies match correctly, allowing us to link their Compustat records to domains. We present results of a verification exercise in Appendix A.1.1 below. We show a worked example of connecting a single firm's lobbying disclosure client names to the firm in Appendix A.1.2. And in full transparency, we describe limitations of this approach (and steps we have taken to mitigate those limitations) in Appendix A.1.3.

A.1.1 Audit of Firm Disambiguation Performance

Our approach holistically performs extremely well with respect to firms that exist today. As noted above, our method requires us to disambiguate each representation of a firm name in each data set before connecting the data. Here we present results from the first stage of disambiguation: taking a canonical firm name (from Compustat) and disambiguating that firm name. In Table A.1, we report, for brevity, the top 50 firms of the S&P 500 along with the domain names our method automatically links to the firms. Through manual audits, we have verified that the entire S&P 500 and all of the top 1,000 largest non-S&P 500 firms resolve correctly. In Table A.2, we present our results for 50 randomly selected publicly traded firms and achieve 96% accuracy.

Firm Name	Ticker	gvkey	Domain	Verified
Apple Inc	AAPL	001690	apple.com	✓
Microsoft Corp	MSFT	012141	microsoft.com	✓
NVIDIA Corporation	NVDA	117768	nvidia.com	✓
Amazon.com Inc	AMZN	064768	amazon.com	✓
Meta Platforms Inc	META	170617	facebook.com	✓
Alphabet Inc	GOOGL	160329	abc.xyz	✓
Berkshire Hathaway Inc	BRK.B	002176	berkshirehathaway.com	✓
Eli Lilly and Co	LLY	006730	lilly.com	✓
Broadcom Inc	AVGO	180711	broadcom.com	✓
JPMorgan Chase & Co	JPM	002968	jpmorgan.com	✓
Tesla Inc	TSLA	184996	tesla.com	✓
UnitedHealth Group Incorporated	UNH	010903	unitedhealthgroup.com	✓
Exxon Mobil Corp	XOM	004503	exxonmobil.com	✓
Visa Inc	V	179534	visa.com	✓
Procter & Gamble Co (The)	PG	008762	pg.com	✓
Costco Wholesale Corp	COST	029028	costco.com	✓
Johnson & Johnson	JNJ	006266	jnj.com	✓
Mastercard Inc	MA	160225	mastercard.us	✓
Home Depot Inc. (The)	HD	005680	homedepot.com	✓
AbbVie Inc	ABBV	016101	abbvie.com	✓
Walmart Inc	WMT	011259	walmart.com	✓
Netflix Inc	NFLX	147579	netflix.com	✓
Merck & Co Inc	MRK	007257	merck.com	✓
Coca-Cola Co (The)	KO	003144	coca-colacompany.com	✓
Bank of America Corp	BAC	007647	bankofamerica.com	✓
SALESFORCE INC	CRM	157855	salesforce.com	✓
Advanced Micro Devices Inc	AMD	001161	amd.com	✓
Adobe Inc	ADBE	012540	adobe.com	✓
Chevron Corp	CVX	002991	chevron.com	✓
PepsiCo Inc	PEP	008479	pepsico.com	✓
Thermo Fisher Scientific Inc	TMO	010530	thermofisher.com	✓
Oracle Corp	ORCL	012142	oracle.com	✓
Linde Plc	LIN	025124	linde.com	✓
Accenture PLC	ACN	143357	accenture.com	✓
McDonald's Corp	MCD	007154	mcdonalds.com	✓
Cisco Systems Inc	CSCO	020779	cisco.com	✓
Wells Fargo & Co	WFC	008007	wellsfargo.com	✓
QUALCOMM Inc.	QCOM	024800	qualcomm.com	✓
Abbott Laboratories	ABT	001078	abbott.com	✓
Intuit Inc.	INTU	027928	intuit.com	✓
GE Aerospace	GE	005047	geaerospace.com	✓
Philip Morris International Inc	PM	179621	altria.com	✓
Texas Instruments Inc	TXN	010499	ti.com	✓
International Business Machines Corp	IBM	006066	ibm.com	✓
Danaher Corp	DHR	003735	danaher.com	✓
Amgen Inc	AMGN	001602	amgen.com	✓
ServiceNow Inc	NOW	171007	servicenow.com	✓
Verizon Communications Inc	VZ	002136	verizon.com	✓
Applied Materials Inc	AMAT	001704	appliedmaterials.com	✓
Intuitive Surgical Inc	ISRG	136725	intuitive.com	✓

Table A.1: *S&P 50 Disambiguation Results:* This table shows our successful disambiguation of the firm names of the top 50 firms of the S&P 500 by market capitalization. Success is defined here by the disambiguation domain matching the expected domain name for the firm. All domains are resolved correctly.

Firm Name	**Ticker**	**gvkey**	**Domain**	**Verified**
ABM Industries Inc	ABM	001410	abm.com	✓
Acadia Healthcare Co Inc	ACHC	192255	acadiahealthcare.com	✓
Advanced Energy Industries Inc	AEIS	061562	advancedenergy.com	✓
Apyx Medical Corp	APYX	012719	apyxmedical.com	✓
Austin Gold Corp	AUST	039937	austin.gold	✓
BARK Inc	BARK	038937	bark.co	✓
Baidu Inc	BIDU	164532	baidu.com	✓
BankFinancial Corp	BFIN	161942	bankfinancial.com	✓
Bioverativ Inc	BIVV	029819	sanofi.com	✓*
Bowman Consulting Group Ltd	BWMN	038754	bowman.com	✓
CRH Medical Corp	CRHM	154773	crhsystem.com	✓
California Resources Corp	CRC	021431	crc.com	✓
Clementia Pharmaceuticals Inc	CMTA	032130	ipsen.com	✓*
Crown Electrokinetics Corp	CRKN	036357	crownek.com	✓
Datawatch Corp	DWCH	025320	datawatch.com	✓
Digital Turbine Inc	APPS	009551	digitalturbine.com	✓
DigitalBridge Group Inc	DBRG	183603	digitalbridge.com	✓
Dynatronics Corp	DYNT	004124	dynatronics.com	✓
Dynavax Technologies Corp	DVAX	141496	dynavax.com	✓
EPR Properties	EPR	065710	eprkc.com	✓
Eaton Corporation plc	ETN	004199	eaton.com	✓
Express Scripts Holding Co	ESRX	025356	express-scripts.com	✓
Extraction Oil & Gas Inc	XOG	027677	extractionog.com	✓
First Connecticut Bancorp Inc	FBNK.1	186344	finexcu.org	✓*
Flex LNG Ltd	FLNG	288586	flexlng.com	✓
FreeCast Inc	FCST	036384	freecast.com	✓
Fusion Pharmaceuticals Inc	FUSN	036571	fusionpharma.com	✓
Gritstone Bio Inc	GRTS	034082	gritstonebio.com	✓
GrubHub Inc	GRUB.2	019906	grubhub.com	✓
Halozyme Therapeutics Inc	HALO	151630	halozyme.com	✓
Immuneering Corp	IMRX	039189	immuneering.com	✓
Insight Enterprises Inc	NSIT	031453	insight.com	✓
Kemper Corp	KMPR	021503	kemper.com	✓
Manchester United Plc	MANU	200480	manutd.com	✓
MaxLinear Inc	MXL	184551	maxlinear.com	✓
Momentus Inc	MNTS	039427	momentus.space	✓
Olympic Steel Inc	ZEUS	029870	olysteel.com	✓
Protara Therapeutics Inc	TARA	036044	protaratx.com	✓
RHC Capital Corp-old	RHC.5	020048	rhccap.com	✓
Reneo Pharmaceuticals Inc	RPHM	038453	reneopharma.com	✓
Reven Housing REIT Inc	RVEN	186078	pacificoakcapitaladvisors.com	✗
Robinhood Markets Inc	HOOD	039293	robinhood.com	✓
Sana Biotechnology Inc	SANA	037925	sana.com	✓
Spring Bank Pharmaceuticals Inc	SBPH	026790	spglobal.com	✗
Syra Health Corp	SYRA	042910	syrahealth.com	✓
TTM Technologies Inc	TTMI	139804	ttm.com	✓
Tenet Healthcare Corp	THC	007750	tenethealth.com	✓
The PNC Financial Services Group Inc	PNC	008245	pnc.com	✓
Travere Therapeutics Inc	TVTX	016848	travere.com	✓
Viking Therapeutics Inc	VKTX	021032	vikingtherapeutics.com	✓

Notes:
* These domains are correct based on our criteria for mergers / acquisitions / firm name changes

Table A.2: *Random Firm Audit Results:* This table shows our successful disambiguation of the firm names of 48 of 50 (96% accuracy) randomly selected firms among the publicly traded companies tracked by WRDS Compustat.

Year Firm Defunct	2006	2008	2010	2015
Accuracy Rate	66%	76%	78%	79%

Table A.3: *Audit of Firms by Year Ceased Trading:* The proportion of companies disambiguated correctly across four audit samples of 100 firms that were at one point publicly traded but ceased trading in the selected year. Performances worsens with older years and becomes significantly worse in 2006.

Our performance declines with respect to firms that ceased trading years ago. As a representative example of our performance for long-bankrupt firms, we search the S&P Compustat database to identify a sample of 100 firms that ceased trading in each of 2015, 2010, 2008 (the most recent major legislative change to the LDA and an inflection point at which our data becomes more comprehensive) and 2006. They may cease trading due to bankruptcy, acquisitions, being taken private, or any other means. We manually code whether we are able to disambiguate the firm name correctly from the Compustat data and report our results in Table A.3. We perform well on this hard task, but the error rate increases the further back we go, and the 2006 sample has significantly more errors.

The errors in the above audits follow a similar pattern to those described above: firms that ceased to exist a long time ago typically have very poor quality search results, including results from heretofore unknown (and thus not automatically rejected) news sites or business directories. In other cases, the liquidation or financing firms that were involved in bankruptcy proceedings (e.g., Blackstone) are incorrectly identified as the firm being liquidated.

This problem is self-correcting in two respects. As we identify errors in our disambiguation, we fix them. Thus, the above accuracy rates are synthetic for the purpose of this appendix — we have already fixed the errors identified in these audits specifically. Second, the poor quality of search results for long-gone firms only affects firms that disappeared before we began running our algorithm. If firms that are correctly disambiguated were to disappear tomorrow, our results would continue to be correct on a go-forward basis.

Using a combination of rules-based approaches and human supervision, we have additionally audited a significant portion of our data, including tens of thousands of client records, thousands of Compustat firm records, and thousands of corporate PACs.

A.1.2 Example of Client Disambiguation: RTX (Raytheon)

More difficult than resolving firm names from a firm database is resolving the other end of the equation: the client name fields in U.S. federal lobbying reports. In this section, we present results that are typical of our performance for this task. in a real-world scenario. Consider RTX Corporation (formerly Raytheon), a large defense contractor. Prior to our audit, 35 unique client names disambiguated to RTX: three of these were in error, and 32 were correct (92% accuracy). We provide a list of all of these representations in Table A.4.

There are two main errors that surfaced in this exercise. The first error, "Copeland, Lowery, Jacquez, Denton, & White", was the pre-2007 name of a lobbying firm now called "Innovative Federal Strategies". An RTX executive was a partner at this firm, and his corporate RTX biography page was, for a time, the highest ranked search engine result for the now-obsolete firm name. Second, two errors related to "Silicon Graphics". Silicon Graphics is a long-defunct business that produced high-end graphics rendering servers. A press release from RTX announcing a new headquarters was, for a time, the top search result for "Silicon Graphics Federal Inc". As noted above, these errors are fixed in our live data.

The RTX data reveals the limitations of distance-based measures: no text-based measure would be capable of connecting "United Technologies Corp." to "Raytheon Missile Systems", but our disambiguation method does so trivially.

Client Name		Client Name	
COPELAND LOWERY JACQUEZ DENTON & WHITE	✗	RAYTHEON TECHNOLOGIES CORPORATION (FKA RAYTHEON COMPANY)	
RAYTHEON		RAYTHEON TECHNOLOGIES CORPORATION AND AFFILIATES (F K A UNITED TECHNOLOGIES COR	
RAYTHEON BLACK BIRD TECHNOLOGIES		RAYTHEON TECHNOLOGIES CORPORATION AND AFFILIATES (F K A UNITED TECHNOLOGIES COR	
RAYTHEON CO		RAYTHEON TECHNOLOGIES CORPORATION AND AFFILIATES (FKA RAYTHEON COMPANY)	
RAYTHEON COMPANY		RAYTHEON TI SYSTEMS	
RAYTHEON COMPANY ON BEHALF OF RAYTHEON SYSTEMS LIMITED		RTX CORPORATION	
RAYTHEON COMPANY SPACE & AIRBORNE SYSTEMS		RTX CORPORATION (FKA RAYTHEON TECHNOLOGIES CORPORATION AND AFFILIATES)	
RAYTHEON CORP		RTX CORPORATION AND AFFILIATES	
RAYTHEON CORPORATION		RTX CORPORATION AND AFFILIATES (FKA RAYTHEON TECHNOLOGIES CORP. AND AFFILIATES)	
RAYTHEON ENGINEERS AND CONSTRUCTORS		RTX CORPORATION AND AFFILIATES (FKA RAYTHEON TECHNOLOGIES CORPORATION AND AFFIL	
RAYTHEON MISSILE SYSTEMS		RTX CORPORATION AND AFFILIATES(RAYTHEON TECHNOLOGIES CORPORATION AND AFFILIATES)	
RAYTHEON SYSTEM CORP		SILICON GRAPHICS FEDERAL INC	✗
RAYTHEON SYSTEMS CO		SILICON GRAPHICS, INC. FEDERAL	✗
RAYTHEON SYSTEMS COMPANY		UNITED TECHNOLOGIES	
RAYTHEON SYSTEMS CORP		UNITED TECHNOLOGIES CORP	
RAYTHEON TECHNOLOGIES CORP. (RTX) (F.N.A. UTC)		UNITED TECHNOLOGIES CORP PRATT & WHITNEY DIV	
RAYTHEON TECHNOLOGIES CORPORATION		UNITED TECHNOLOGIES CORP.	
RAYTHEON TECHNOLOGIES CORPORATION (FORMALLY RAYTHEON COMPANY)		UNITED TECHNOLOGIES CORPORATION	

Table A.4: *Raytheon Disambiguation Results:* All lobbying client names that disambiguated to RTX (Raytheon). Our accuracy rate in this disambiguation was 91%; three errors are noted above in red. All errors are fixed in our live data.

A.1.3 Limitations

This method is not perfect. Researchers who use traditional text-distance methods on raw data currently commit thousands of errors by failing to connect firms that are actually connected. Our method will reduce these errors (though not eliminate them) and may, in some cases, produce the opposite error: identifying two firms as connected when they are not. In this subsection, we describe several cases where our results may be suboptimal and the steps we take to mitigate these errors.

Mergers and Acquisitions: Some limitations are conceptual. Suppose a firm is publicly traded and has political activity. Later, it is acquired by another publicly traded firm and becomes a wholly owned subsidiary or is incorporated into the firm. Eventually, domain results will converge, and the two firms will appear to this method as being the same firm. Researchers might have particular theories about whether it makes sense to treat these firms as identical after the merger, across the span of the dataset, or perhaps never; this preference might differ from firm to firm. Our approach captures acquisitions when the acquiring entity assumes the internet presence of the acquired entity, but not when the acquired entity retains a separate consumer-facing brand. This can be complicated further by firms that are acquired, spun out, and acquired again: it is a matter of conceptual debate whether eBay and PayPal should be considered a single company or two companies at any given point in time.

No Consumer Presence: Other limitations are practical. Some classes of firms are hardly consumer facing at all. In particular, firms that provide upstream energy services or commercial logistics operate with little public-facing presence. Firms without a public-facing presence typically have their results dominated by news coverage, business directories, and public documents on government websites. Even excluding all of these, we might still be unable to locate a public-facing domain for the company. In the present analysis, we believe this limitation is *de minimis* because very few, if any, of these firms will be publicly traded.

Long-Bankrupt Firms: A search engine represents the state of the art as of the time of the search. But the approach is weaker for firms that no longer exist and have not existed for some time. Suppose a firm ceases trading and eventually, years later, its domain registration lapses. Over time, results related to this firm will degrade. Eventually, news sources, directories, and other third-party mentions of the firm will persist, but the firm itself will have been delisted. We mitigate this by maintaining a blocklist of domains for which results should never be correct (including news blogs, corporate directories, stock tip websites, etc.) At the time of writing, this list comprises over 2,200 domain names. We additionally augment our results; for firms that have Wikipedia pages, we attempt to identify the firm's prior domain name through those as well. This allows for simple correction of firms whose domains cannot resolve using conventional search. Finally, as noted in Appendix A.1.1, we refrain from analyzing data before 2006 in order to mitigate the worst of this problem.

Other Limitations: In the current analysis, we need only correctly disambiguate publicly-traded firms. For future analyses, we have identified five additional categories of lobbying client representations that are likely to resolve incorrectly and implemented fixes for them.[14]

Although we cannot warrant that all errors have been removed from the data, our basic claim is that they are rare and unlikely to be errors that systematically exacerbate (or minimize) our quantities of interest.

[14] Those categories are: (1) Clients that are individual persons; (2) Clients whose firm names are dictionary words; (3) Ad Hoc interest groups; (4) Lobbying with pass-through clients operating "on behalf of" an ultimate client; (5) Firms that share domain names with other firms despite not being the same firms. We are happy to provide additional detail for those interested in how we overcame these issues.

A.2 Committee Power

All committees are not equal in the U.S. Congress. Our measure is designed so that even if every committee were equally likely to be lobbied on a per-bill basis, exposure weights would still vary because committees handle significantly different numbers of bills. In this section, we give an overview of how committees differ in the business they consider and how this affects committee-industry exposure weights. We begin by showing the distribution of bills considered per committee — which we call *productivity* — and the distribution of how many bills are lobbied by committee — which we call *power*. In Figure A.1, we show the general relationship between these two attributes.

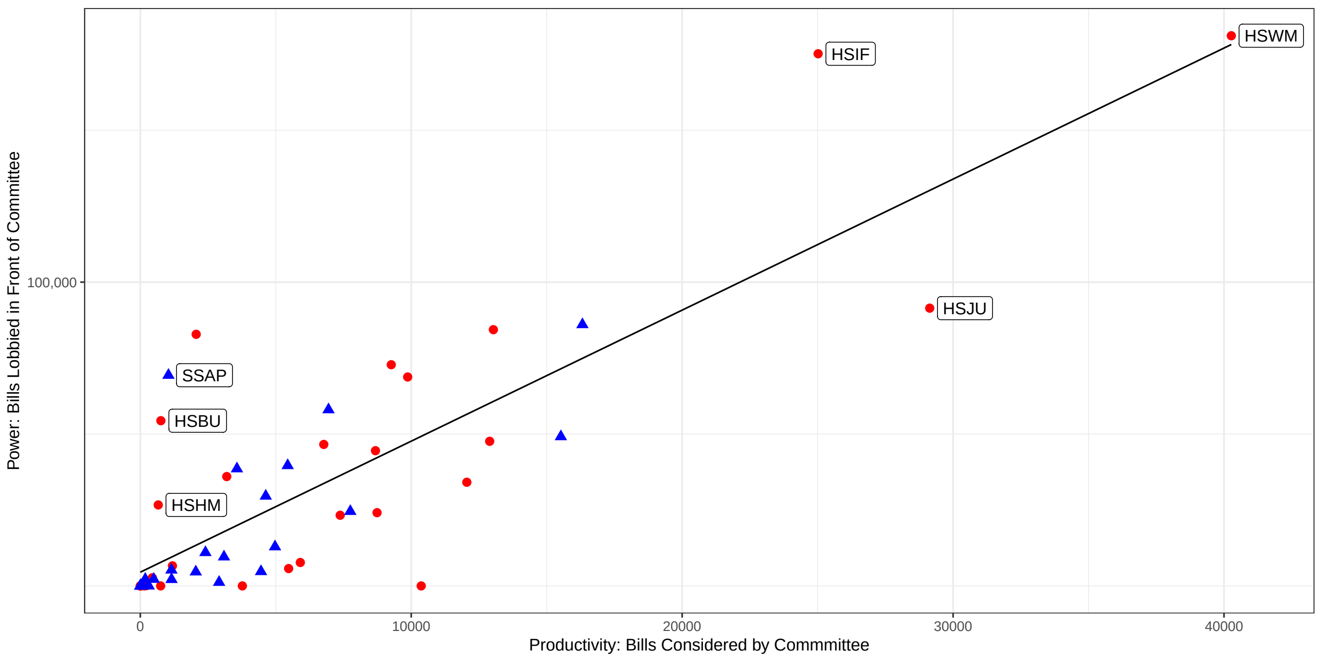

Figure A.1: *Committee Power*: This plot shows the relationship between committee productivity (how many bills are considered by a committee) and power (how much lobbying occurs on those bills). Committee chamber is depicted by color and symbol: House committees as red circles and Senate committees as blue triangles. There is significant variation in committee power and productivity. Committee with extreme values of productivity, power, or both are manually labeled and map well to traditional views of "Power Committees".

How does varying committee productivity and power translate in our ultimate weights? In Figure A.2, we show the exposure weight assigned to each committee, averaged across industries (and using the version of our weights that normalize separately by chamber).[15] These exposure weights do not track exactly with canonical notions of committee power. For example, we identify "power committees" on the basis of their Groseclose-Stewart scores, which attempt to quantify the desirability of a committee posting among members (Groseclose and Stewart, 1998; Stewart and Groseclose, 1999). Although the most exposed committees in our data are among the GS power committees, the two measures also differ in obvious ways: Senate Foreign Relations and House Rules have relatively low exposure despite being high-GS committees. By contrast, committees with the power of the purse rank highly by both measures. Regardless, as we establish in the main text, these high-GS score committees make up perhaps one third of the average firm's exposure, and so we are confident that a data-driven exposure measure best captures a firm's holistic interests.

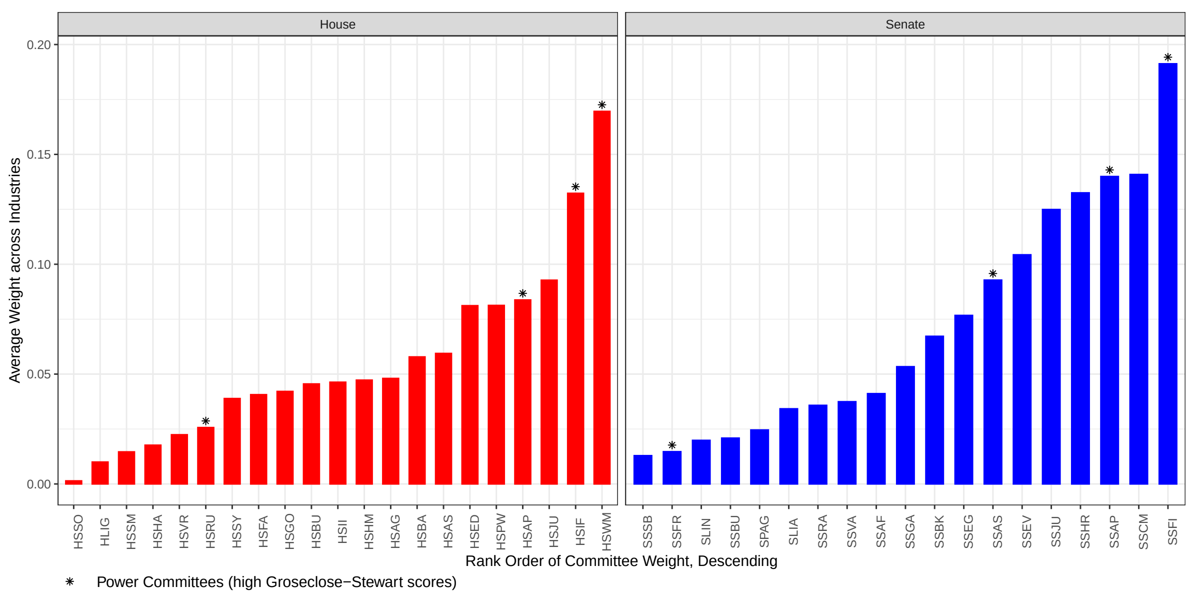

Figure A.2: *Average Exposure by Committee*: This figure shows the average exposure weight across all industries by committee and chamber. Committees with high Groseclose-Stewart scores are denoted with asterisks. Committees with significant spending powers tend to rank highly in both assessments, while Senate Foreign Relations and House Rules are high-GS score but low exposure. There is significant heterogeneity in exposure by committee.

[15]These averages are taken among industries that have some exposure to the committee, so industries with no exposure at all do not attenuate the average weight for a committee.

A.3 Face Validity of Industry-Committee Mappings

We claim that our method produces a facially valid mapping of industries to committee exposure; that is to say, we correctly measure industry lobbying of bills in front of committees, and that industries lobby bills in front of the committees that we expect to regulate the industry in question. Because no ground truth exists for this condition, we apply a simple test of facial validity: armed only with substantive knowledge of an industry's activity and the universe of congressional committees, which committees would we expect to be the top committee for a given industry? If the committees with the highest weights seem to reflect this substantive knowledge, then our measure has facial validity. Table A.5 presents the results of this exercise, which broadly conforms with our facial expectations.

Industry	# Firms	Top House Committee	Weight	Top Senate Committee	Weight
Chemical Mfg. (Pharma)	131	Energy *	0.17	Health	0.18
Computer/Electronic Mfg.	102	Appropriations	0.13	Appropriations	0.19
Credit Intermediation	64	Finance	0.18	Banking	0.21
Insurance	59	Ways and Means	0.18	Banking	0.27
Utilities	56	Ways and Means	0.15	Finance	0.22
Professional Services (Law)	43	Appropriations	0.14	Agriculture	0.20
Machinery	41	Ways and Means	0.13	Finance ‡	0.18
Transportation Manufacturing	39	Appropriations	0.14	Appropriations	0.20
Securities, Investments	39	Ways and Means	0.24	Finance	0.27
Food Manufacturing	24	Energy	0.14	Health	0.23
Electrical Manufacturing	24	Ways and Means	0.19	Appropriations	0.20
Web Search and Archives	24	Appropriations	0.11	Health	0.21
Mining (non-Oil/Gas)	23	Natural Resources	0.14	Energy	0.28
Telecommunications	23	Energy †	0.20	Commerce + Science §	0.42

Notes:
* House Energy includes Health Subcommittee
† House Energy includes Telecommunications Subcommittee
‡ Senate Finance includes Natural Resources Subcommittee
§ Senate Commerce includes Internet Subcommittee

Table A.5: *Face Validity Exercise:* This table lists the top House and Senate committee by exposure weights for each of the Top 14 Industries, sorted by number of politically active firms (firms that donate) in our database descending. The chosen committees accord with the committees we expect based on a qualitative understanding of these industries composition and interests. The notes section depicts non-obvious cases of clearly relevant subcommittees nested inside committees whose names are ambiguous.

A.4 NAICS and Industry Aggregation

Our primary specification aggregates industries to the 3-digit NAICS code (sub-sector) level, but our method, in principle, works with any possible industry aggregation. In this section, we illustrate how NAICS codes are constructed in a hierarchically nested manner and continue by showing how higher or lower aggregations would impact the industry-committee linkages in a worked example.

6-digit NAICS codes are assigned to firms on the basis of the business in which they are engaged. Firms can have multiple NAICS codes for different areas of business activity. We measure a single NAICS code per firm according to their primary business. Consider Apple Inc., which is represented in our data set by the 6-digit NAICS code 334221 (Wireless Communications Equipment) by S&P Compustat, based on its iPhone business. It could also credibly be assigned NAICS code 334111 (Computer Equipment). In Figure A.3 we illustrate how this six-digit NAICS code encodes a hierarchy of industries.

Figure A.3: *Deciphering NAICS Code Hierarchies:* This figure uses Apple Inc. to illustrate how a NAICS Code is constructed. Each successive digit encodes an increasing level of precision, hierarchically nested under higher levels of aggregation. In our primary model, we aggregate to 3-digit sub-sector groupings, and thus Apple Inc. is classified as "334".

As noted, in our primary results, firms are aggregated to the sub-sector (3-digit) level. In Apple's case, NAICS sub-sector 334 (Computer and Electronics Product Manufacturing). This sub-sector is nested under sector 33 (Manufacturing), which also includes sub-sectors 332 (Fabricated Metal e.g., Stanley Black & Decker), 333 (Machinery Manufacturing e.g., Deere & Company, Caterpillar), 335 (Electrical Appliances e.g., Whirlpool, Maytag), and 336 (Transportation Equipment e.g., Ford, Tesla, Boeing). 334 includes multiple industry group (4 digit) components, including 3341 (Computer and Peripheral Electronic Manufacturing e.g., Dell, HP), 3342 (Communications Equipment), 3344 (Semiconductors e.g., Intel, NVIDIA), and 3345 (Sensors and Instruments e.g., Thermo Fisher). And indeed, further down, industry group 3342 includes industries (5-digit) levels such as 33421 (Telephones e.g., Netgear), and 33422 (Wireless Communications).

A higher aggregation makes industry interests more diffuse. It is clear that Apple and Boeing, although both manufacturing products, are exposed to very different legislative committees. This must be balanced against data support. As we discuss in the main paper, it would be possible to calculate firm exposure weights at the firm level without any aggregation. But many firms, including firms that are politically active via donations, do not lobby. Even at lower NAICS aggregations, sparsity of data becomes a problem. Some industry groups have few publicly traded firms or limited political activity, others might be dominated by one or a smaller number of firms.[16] We illustrate this tradeoff in Table A.6.

We choose 3-digit NAICS codes as our primary aggregation because doing so provides many clusters with strong data support while maintaining facially coherent shared interests. Regardless, our results are robust to higher (2-digit NAICS sector codes) and lower (4-digit NAICS industry group codes) aggregations of industries, and we reproduce those results in Appendix A.9.2. The choice of appropriate level of aggregation for more targeted substantive research depends on the precise question being asked. Finally, nothing about our method is intrinsic to NAICS codes; any alternative clustering of firms into industries could be used.

[16] We guard against this kind of leverage in the smallest industries among our chosen aggregation in Appendix A.11.

NAICS	Industry Name	# Firms	Top Firm	Top Commmittee	Weight	# Committees	Eff. # Committees
21	Mining	400	Rio Tinto	Energy	0.16	17	8.8
32	Materials Mfg.	1,054	Shell	Health *	0.16	19	9.9
33	Product Mfg.	1090	Apple	Appropriations	0.16	18	10.6
├331	Metal Mfg.	48	Nucor	Environment	0.24	13	6.3
├332	Fabricated Metal Mfg.	40	Parker-Hannifin	Appropriations	0.27	13	7.2
├333	Machinery	127	Caterpillar	Finance	0.18	16	9.1
├334	Computer / Electronic Mfg.	530	Apple	Appropriations	0.19	18	9.3
│├3341	Computer / Peripheral Mfg.	47	HP	Appropriations	0.28	15	6.4
│├3342	Communications Mfg.	82	Apple	Armed Services **	0.27	14	6.6
│├33421	Telephone Mfg.	27	Alcatel-Lucent	Commerce + Science	1.00	1	1.0
│├33422	Wireless Mfg.	44	Apple	Armed Services **	0.31	13	6.0
│└33429	Other Comm. Mfg.	10	Tyco	Appropriations	0.49	9	3.1
│–3344	Semiconductor Mfg.	183	Taiwan Semiconductor	Appropriations	0.18	17	9.2
│–3345	Sensor / Control Mfg.	206	Thermo Fisher Scientific	Appropriations	0.21	17	7.8
├335	Electrical Mfg.	81	Sony Group	Appropriations	0.20	17	8.4
├336	Transportartion Mfg.	99	Toyota Motor	Appropriations	0.20	18	8.7
├339	Misc. Mfg.	145	Stryker	Judiciary	0.21	11	7.9
51	Information Services	768	Alphabet	Commerce + Science	0.22	18	9.4
52	Finance / Insurance	1108	UnitedHealth	Banking †	0.20	19	9.2

Notes:

* Senate Health includes Employment and Workplace Safety Subcommittee

** Senate Armed Services includes Cybersecurity and Defense Procurement ("Airland") Subcommittees

† Senate Banking includes Securities and Investment Subcommittees

Table A.6: *Aggregation and Committee Maps:* This table shows how the choice of aggregation impacts the committee mappings. At each NAICS level, we show a representative top firm, the top Senate committee this level of aggregation is exposed to, the weight of that top committee (normalized to Senate-chamber weights, excluding House committees), how many total committees that industry is exposed to, and the effective number of committees. The effective number of committees is the inverse Herfindahl-Hirschman Index. Higher levels of aggregation have more firms; groupings with more firms tend to have more diverse interests. Our main model chooses the 3-digit (sub-sector) level of aggregation.

A.5 Weight Distribution and Feasible Effect Sizes

Each of our result coefficients describes the effect of a one unit increase in the treatment variable, committee weight, but it is virtually impossible to achieve such an extreme weight. A weight of 1 is nearly impossible for many reasons: first, our primary weights are calculated across the Senate and House.[17] This means that a legislator can only obtain the maximum weight if the industry lobbies in only a single chamber, or if the legislator serves in both chambers in the same congress. Second, even within a chamber of Congress, several congressional committees (so-called "exclusive committees") explicitly preclude members from serving on other committees. Third, members serve on a very limited number of committees per term (mean 3.77, std. dev. 3.35, of 58 committees in our database); meanwhile, industries with sufficient numbers of firms have fairly diverse committee exposure. Thus, the most principled way to interpret our model results is to scale the effect coefficient according to the feasible scale of the treatment variable.

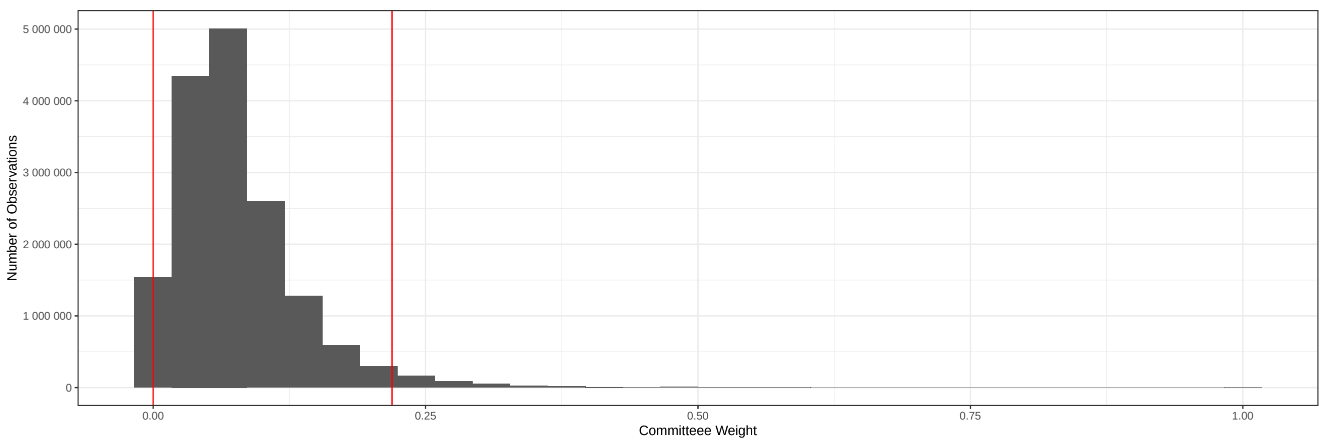

Figure A.4: *Histogram of Exposure Weights*: This histogram depicts the distribution of exposure weights across our entire panel (all legislator-firm-congress units). Red vertical lines reflect the 2.5% and 97.5% marks of the distribution, 0.00 and 0.219 respectively. The "elbow" occurs at approximately the 97.5% mark, which we suggest as a good benchmark for a large but feasible treatment value.

[17] In Appendix A.5.1, we instead normalize weights per chamber; this has the effect of approximately doubling firm-legislator weights and approximately halving the associated effect size coefficients [which mechanically scale to the unit of treatment] and does not disrupt any of the observed results.

In our main text we present the marginal effect of moving across the distribution of exposure weights, from 0th to 33rd to 66th to 99th percentile, holding all control variables at their mean (or mode). Here, we assess the feasible range of exposure weights more deeply. In Figure A.4 we present a distribution of committee weights across all observations to assess this strategy. The distribution shows a sharp elbow and a long tail. We place lines indicating the 2.5% quantile (0) and 97.5% quantile (0.219).

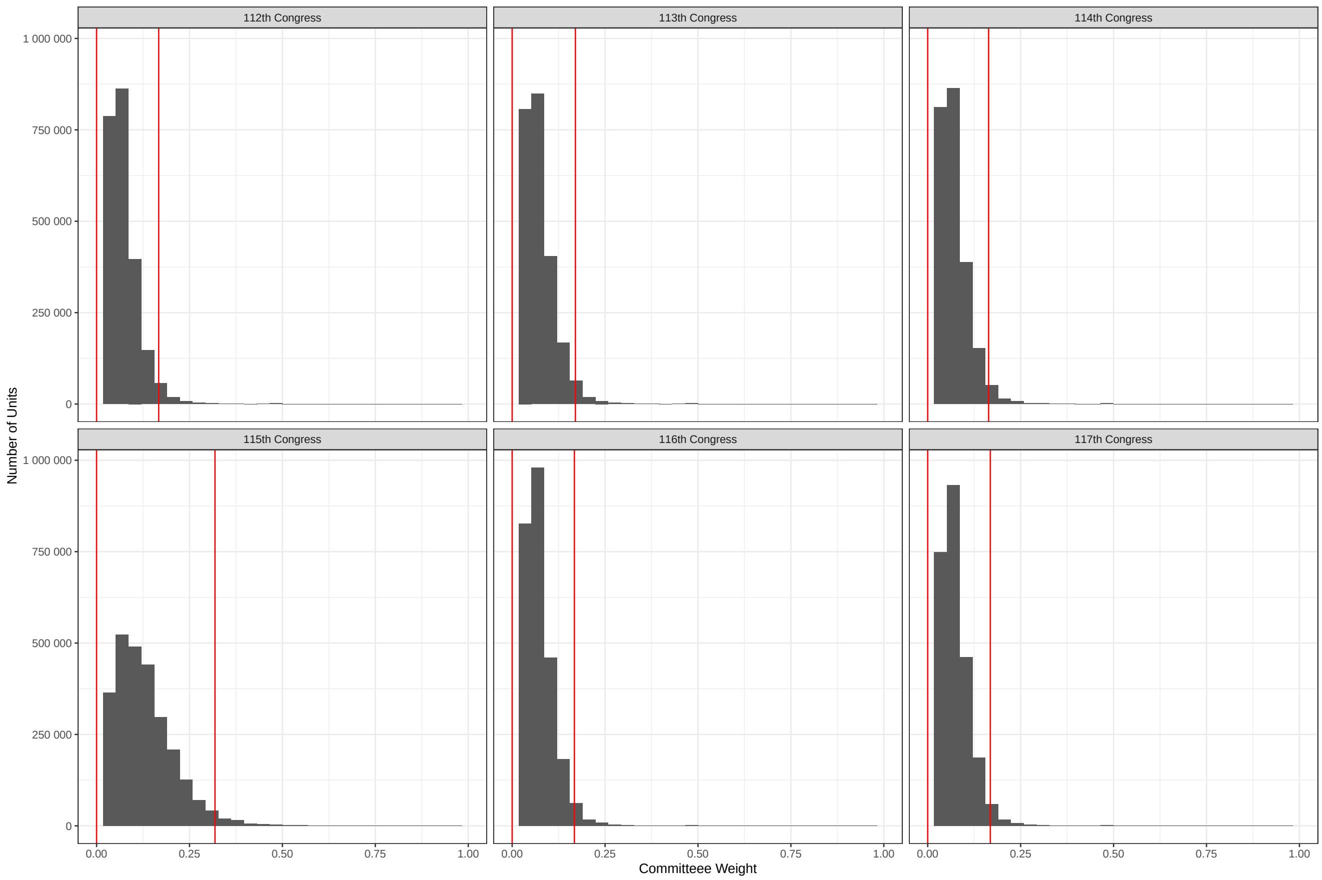

Figure A.5: *Histogram of Exposure Weights, By Congress*: This histogram depicts the distribution of exposure weights across our entire panel, disaggregated by session of Congress. Red vertical lines reflect the 2.5% and 97.5% marks of the distribution for each session of Congress, which are consistent with the aggregated distribution, except for the 115th Congress.

The purpose of this exercise is simply a graphical diagnostic of the reasonable range of values. There is no "correct" value to take, and the choice to take the 97.5% quantile is arbitrary. Just the same, it has a feature that is useful: it represents the end of the "elbow" in density in the distribution. Additionally, the value of 0.219 is relatively consistent across various specifications: in Figure A.5 we present the distribution by Congress, showing that the 97.5% quantile is very nearly identical in all congresses except the 115$^{\text{th}}$ Congress. The IQR is spans 0.045 to 0.1027 and the standard deviation is 0.057. Although we omit further figures for brevity, we also explore the weight distribution among only active firms (97.5% at 0.218) and across industries (mean industry-level 97.5% quantile = 0.283).

In our main results, which employ linear models, adjusted effect sizes can be obtained by multiplying the effect size in the table by the feasible weight (in effect, reducing effects between 4-fold and 5-fold).

A.5.1 Per-Chamber Weights and Feasible Effect Sizes

Here we depict the distribution of Exposure Weights, analogous to those presented in Figure A.4, except normalized per chamber. This has the effect of approximately doubling weights, and the largest feasible treatment weight is now 0.468 rather than 0.219. In Appendix A.9.4, we show our main results using these weights in lieu of our primary weights.

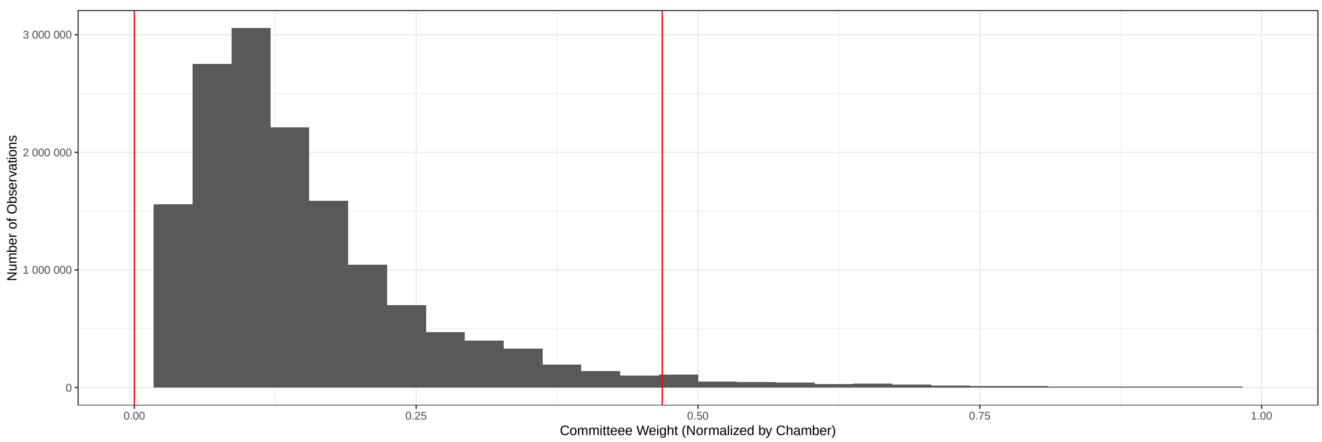

Figure A.6: *Histogram of Exposure Weights, Normalized by Chamber*: This histogram depicts the distribution of exposure weights across our entire panel (all legislator-firm-congress units). Red vertical lines reflect the 2.5% and 97.5% marks of the distribution, 0.00 and 0.468 respectively.

A.6 Confounding Variables

In this section we describe the source, operationalization, and details of the confounding ("control") variables included in our models in Table A.7. Our identification relies on the idea that treatment assignment is plausibly exogenous, conditional on the inclusion of these variables. In other words, we model the factors that predict both firm-legislator donation behavior and legislator committee assignment. Further discussion of the `legis_pastexp` (past elected experience) variable coding is found in Appendix A.6.1 below.

Variable	Source	Details
Legislator Variables:		
legis_gender	Congress	Dichotomous categorical variable for legislator gender
legis_party	Congress	Categorical variable for legislator party (including "Independent")
legis_term	Congress	Integer variable for ordinal number of congresses legislator has served in. Legislators who have non-contiguous service will have their number continue to increment.
legis_chamber	Congress	Dichotomous categorical variable for which chamber the legislator is serving in
legis_majority	Congress	Is the legislator a member of their chamber's majority? For ambiguous cases, such as independents, this variable defaults to false.
legis_chair	Congress	Is the legislator a committee chair?
legis_leader	Congress	Is the legislator otherwise a member of congressional leadership? Leadership is defined to include Speaker of the House, House Majority and Minority Whip, House Majority and Minority Leader, Senate Majority and Minority Leader, and Senate Majority and Minority Whip.
legis_pastexp	Congress	Did the legislator serve in lower (or higher) elected office prior to their current position?
legis_state	Congress	Categorical variable encoding the state the legislator serves. Territorial delegates and the D.C. Representative are excluded from our dataset.
nominate_dim1	Voteview	First-dimensional NOMINATE score (range -1 to 1) for the legislator in this congress. Legislators who did not vote a sufficient number of times to be assigned a score are missing a value for this variable.
nom2	(Derived)	Squared value of NOMINATE score, included to capture ideological extremity
legis_margin	FEC	Legislator's most recent electoral margin of victory over the second place contestant. This variable is capped at 25% to avoid issues with uncontested or substantially uncontested races. Legislators for whom this value is missing are imputed to their previous margin of victory or when this is missing, their party's average margin of victory in the election.
Firm Variables:		
revt	S&P Compustat	Firm's annual revenue (averaged across the two years of the associated congress), log-transformed at the model stage. Firms with negative revenue have their revenue truncated to 0 and the variable neg_revt set to 1.
capx	S&P Compustat	Firm's annual capital expenditure (averaged across the two years of the associated congress), log-transformed at the model stage. Firms with negative capital expenditure have their capital expenditure truncated to 0 and the variable neg_capx set to 1.
emp	S&P Compustat	Firm's employee headcount (thousands), log-transformed at the model stage.
neg_revt	(Derived)	Binary indicator which takes the value 1 if the firm's revenue was negative in this row before being changed to 0.
neg_capx	(Derived)	Binary indicator which takes the value 1 if the firm's capital expenditure was negative in this row before being changed to 0.
naics	S&P Compustat	3-digit version of firm's NAICS code, which describes its industrial classification.
vapl	(Derived)	Value Added Per Worker, calculated by revenue less capital expenditure per employee at the firm. Value is missing for firms missing any one of the three constituent variables.
Firm-Legislator Variables:		
same_state	(Derived)	Does the legislator's state and the firm's home state match? Firms that are headquartered outside the U.S. but included due to trading on a U.S. exchange have the value FALSE for this variable.

Table A.7: *Codebook of Confounding ("Control") Variables:* The above table lists each of the control variables in our models with information about sourcing and operationalization.

A.6.1 Coding Past Legislative Experience

Legislators have a diverse range of backgrounds. Previous work (Jacobson and Kernell, 1983; Dal Bó and Finan, 2018) has identified that legislators with prior elected experience are both viewed as "quality candidates" (predicting donations) and as effective legislators (predicting committee assignments). To assess whether a legislator has had prior electoral experience, we extract their complete biography from the U. S. Congress Bioguide. We identify relevant past experience and code any legislator with past experience as 1, and all other legislators as 0. In the quoted passage below, we show an example biography, highlighting text that was used to determine past experience.

> ROMNEY, Willard Mitt (Mitt), a Senator from Utah; born in Detroit, Mich., March 12, 1947; B.A., English, Brigham Young University, 1971; J.D., Harvard Law School, 1975; M.B.A., Harvard Business School, 1975; investment company founder and chief executive officer; consulting firm chief executive officer; president and chief executive officer of the Salt Lake Organizing Committee for the Olympic Winter Games of 2002; **governor of Massachusetts 2003-2007**; was an unsuccessful candidate for the Republican presidential nomination in 2008; was an unsuccessful Republican nominee for President of the United States in 2012; elected as a Republican to the United States Senate in 2018 for the term ending January 3, 2025.

Table A.8: *Congressional Bioguide Biography:* This excerpt depicts the Congressional Bioguide biography of Mitt Romney (R-UT). The highlighted text indicates that, on account of his past service as Governor of Massachusetts, we identify Mitt Romney as having held past elected office.

The primary roles that our coding identifies as having past experience are: state House and Senate; state Governor, Lt. Gov., and Attorney General; other state elected positions (Comptroller, Treasurer); elected Judgeships; Mayor and city council positions; U.S. House and Senate. A variety of edge cases exist: for example, when a position is elected in some states and not others (e.g., State Treasurer), we view it as prior experience for comparability across states. We exclude positions in quasi-elected intergovernmental bodies (for example, Sen. John Thune serving as the Executive Director of the South Dakota Municipal League is not counted as prior elected experience). Our results are not sensitive to the inclusion or exclusion of this variable.

A.7 Placebo Test Results

To guard against a situation where certain committees are consistently high-weight and also attract donations for reasons *other* than the mechanism of actual strategic importance to industries, we run a placebo test ($n = 1000$ iterations) where industry weights are randomly assigned to another industry. The resulting distribution is not centered at zero (such that there is some dependence) but it is greatly attenuated compared to the appear result, which is the highest magnitude observation in the entire placebo distribution ($p < 0.001$).

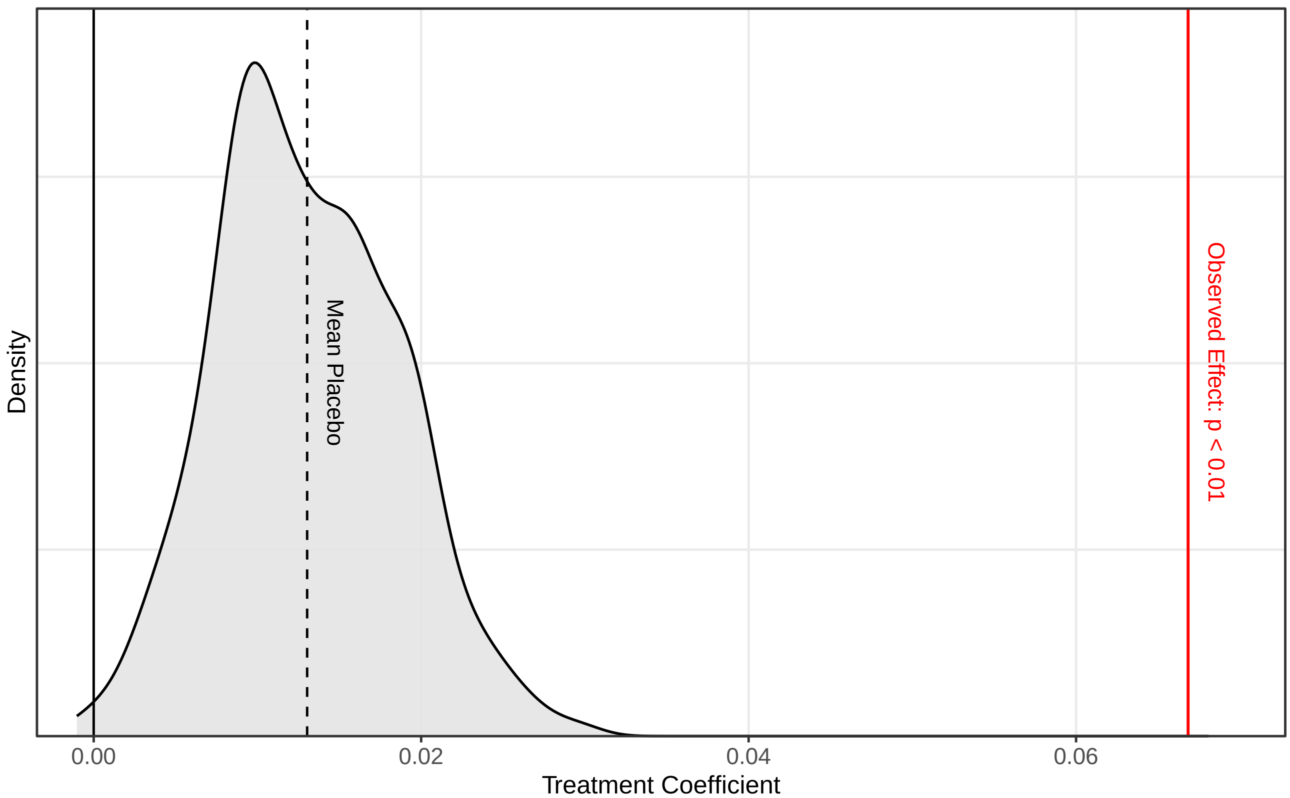

Figure A.7: *Placebo Distribution Plot*: This plot depicts the results of a placebo test where industry-committee exposures were randomly permuted to ensure the effects did not derive from persistently high-weight committees. Although the placebo is not centered at exactly 0, its mean is substantially below the observed effect size, and the observed effect size is large enough to reject the null hypothesis that the effect would emerge by chance.

A.8 Alternate Specifications, Firm-Legislator Estimates

A.8.1 Bivariate Specification

Table A.9 runs each of our samples as bivariate regressions, suppressing controls and fixed effects. The results are substantively unchanged. Our results are not at all sensitive to the inclusion of confounding variables or the fixed effect structure. Although we suppress additional models for brevity, we can confirm that our results are substantively unchanged running models with and without each set of confounders (no confounders, legislator characteristics, firm characteristics) and with and without fixed effects and include the code to do so in our replication archive.

Dependent Variables:	Donated?			Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.059***	0.069***	0.361***	0.468***
	(0.010)	(0.014)	(0.044)	(0.084)
Fit statistics				
Observations	16,040,562	2,910,562	2,689,443	16,039,225

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.9: *Effect of Committee Exposure on Donation Amount (Mininal Specification):* These models re-estimate our base specification on a bivariate basis, with no controls or fixed effects Results are substantively comparable to our primary models.

A.8.2 Donation Amount D.V. (Poisson)

In Model (4) in our primary specifications, we employ the donation amount as the dependent variable. We subset to only observations that contain donations ($Amount > 0$) to calculate the intensive margin of the treatment effect. Here, we run the model on the entire panel, capturing both intensive and extensive margins. We run the model as a Poisson regression because running a linear regression with a log-transformed dependent variable that contains many 0s using the familiar $log(X + 1)$ or inverse hyperbolic sine transformations produces distorted results (Chen and Roth, 2024; Mullahy and Norton, 2024) (though, running the model would produce positive and significant results, regardless). The resulting coefficients here are multiplicative increases, not scaled to the feasible effect size, and combine intensive and extensive margins, so they are not easily interpreted. Taking the average partial derivative (i.e., holding all observations and all covariates fixed and adjusting treatment value from 0 to a high feasible treatment value, see Appendix A.5) implies $690 million in additional donations combining the two effects.

Dependent Variable:	amount
Model:	(1)
Variables	
Exposure Weight	4.69***
	(0.503)
Same State?	1.62***
	(0.154)
Legislator Characteristics:	Yes
Firm Characteristics:	Yes
Fixed-effects	
Congress	Yes
Industry	Yes
Fit statistics	
Observations	9,353,694

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.10: *Effect of Committee Exposure on Donation Amount (Poisson):* This model estimates the effect of committee exposure on donation amount at the firm-legislator-congreess level, estimated on the full sample. The result, thus, combines the extensive margin of moving from no donation to some donation, and the intensive margin of the value of that donation.The result is positive, significant, and large.

A.8.3 Only Among Legislators on Firms Top Committees

Here, we subset our data to only those firm-legislator dyads where the legislator serves on the firm's top committee. In other words, all of the variation in the treatment in these models emerges from whether or not the legislators are present on additional committees of interest to the firm. The result loses significance for first-term congresspersons (though we note that the substantive magnitude is unchanged, and we attribute the loss of significance to a reduced subsample size). Our results show that methods that assign a single committee to a firm underestimate the effect size by ignoring the marginal contributions of other committees.

Dependent Variables:	Donated?			Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.024**	0.023	0.113*	0.187**
	(0.008)	(0.014)	(0.043)	(0.066)
Same State?	0.035***	0.049***	0.162***	0.286***
	(0.007)	(0.008)	(0.018)	(0.059)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	1,357,170	129,198	244,192	1,357,170

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table A.11: *Effect of Committee Exposure on Donations (Additive Variation Within Top Committee Members):* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level, restricted to legislators serving on their industry's highest-exposure committee (by chamber). Standard industry exposure weights and the main covariate set are used. Identification comes from variation in total exposure among legislators who share the industry's top committee assignment. Model 2 repeats the estimation for first-term congresspersons. Model 3 repeats the estimation for firms that have at least one donation. Model 4 uses a logged dollar amount dependent variable.

A.8.4 Alternative Fixed Effects

Elsewhere, our fixed effect structure includes industry and congress fixed effects, so models capture variation within-industry and within-congress. In Table A.12, we show our main specification with firm fixed effects rather than industry fixed effects; with industry and legislator fixed effects; and with firm and legislator fixed effects. Finally, we show interacted fixed effects (Industry × Congress rather than Industry + Congress). Results are comparable.

Dependent Variable:	Donated?			
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.079***	0.078***	0.078***	0.080***
	(0.012)	(0.013)	(0.013)	(0.012)
Same State?	0.030***	0.031***	0.030***	0.031***
	(0.006)	(0.006)	(0.006)	(0.006)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	
Firm	Yes		Yes	
Industry		Yes		
Legislator		Yes	Yes	
Industry-Congress				Yes
Fit statistics				
Observations	9,502,233	9,502,750	9,502,233	9,502,750

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.12: *Effect of Committee Exposure on Donations (Alternate Fixed Effects):* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level. Model 1 substitutes industry for firm fixed effects. Model 2 includes legislator and industry fixed effects. Model 3 includes legislator and firm fixed effects. Model 4 includes Industry × Congress Fixed Effects. Results are comparable to our main estimates in all cases.

A.8.5 Using `did_multiplegt_dyn` Estimator

TWFE estimators have been subject to well-documented criticisms regarding the effects they actually identify, which we discuss above in the main text. Here, we run a modern heterogeneous treatment effect estimator version of our main results, derived in De Chaisemartin and d'Haultfoeuille (2026). We selected this estimator because it is one of the few that allows for panels where all units are continuously treated with a variable treatment dose (our exposure weights). The results presented are a version where the effect size is scaled to be normalized to unit-dose (i.e., as comparable as possible to our primary TWFE LPM), though the differences in estimand mean that we caution against directly comparing magnitudes. n differs because identification relies only on groups with two or more observations. The results are unchanged.

Dependent Variables:	Donated?		Dollars (log)
Model:	(1)	(2)	(3)
Variables			
Exposure Weight	0.085***	0.487***	0.681***
	(0.020)	(0.114)	(0.163)
Legislator Characteristics:	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes
Fixed-effects			
Congress	Yes	Yes	Yes
Industry	Yes	Yes	Yes
Fit statistics			
Observations	3,156,341	585,653	3,156,341

*Signif. Codes: ***: 0.001*

Table A.13: Effect of Committee Exposure on Donations (`did_multiplegt_dyn` Estimator): We run our main model using the estimator described in (De Chaisemartin and d'Haultfoeuille, 2026), which is a modern Heterogeneous Treatment Effect estimator which identifies based on the movement of "switcher" observations to mitigate the inferential risks associated with our stock TWFE model. Model 1 uses our primary specification; Model 2 subsets to active firms; and Model 3 substitutes logged dollar amount of contributions as the dependent variable. We do not run the first-term subset because this estimator relies on within-unit variation. Results are positive, significant, and large (though differences in the estimand make direct comparison difficult), and consistent with our primary results.

A.9 Alternate Treatments, Firm-Legislator Estimates

A.9.1 Restricting Committee Counts

In the main text, we contrast our holistic approach to committee exposure with existing research, which typically assigns firms to a single committee of interest based on industry. In Table A.14, we present results that capture varying degrees of compromise between those two extremes: restricting weight calculation to the top 10, top 5, and top 1 committee by industry. These specifications also get closer to core industry-level activity, but at the risk of smoothing out heterogeneity within a given industry. Selecting the top 1 committee reflects a data-driven version of prior practice. Here we simultaneously show that our result is not directionally sensitive to the precise construction of the exposure weight measure and that prior work seriously understates the effect magnitude when it assigns a single committee.

Dependent Variable:	Donated?			
	All Committees	Top 10	Top 5	Top 1
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.079***	0.063***	0.044***	0.005
	(0.012)	(0.011)	(0.009)	(0.002)
Same State?	0.031***	0.033***	0.036***	0.037***
	(0.006)	(0.007)	(0.007)	(0.009)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	9,502,750	6,068,516	3,454,643	764,831

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.14: *Effect of Committee Exposure on Donations, Robustness to Alternate Exposure Measures:* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level. Models differ in how weights are constructed: using all committees lobbied by an industry, just the top 10 committees, just the top 5, and just the top committee. Effects attenuate but remain large and significant across specifications.

A.9.2 Alternate Industry Aggregations

Our method does not require a specific grouping of firms within industries. Our default model uses NAICS 3-digit sub-sector codes to isolate industries, but we invite research into other methods of grouping firms as necessary. In Table A.15 we show that alternate industry aggregations using 2-digit NAICS sector codes or 4-digit NAICS industry group codes produce similarly large and significant results; we remain agnostic and do not offer any further commentary on the optimal conceptual trade-off between the precision offered by lower levels of aggregation and the smoothing provided by higher levels of aggregation.

Dependent Variable:	Donated?	
	NAICS 2-digit	NAICS 4-digit
Model:	(1)	(2)
Variables		
Exposure Weight	0.096***	0.073***
	(0.022)	(0.009)
Same State?	0.030***	0.032***
	(0.006)	(0.007)
Legislator Characteristics:	Yes	Yes
Firm Characteristics:	Yes	Yes
Fixed-effects		
Congress	Yes	Yes
Industry (NAICS2)	Yes	
Industry (NAICS4)		Yes
Fit statistics		
Observations	9,696,281	8,640,326

*Signif. Codes: ***: 0.001*

Table A.15: *Effect of Committee Exposure on Donations, Robustness to Alternate Industry Aggregations:* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level. These models use 2-digit NAICS sector codes and 4-digit NAICS industry group codes rather than the 3-digit NAICS sub-sector codes used in our main model to construct industry aggregations. Effects remain large and significant.

A.9.3 Weights Assigned to Subcommittees

In our main paper, we assign exposure weights based on committee membership and exclude subcommittees from consideration. In recent decades, Congress has increasingly outsourced legislative work from committees to subcommittees. Many of the highest exposure committees are those with the power of the purse, which in turn contain subcommittees that more directly match the interests of individual industries. Consider the U.S. Senate Finance Committee, which presently has six subcommittees: (1) Energy, Natural Resources, and Infrastructure; (2) International Trade, Customs, and Global Competitiveness; (3) Health Care; (4) Social Security, Pensions, and Family Policy; (5) Taxation and IRS Oversight; and (6) Fiscal Responsibility and Economic Growth. Given this hierarchy, it may be more appropriate to measure industry interests in a more tailored fashion at the subcommittee level, though doing so presents some measurement challenges (significantly fewer legislators are members of any given subcommittee, lowering average exposure weights; moreover, we only have subcommittee membership data for Congresses beginning with the 115$^{\text{th}}$). Finally, even given the presence of full-panel subcommittee data, the purview, name, and in some cases existence of subcommittees differ across time, making it harder to exploit the pre-treatment lobbying activity while measuring legislator weights in later congresses. Regardless of these disclaimers, in Table A.16 we show that alternate exposure weights calculated at the subcommittee level (excluding committees) show positive and significant results, though somewhat attenuated compared to the committee-level weights.

Dependent Variables:		Donated?		Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.041***	0.029	0.217***	0.321***
	(0.009)	(0.010)	(0.046)	(0.074)
Same State?	0.026***	0.028***	0.144***	0.216***
	(0.007)	(0.006)	(0.022)	(0.055)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	3,451,937	622,885	576,920	3,451,937

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.16: *Effect of Committee Exposure on Donations, Subcommittee Exposure Weights:* Linear probability models estimating the effect of subcommittee exposure on probability of donation at the firm-legislator-congress level including legislator and firm controls and two-way industry and congress fixed effects. These models use weights assigned to subcommittee exposure rather than committee exposure. Model 2 repeats the estimation for first-term congresspersons. Model 3 repeats the estimation for firms that have at least one donation. Model 4 uses a logged dollar amount dependent variable. Effects remain large and significant, although somewhat attenuated compared to our primary models. Data is limited to the 115th and subsequent Congresses.

A.9.4 Weights Normalized Per-Chamber

In our primary specification of industry-committee weights (and thus firm-legislator weights), weights are calculated across all committees in both the House and Senate. As described in Appendix A.5, this typically creates a mechanical restriction on weights, since industries lobby both House and Senate bills, and legislators serve in only one chamber at a time (excepting vanishingly few bridging observations, such as Representatives who resign their seat mid-session after appointment or special election to the Senate). In this subsection, we instead normalize weights per chamber, such that $w^h_{ic} = \frac{w_{ic}}{\sum_{m \in h} w_{im}}$ where h refers to a chamber, m to a committee that is part of the chamber, i to an industry, and c to the committee at hand. Appendix A.5.1 shows how these weights propagate to the firm-legislator level. Here, we present our primary model estimates using these weights. Weights are approximately doubled across the board, while the resulting effect coefficients are approximately halved, leading to no change in interpretation.

Dependent Variables:		Donated?		Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.039***	0.032***	0.204***	0.312***
	(0.006)	(0.006)	(0.024)	(0.047)
Same State?	0.031***	0.035***	0.150***	0.251***
	(0.006)	(0.007)	(0.019)	(0.052)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	9,502,750	1,690,717	1,751,259	9,502,750

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.17: *Effect of Committee Exposure on Donations (Chamber Weights):* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level including legislator and firm controls and two-way industry and congress fixed effects. Model 2 repeats the estimation for first-term congresspersons. Model 3 repeats the estimation for firms that have at least one donation. Model 4 uses a logged dollar amount dependent variable. Relative to the main results, these models use weights normalized by chamber, which approximately doubles weights and approximately halves coefficient sizes. All results remain large and statistically and substantively significant.

A.9.5 Rolling Pre-Periods For Weight Calculation

Our primary specifications rely on a clearly defined "pre-period", consisting of the 105$^{\text{th}}$-111$^{\text{th}}$ Congresses. One possible alternative formulation would be to use a rolling pre-period. In this sub-section we present weights calculated via three different rolling strategies: in the first strategy, weights use all lobbying before the reference Congress (i.e., observations in the 117$^{\text{th}}$ Congress use lobbying up to the end of the 116$^{\text{th}}$ Congress). In the second, weights use a narrow band of Congresses before the reference Congress (i.e., observations in the 117$^{\text{th}}$ Congress use lobbying from the 112$^{\text{th}}$ through 116$^{\text{th}}$ Congresses). In the final strategy, weights are rolling but exclude the three Congresses immediate preceding the reference session (i.e., observations in the 117$^{\text{th}}$ Congress use lobbying through the end of the 113$^{\text{th}}$ Congress). Effects are comparable in magnitude and remain significant.

Dependent Variable:		Donated?	
	Rolling	Recent	Lagged
Model:	(1)	(2)	(3)
Variables			
Exposure Weight	0.088***	0.083***	0.070***
	(0.013)	(0.013)	(0.012)
Same State?	0.031***	0.031***	0.031***
	(0.006)	(0.006)	(0.006)
Legislator Characteristics:	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes
Fixed-effects			
Congress	Yes	Yes	Yes
Industry	Yes	Yes	Yes
Fit statistics			
Observations	9,571,625	9,496,194	9,341,084

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.18: *Effect of Committee Exposure on Donations, Rolling Weights:* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level. These models use rolling pre-period weights rather than a fixed pre-period. Weights are calculated, respectively, using all lobbying before the reference congress; lobbying during the four most recent Congresses; and lobbying excluding the three most recent Congresses. Effects remain large and significant.

A.10 Disaggregation by Congress, Firm-Legislator Estimates

Table A.19 presents our primary specification dis-aggregated by each Congress from the 112th (2011-2013) to present. We perform this analysis to observe that our findings are not dependent on any particular Congress, and that there is no noticeable time trend. Although there is some heterogeneity in effect size across sessions of Congress, the sign and significance are unchanged and effect size remains approximately the same. We observe an attenuated effect size in the 115th Congress (which may also stem from slightly larger average legislator weights in that Congress). Because the results are disaggregated by Congress, we drop the majority party status predictor from Models 3 and 5 (because in each Congress, majority party status is collinear with one of the party fixed effects).

Dependent Variable:	Donated?					
	112th	113th	114th	115th	116th	117th
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Exposure Weight	0.095***	0.111***	0.109***	0.061***	0.098***	0.087***
	(0.014)	(0.016)	(0.018)	(0.009)	(0.019)	(0.015)
Same State?	0.036***	0.034***	0.035***	0.034***	0.027***	0.020***
	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)	(0.005)
Legislator Characteristics:	Yes	Yes	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes	Yes	Yes
Fixed-effects						
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	1,514,401	1,563,775	1,544,655	1,684,435	1,808,126	1,387,358

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.19: *Effect of Committee Exposure on Donations, Disaggregated By Congress:* Models estimating the effect of committee exposure on probability of donation at the firm-legislator level, disaggregated by congress. Results mirror main specification results across the board and are positive, significant, and large with little heterogeneity by Congress.

A.11 Robustness to Firm Leverage, Firm-Legislator Estimates

In this section, we re-estimate our primary model, excluding industries with relatively few firms (industries for which one firm generates more than 10% of all lobbying activity). To the extent that treatment is intended to be exogenous to a firm but is composed in part of firm-level activity, then these high-lobbying firms can be viewed analogously to market price-makers, and they may disproportionately impact exposure weight estimates within their industries. We drop these industries (and thus these firms) from the analysis in Table A.20, and find that the results remain robust, large, and statistically significant.

Dependent Variables:		Donated?		Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.092***	0.066***	0.509***	0.739***
	(0.016)	(0.017)	(0.065)	(0.133)
Same State?	0.030***	0.034***	0.150***	0.240***
	(0.007)	(0.008)	(0.024)	(0.062)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	7,697,047	1,369,072	1,363,938	7,697,047

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.20: *Effect of Committee Exposure on Donations, Robustness to Excluding 'Weight-Makers':* Linear probability models estimating the effect of committee exposure on probability of donation at the firm-legislator-congress level. Compared with our primary specifications, these specifications exclude industries where exposure weights are dominated by few firms (any industry where a firm contributes >10% of exposure weights). Each model maps to @tbl-main; results are positive, significant, large, and consistent with our primary results.

A.12 Robustness to Missingness, Firm-Legislator Estimates

Our more complex model specifications, including our preferred specification, display a significant degree of data missingness, driven by limits on the availability of firm-level data in S&P Compustat, the external data source we use. In Table A.21, we describe the source of the missingness variable by variable.

Variable	**Missingness**
`nominate_dim1`	0.03%
`revt`	24.73%
`capx`	16.88%
`emp`	28.39%
`vapl`	37.21%

Table A.21: *Missingness Summary:* This table summarizes sources of missingness in our data. Missingness in nominate_dim1 stems from members who voted too infrequently to be assigned a NOMINATE score, typically because they were elected in a special election for a limited period. The overwhelming majority of missingness stems from firm-year level revenue, capital expenditure, and employee data (which are in turn used to produce the VAPL measure; VAPL cannot be produced if any of revenue, capex, and employee headcount are missing).

Missingness in our corporate data is unlikely to confound the relationship we study because, to threaten our inference, the pattern of missingness would have to correlate with both industry-committee exposure and the propensity to donate. Smaller firms might be more likely to have missing firm-level data and less likely to donate, but they are unlikely to differ systematically in their industry-committee exposure. Nevertheless, we respond to missingness using two estimation strategies: first, interpolating missing firm-year data for firms that have non-missing years. We choose linear interpolation because more complex spline-based or higher-order polynomial approaches can produce extreme values with missingness near endpoints (De Boor, 1978). Interpolation cuts the degree of missingness approximately in half. We report these results in Table A.22. Our results are not affected by this approach.

Second, we use multiple imputation to impute the missing values. Multiple imputation has historically been used by political scientists as a tool to deal with missingness (King et al., 2001). We take no position on criticisms that, under highly non-random missingness, multiple imputation may produce biased results compared to listwise deletion (Pepinsky, 2018). Rather, we observe

Dependent Variables:	Donated?			Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.081***	0.062***	0.428***	0.647***
	(0.014)	(0.015)	(0.055)	(0.110)
Same State?	0.031***	0.035***	0.148***	0.254***
	(0.006)	(0.007)	(0.017)	(0.051)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	12,373,522	2,194,746	2,348,138	12,373,522

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.22: *Effect of Committee Exposure on Donations, Missing Data Linearly Interpolated:* Models estimating the effect of committee exposure on probability (or amount) of donation at the firm-legislator-congress level. Firm-level covariates displaying missingness have been linearly interpolated. The effects are comparable to the models presented in the main paper. All results are large and statistically and substantively significant.

that our results are consistent across all three strategies for dealing with missingness. To perform multiple imputation, we select a broad basket of corporate financial variables in S&P Compustat and borrow strength from those when imputing our variables of interest.[18] This approach further reduces the missingness in our data (not to zero, because some firms lack any of the "support variables" we use in order to help impute our variables of interest). The results for this exercise are reported in Table A.23. Our results are not affected by this approach.

[18]Those variables are, by Compustat codebook name: `capx`, `emp`, `revt`, `at`, `mkvalt`, `xopr`, `ppegt`, `ppent`, `gp`, `dp`, `crv`, `cogs`, `ivncf`, and `oancf`.

Dependent Variables:		Donated?		Dollars (log)
Model:	(1)	(2)	(3)	(4)
Variables				
Exposure Weight	0.075***	0.060***	0.431***	0.603***
	(0.013)	(0.015)	(0.056)	(0.106)
Same State?	0.029***	0.032***	0.148***	0.231***
	(0.006)	(0.006)	(0.017)	(0.046)
Legislator Characteristics:	Yes	Yes	Yes	Yes
Firm Characteristics:	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Industry	Yes	Yes	Yes	Yes
Fit statistics				
Observations	14,052,789	2,484,763	2,452,920	14,052,789

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.23: *Effect of Committee Exposure on Donations, Missing Data Multiply Imputed:* Models estimating the effect of committee exposure on probability (or amount) of donation at the firm-legislator-congress level. Firm-level covariates displaying missingness have been multiply imputed. The effects are comparable to the models presented in the main paper. All results are large and statistically and substantively significant.

A.13 Robustness to Non-Linear Functional Form, Firm-Legislator Estimates

A.13.1 Logistic Regression Coefficient Estimates

Our primary functional form, for ease of reporting and interpretation, is the linear probability model. But there are reasons to believe the true functional form is non-linear; perhaps, for instance, firms measure legislator exposure in coarse terms, and rather than linearly doling out additional donations through higher exposure, they identify legislators that cross a threshold of exposure for donation. This suggests something closer to a sigmoid relationship between exposure and the probability of donation. Thus, in Table A.24, we report estimates of our specifications run as logistic regressions (e.g., with functional form $Pr(y_{i[j]f,t} = 1) = logit^{-1}(.))$. The sign and significance of our results are unchanged, and their magnitudes are higher than the linear probability models presented above.

We refrain from offering familiar guidance about the interpretations of logistic regression coefficients except to note three points of interest: first, in sections Appendix A.13.2 and Appendix A.13.3 we report the average partial derivatives of our actual data, which reflect our preferred interpretation of the logistic regressions with and without adjusting the effect size to reflect the feasibility concerns discussed in Appendix A.5. Second, our logistic regressions discard fixed effects that would otherwise result in perfect separation, which improves the numerical stability of our estimates. Third, although we offer pre-computed versions of our models in our replication archive, these models specifically cannot be computed on desktop-class computer hardware due to memory requirements.

Because we are computing a logistic regression on panel data, we comment on, and correct for, the well-known incidental parameter problem (Heckman, 1978) in Appendix A.13.4.

Dependent Variable:	Donated?				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Exposure Weight	3.29***	5.22***	4.95***	5.56***	5.29***
	(0.545)	(0.560)	(0.557)	(0.532)	(0.530)
Same State?		1.22***	1.35***	1.63***	1.74***
		(0.111)	(0.122)	(0.133)	(0.143)
Legislator Characteristics:			Yes		Yes
Firm Characteristics:				Yes	Yes
Fixed-effects					
Congress		Yes	Yes	Yes	Yes
Industry		Yes	Yes	Yes	Yes
Fit statistics					
Observations	15,425,275	15,279,858	15,279,858	9,353,694	9,353,694

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.24: *Effect of Committee Exposure on Donations, Logistic Link:* Logistic regression models estimating the effect of committee exposure on the probability of donating at the firm-legislator-congress level. Results are positive, significant, and large in magnitude.

A.13.2 Average Partial Derivatives

As discussed above, we report the average partial derivatives (APDs) of our data here. This is accomplished by producing two synthetic versions of our dataset: the first, with all other covariates held at their actual values and the weight variable set to zero for all observations, and the second with all other covariates held at their actual values and the weight variable set to a comparison level (respectively 1 and 0.219 in this subsection and Appendix A.13.3); we then calculate the treatment effect per observation (the synthetic $y_{i[j]f,t}(1) - y_{i[j]f,t}(0)$), and average across all observations. Unlike with a linear probability model, in a GLM the partial derivative differs depending on the unit because a fixed increase in the observed space transforms non-linearly in terms of increases in the probability-space of the outcome through the link function.

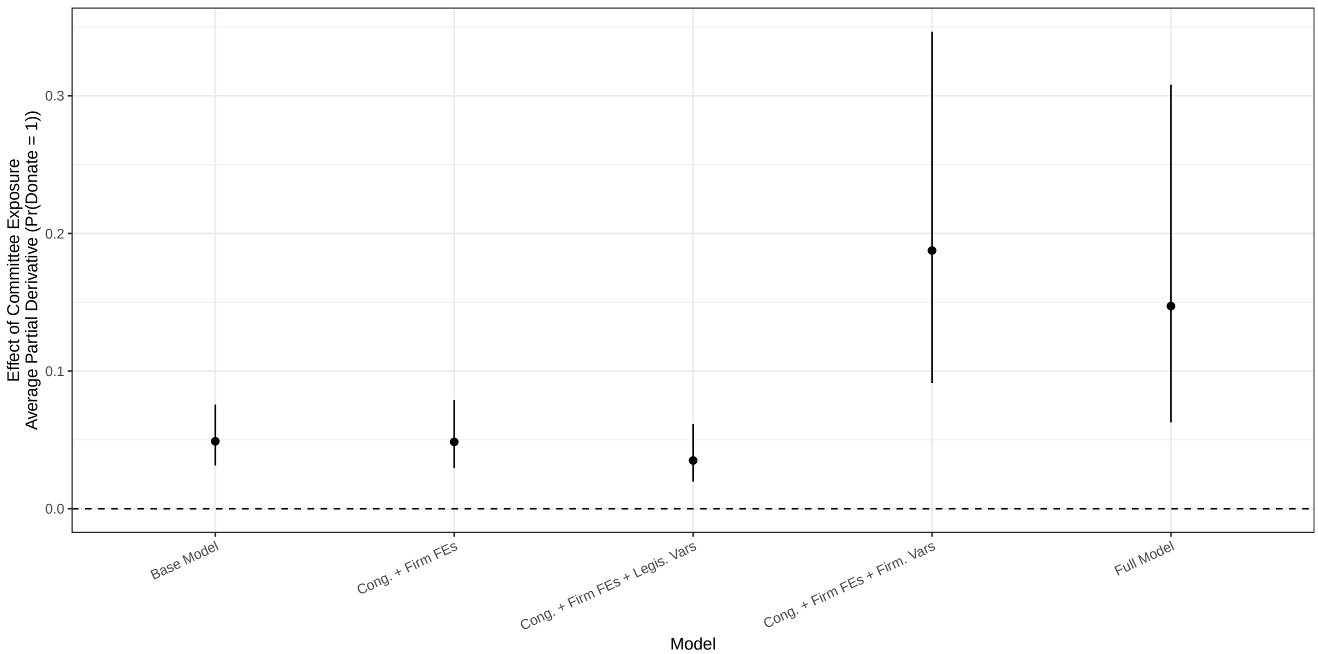

Figure A.8: *Effect plot of the Average Partial Derivative (APD)*: Synthetic versions of our data are generated, replacing the committee weight variable with low (0) and highest possible (1) values. Unit-specific partial derivatives are computed, holding all other variables at their true values. See Section A.13.3 below for an alternative computation taking into account concerns raised in Section A.5.

A.13.3 Average Feasible Partial Derivatives

In this subsection, the average partial derivative reflects a change in the treatment variable, committee weight, of 0.219, which reflects the discussion in Appendix A.5. This mechanically attenuates the effect presented, but the sign, significance, and substantive interpretation are unchanged.

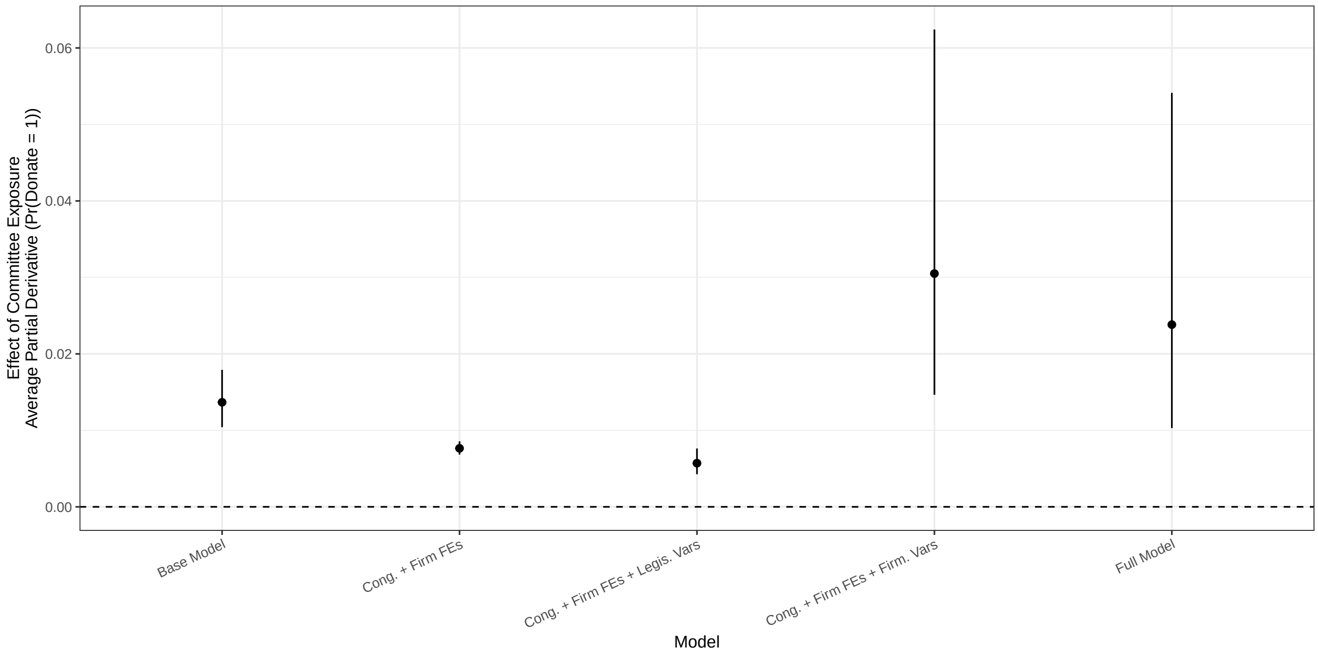

Figure A.9: *Effect plot of the Average Partial Derivative (APD), Feasible Effect:* Synthetic versions of our data are generated, replacing the committee weight variable with low (0) and feasibly high (0.203) values. Unit-specific partial derivatives are computed, holding all other variables at their true values. The effects remain substantively large and significant.

A.13.4 Incidental Parameter Problem

Because our panels are relatively short ($t = 6$ congresses) and feature unit and time fixed effect structures, our GLM results are vulnerable to the incidental parameter problem, wherein coefficient estimates are inconsistent. We implement the split-panel jackknife correction described in Fernández-Val and Weidner (2016) to correct for this problem: the correction applies to Models 2-5 in Table A.25.[19] The changes to the coefficients of interest are immaterial: the results remain positive, significant, and substantively large.

Dependent Variable:			Donated?		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Exposure Weight	3.29***	4.72***	4.48***	4.93***	4.70***
	(0.545)	(0.560)	(0.557)	(0.532)	(0.530)
Legislator Characteristics:			Yes		Yes
Firm Characteristics:				Yes	Yes
Fixed-effects					
Congress		Yes	Yes	Yes	Yes
Industry		Yes	Yes	Yes	Yes
Fit statistics					
Observations	15,425,275	15,279,858	15,279,858	9,353,694	9,353,694

Clustered (Industry) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.25: *Effect of Committee Exposure on Donations, Logistic Link, Correcting For Incidental Parameter Problem:* Logistic regression models estimating the effect of committee exposure on the probability of donating at the firm-legislator-congress level. Models 2-5 have been subject to bias corrections using the generalized split-panel jackknife, correcting for the incidental parameter problem caused by the inclusion of unit and time fixed effects in our models. The results are comparable to the original uncorrected models and remain positive, significant, and large in magnitude.

[19] Fernández-Val and Weidner (2016) provides both analytical and simulation approaches for estimating; we implement the simulation approach for reasons of expediency. In the authors' results, both corrections substantially improve coverage compared to uncorrected estimates.

A.14 Robustness, Firm-Committee Estimates

A.14.1 Alternative Fixed Effect Structures

Our primary firm-committee estimates use congress, industry, and committee fixed effects. In this appendix, we substitute an alternative fixed-effect structure, using congress + firm + committee fixed effects. The analysis is robust to this alternative specification. For space purposes we omit additional estimates which cluster standard errors by firm rather than firm × committee, but these too are positive and significant.

Dependent Variables:	Donated?		Amount (logged)	
Model:	(1)	(2)	(3)	(4)
Variables				
Treatment (proportion)	0.316***		4.27***	
	(0.032)		(0.305)	
Treatment (count)		0.009***		0.126***
		(0.0008)		(0.008)
Firm characteristics	Yes	Yes	Yes	Yes
Fixed-effects				
Congress	Yes	Yes	Yes	Yes
Firm	Yes	Yes	Yes	Yes
Committee	Yes	Yes	Yes	Yes
Fit statistics				
Observations	802,939	802,939	802,587	802,587

Clustered (Firm × Committee) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.26: *Firm-Committee Effects, Alternative Fixed Effects*: This figure presents the same models as our primary firm-committee models, but with an alternative fixed effect structure, including congress, firm, and committee fixed effects. Results remain significant.

A.14.2 Disaggregation by Congress

To verify that our results are not driven by a particular session of Congress with unusually high turnover or donation activity, and that they are not an artifact of a particular election cycle, we disaggregate our firm-committee estimates by Congress. All sessions of Congress in the sample remain positive, significant, and substantively large.

Dependent Variable:	Donated?					
	112th	113th	114th	115th	116th	117th
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Treatment (proportion)	0.424***	0.297**	0.510***	0.352**	0.307**	0.580***
	(0.118)	(0.107)	(0.139)	(0.114)	(0.113)	(0.152)
Firm characteristics	Yes			Yes	Yes	Yes
Fixed-effects						
Industry	Yes	Yes	Yes	Yes	Yes	Yes
Committee	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	130,571	132,767	135,146	142,787	139,853	121,817

Clustered (id_firm_ committee) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01*

Table A.27: *Firm-Committee Turnover Instrument by Congress*: We disaggregate our firm-committee results by congress and present our primary analysis (dependent variable: probability of donation, treatment measure: proportion of turnover). The results remain positive, significant, and comparable in magnitude across each Congress in the sample.

A.14.3 Robustness to Non-Linear Functional Form

One of our two dependent variables for the firm-committee analysis, probability of donation, has a binary dependent variable. We verify that our observed results are not sensitive to the linear probability model specification by re-estimating them using a logistic regression. The results are unchanged.

Dependent Variable:	Donated?	
Model:	(1)	(2)
Variables		
Treatment (proportion)	5.10***	
	(1.20)	
Treatment (count)		0.138***
		(0.030)
Firm characteristics	Yes	Yes
Fixed-effects		
Congress	Yes	Yes
Industry	Yes	Yes
Committee	Yes	Yes
Fit statistics		
Observations	798,767	798,767

Clustered (Firm × Committee) standard-errors in parentheses
*Signif. Codes: ***: 0.001*

Table A.28: *Firm-Committee Turnover Instrument (Logit):* We re-estimate our primary firm-committee models with the probability of donation dependent variable using a non-linear functional form. The estimates remain positive, significant, and substantively large.