Why Firms Fund Politicians: Recovering Motives with Large Language Models

Stefan Lewellen* Devin Zhang†

August 2026

Abstract

Corporate political donations are public to the dollar, but the motives behind them are unobserved, forcing the literature to estimate returns across pooled purposes. We measure corporate intent using a large language model. By reading legislative remits alongside industry definitions, the model maps the statutory jurisdiction of House committees against all industries. Because the model judges the legal mapping rather than the observation itself, the measure is deterministic, immune to look-ahead bias, and auditable. This approach evaluates variables concurrently, allowing donations to tick multiple, overlapping motives without imposing an artificial priority ordering. Applying these judgments to 285,072 PAC donations over 2000–2022 reveals that firms buy jurisdiction over their own operations, not general committee access. A committee governing the donor's industry raises the probability that the firm funds one of its members by 3.6 percentage points against a 5.8% base, whereas a committee reaching all industries alike is worth zero. Money arrives when a candidate gains jurisdiction over a donor's industry and does not leave when the jurisdiction goes. Jurisdiction alone accounts for half of corporate PAC giving; evaluated concurrently, the recorded motives together explain roughly nine-tenths of all corporate PAC contributions.

Keywords: Campaign finance, Corporate political donations, Committee jurisdiction, Congress, Large language models

JEL classifications: C81, D22, D72, G38

*Pennsylvania State University.
†Pennsylvania State University.

1 Introduction

The Federal Election Commission records every dollar a corporate political action committee gives to a federal candidate, and records nothing about why it was given. Over the twelve election cycles from 2000 to 2022, the corporate PACs of 1,322 Compustat firms sent $895 million to 2,869 U.S. House candidates in 285,072 separate transactions, each public to the dollar and the day. None of them states a reason. The central object in the economics of corporate political activity—what the firm was trying to accomplish—is the one thing the data do not contain.

A large body of work prices the donation and reports returns that are positive but modest, fragile to specification, and hard to attribute (Aggarwal et al., 2012; Cooper et al., 2010; Fowler et al., 2020). A return is a return to something, and without a measure of what the firm was buying, an estimate pools donations made for reasons that should carry entirely different payoffs. The puzzle Ansolabehere et al. (2003) pose—why firms spend so little if contributions buy policy—is a question about purpose, and looking harder at outcomes will not settle it.

Lacking a measure, the literature infers motive from coincidences between donor and recipient: a shared district, a committee seat the researcher deems relevant, a shared partisan lean. Each records an opportunity rather than a reason. The committee proxy is the most used and the least discriminating—an indicator for holding any seat fires on 90.9% of the donations studied here—and its failure is conflation rather than noise. A committee seat is worth something to a firm when the committee's remit is written about what the firm makes, sells or operates: the Food, Drug, and Cosmetic Act to a drug manufacturer, the Mine Act

to a coal producer. It is worth much less when the committee reaches the firm only through law that reaches every firm alike—antitrust, bankruptcy, ordinary employment. Both are committee seats, and no dummy built from a roster separates them.

Separating them at scale requires reading the intersection of firm operations and committee jurisdiction directly. To recover these motives, we deploy a Large Language Model (LLM) to read the legislative remit of every committee alongside the business description of every donating firm. Crucially, rather than forcing donor intent into a strict hierarchical classification, the model evaluates variables concurrently to tick multiple, overlapping motives. This multi-motive approach captures the reality that a single donation can simultaneously target industry-specific jurisdiction, local district operations, and party leadership without imposing an artificial priority ordering.

Firms buy the first kind of jurisdiction and not the second. Holding the firm and the candidate fixed within the cycle and entering both kinds of reach together, a committee with jurisdiction over the donor's own industry raises the probability that the firm funds that candidate by 3.6 percentage points on a base rate of 5.8%, while a committee reaching the donor only as it reaches everyone is worth -0.22 points, indistinguishable from zero. The purchase is a choice of candidate rather than a sum of money: conditional on giving at all, the same two kinds of reach move the amount by 14.1% and -1.2%. And the money follows the jurisdiction in one direction only. When a candidate gains a committee with jurisdiction over a donor's industry, the probability that the donor gives rises 4.6 points and stays up; when the same candidate loses that jurisdiction but stays in the House, nothing happens at all.

Counted together, these concurrent reasons leave little of the money unexplained. Juris-

diction over the donor's own industry alone accounts for 53.6% of donations and 54.5% of dollars, and adding the district where the firm operates, the tax and spending committees, and party leadership takes the total explained to roughly nine-tenths. The residue is itself informative: 70.9% of it backs candidates who hold no committee seat at all, where a bet on a future member is the only thing on offer, and among donations to committee-holding incumbents—the case the access literature is about—3.4% are unexplained.

Measuring jurisdiction against access requires a fact no dataset contains. Which committees govern which industries is written in the Rules of the House and in the statutes themselves, and it has never been tabulated against an industry classification. Section 3 builds that table with a large language model. The model judges, once per cell and in isolation, whether each of nineteen committees reaches each of seventy-six industries through a statute written about that industry, naming the statute wherever it credits one; the 285,072 donations then inherit the 1,444 judgments by mechanical lookup. We call this design *judging the mapping rather than the observation*, and its guarantees hold by construction rather than by argument: no cell is judged twice, so identical inputs cannot receive different labels; the question contains no firm, no candidate, no date, and no election, so the model cannot look ahead; and the output is small enough to publish whole, so a reader disputes a cell and its named statute rather than a black box. Misclassification between the two committee codes drives their estimated effects toward each other, so the contrast the paper reports is a lower bound on the true one. Everything the measure consumes is dated to what a donor could observe—the committees of the sitting Congress, the district map on the ballot—and Section 2 sets out the institutional facts behind those conventions.

Section 4 builds the other thing the question needs: the alternatives. A file of donations

records only the gifts that happened, so it can say how large a gift was but never why that candidate rather than another. For every firm-cycle, the choice set therefore lays out every House candidate who drew corporate PAC money that cycle—4,880,936 cells, on 5.84% of which a donation occurred—with the motive variables filled in for the candidates the firm passed over, and it reconstructs the realized donations exactly before it is used.

Section 5 reports the composition behind the headline shares and what the standard proxies do to them. The committee dummy overstates the committee motive by a factor of 1.70, and one-directionally, since jurisdiction implies a seat. The geographic proxies understate, and the state-level one fails in a way an aggregate check would pass: a home-state indicator fires on 14.5% of donations against 12.8% for a measure counting the firm's local operations, defined by a 250-employee district cutoff—levels two points apart—yet agrees with that measure on only 80.1% of donations, because most qualifying operations sit outside the headquarters state. A study using both proxies gets the ratio between the two motives wrong by roughly an order of magnitude.

Section 6 shows why none of this is visible without the cross-classified structure. Pooled across candidates, jurisdiction is associated with 3.2% more money; adding firm-by-cycle effects leaves 2.4%, indistinguishable from zero; only when candidate-by-cycle effects force the comparison across donors from different industries funding the same person does the coefficient reach 14.6%. Candidates whose committees reach many industries are senior, safe, and heavily funded from every direction, and a pooled regression mixes the jurisdiction premium with whatever else makes them attractive—a caution for any study that uses a committee proxy without conditioning on the recipient.

The same section stress-tests the central distinction. Restricting to candidates who hold

at least one seat, so the comparison is a seat with jurisdiction against a seat without, barely moves the estimates; the effect is monotone in the number of committees with jurisdiction and concave in it; and compounding both margins over the whole choice set, jurisdiction multiplies the dollars a firm sends a candidate by 1.9 and a presence in the district by 3.3.

Section 7 turns the asymmetry into an event study. The probability of giving is flat in the cycles before a candidate gains jurisdiction over a donor's industry, jumps 4.6 percentage points in the cycle the assignment takes effect, and holds for at least three cycles after; losing the jurisdiction while staying on committees changes nothing, and money leaves only when the member leaves the House. Donation relationships are formed on jurisdiction and maintained on something else—which also supplies a candidate account of the giving that jurisdiction cannot explain at a point in time.

Two further patterns cut against one-mechanism stories. Estimated cycle by cycle, the pull of jurisdiction rises across the window while the pull of geography falls, though geography remains the larger throughout. And the mix shifts with the scale of the operation: among firms funding a median of three candidates, a presence in the district raises the probability of a gift by eighteen times the base rate and dwarfs every other reason, while among firms funding a median of eighty-eight, geography's edge over jurisdiction nearly disappears.

The paper speaks to three literatures. Kroszner and Stratmann (1998) explain why the institution takes this shape—standing committees with fixed jurisdictions let legislators build reputations with the interests they regulate, so money should track jurisdiction rather than office—and Powell and Grimmer (2016) come closest to testing it, showing that contributions follow members onto the committees they join after involuntary transfers. We supply the subject-matter side the test has lacked, the same side Bertrand et al. (2014) show is decisive

for lobbying, and the data find for jurisdiction. To work on corporate political activity we supply a measure of what a donation was for, built without reference to what it achieved, so that returns can be attributed to motives rather than estimated across them. And to research using language models in empirical economics (Mullainathan, 2025), we offer a design rather than an application: judging a small, closed, publishable mapping once, and looking observations up in it, delivers determinism, immunity to look-ahead, auditability and scale as structural properties rather than as claims a reader is asked to accept.

The paper has no experiment claims. Party steering committees make assignments, and a member seeks the committee that matters to the industries in the district and to the donors the member already has; no fixed effect removes that. The within-pair estimates bound how much of the association could run backwards, since the assignment is dated to the Congress sitting before the donation and a firm's own industry almost never moves, but they do not make the assignment random. The measure records whether a reason was available rather than whether it was operative. And its external validation—testing the grid against the committees firms name in their own lobbying disclosures—remains to be done.

2 Institutional Background

Three institutional facts govern what a corporate donation can be for, and each one narrows the set of reasons a firm could plausibly have. Corporate money reaches candidates through a vehicle whose size is capped and whose donor base is fixed by law. The candidate's capacity to act on a firm's behalf runs through committee jurisdiction, which is assigned by the parties and is sticky across Congresses. And the donor decides before the election, so anything the

election determines is unavailable at the moment the money moves. This section sets out the three; Section 3 builds the measure they imply.

2.1 Corporate PACs and What the Law Allows

A corporation may not give corporate treasury money to a federal candidate, and has not been able to since 1907. What it may do is administer a separate segregated fund—a political action committee—financed by voluntary contributions from its own executives and shareholders, and give that money away in its own name. The Federal Election Campaign Act caps a PAC's gift at $5,000 per candidate per election, so $10,000 across a contested primary and general, and the cap has never been indexed. Two consequences follow for what the data can show. The amounts are small relative to anything a regulatory decision is worth, which is the puzzle Ansolabehere et al. (2003) pose and which no subsequent work has dissolved: if contributions bought policy, firms would buy far more of it than they do. And the amounts are lumpy: 59.0% of the donations studied here sit on exactly $1,000, $2,000, $2,500 or $5,000.

The puzzle has generated two literatures that this paper sits between. One prices the donation, asking what a firm's political spending is worth to its shareholders, and finds returns that are positive but modest and hard to attribute (Aggarwal et al., 2012; Cooper et al., 2010), or that narrow sharply once the comparison is made within close elections (Fowler et al., 2020). The other studies the adjacent technology of lobbying, where the sums are an order of magnitude larger and the object purchased is better identified (Bombardini and Trebbi, 2020; Kang, 2016). Both take the donation as given and ask what it buys. This paper asks the prior question of what the firm was trying to do, on the ground that a return

cannot be attributed to a motive that has not been measured.

The lumpiness matters for where a motive can be looked for. If a firm's reason for giving showed up in how much it gave, the reason would have to move a variable that takes four values on three-fifths of the sample and is bounded above by statute. The choice of *whom* to fund carries no such constraint: a firm faces several hundred candidates each cycle and funds a few dozen. Section 6 finds the motives almost entirely on that margin, which is what the institutional setting predicts.

2.2 House Committees and Jurisdiction

A member of the House affects an industry through the committees on which the member sits. Committees hold the power to originate, amend and bottle up legislation in their own subject areas, and those areas are set by the Rules of the House rather than negotiated deal by deal. Energy and Commerce writes the Food, Drug, and Cosmetic Act and the Clean Air Act; Financial Services writes bank regulation and securities law; Agriculture writes the farm bill and the commodity exchange statutes. A firm that wants a hearing on a rule affecting its product is asking a member of the committee that governs that rule, not a member of the House at large.

Jurisdiction is what makes a committee seat worth something to a particular firm, and committee seats are not interchangeable in this respect. Kroszner and Stratmann (1998) show why the institution takes this form: standing committees with fixed jurisdictions allow legislators to build reputations with the interests they regulate, and the repeated dealing that follows is what makes a contribution informative rather than a bribe. The prediction

that money should track jurisdiction rather than office is theirs; what has been missing is a measure of jurisdiction fine enough to test it against the alternative that money tracks access. Powell and Grimmer (2016) come closest, using involuntary committee transfers to show that contributions follow members onto the committees they join.

Two facts about how seats are assigned make the resulting variation usable. Assignments are made by the party steering committees at the start of each Congress, after the election that the donation precedes, so a donor observes the seat a member currently holds rather than the one the member will hold (Kellermann and Shepsle, 2009). And assignments persist: across the 2000–2022 window, 2,927 of 7,898 candidate–cycles show any change at all in the committee portfolio, so the seat a donor observes is usually the seat that survives.

Ways and Means and Appropriations sit apart from the other committees and have to be treated apart. Ways and Means holds taxation, tariffs, Social Security and Medicare; Appropriations holds the discretionary budget, through subcommittees whose membership the data do not record. Neither has an industry-specific remit to test a firm against, because every corporation has a tax interest and every corporation can be a claimant on federal spending. Both are also exclusive under House rules: a member assigned to either generally holds no other standing committee, which is why a member of Ways and Means looks, on every industry-specific measure, like a member with no relevant jurisdiction at all.

2.3 What the Donor Observes

The donor decides before the election, and the analysis holds itself to what the donor could have known. Committee variables are dated to the Congress sitting when the donation

was made, never to the Congress seated afterwards: for the 2012 cycle, the committees are the 112th Congress's. Using the incoming Congress would require assuming that both the election outcome and the subsequent assignment were predictable in advance, and would put information realised after the donation into a variable meant to explain it. Congressional district boundaries run the other way and are taken from the incoming Congress, because a candidate runs for a seat defined by the new map and redistricting is completed and published in law before the vote. The two conventions differ because they assume different things, and the pair is the combination that assumes least.

Election outcomes are excluded outright. The sample carries no vote share, no winner and no subsequent legislation, and it is built that way at source rather than filtered afterwards. Nothing in the measure can be a function of how the race turned out, which is what allows the second half of this research programme—motives predicting outcomes—to be asked without circularity.

3 Measuring the Reason for a Donation

A donation record shows a firm, a candidate, an amount and a date. It does not show the reason, and no amount of care with the record will make it. This section builds a measure of the reason from outside the record, states what the measure can and cannot establish, and explains where a language model does the work and where it is deliberately kept out.

3.1 Why Revealed Correlations Do Not Identify a Reason

The literature infers motive from revealed correlations, and the standard proxies conflate the reason for a donation with the opportunity to make one. Three are in common use: whether the candidate's district contains the firm's headquarters, whether the candidate sits on a committee the researcher judges relevant, and whether the candidate shares the firm's partisan lean. Each is a fact about a coincidence between firm and candidate, and each is compatible with several different reasons for the money. A firm headquartered in a committee chair's district loads on all three at once, and no combination of them says which one moved the gift.

The distinction the proxies cannot draw is the one Bertrand et al. (2014) draw for lobbying, where the question is whether a lobbyist is retained for connections to particular legislators or for expertise in particular issues, and the answer requires observing both the lobbyist's contacts and the subject matter of the bills. Donations have had no equivalent, because the subject-matter side has never been measured: a committee dummy records that a legislator sits somewhere, not that the somewhere governs this donor.

The committee proxy is the one this paper repairs, and its defect is specific rather than general. An indicator for holding a committee seat fires on 90.9% of the donations studied here, because almost every incumbent sits on something. What a firm cares about is not whether the candidate holds a seat but whether the seat governs the firm's own business, and the two coincide on only 59.0% of the donations where the candidate holds any seat. The proxy is wrong in one direction only—jurisdiction over an industry implies a seat, so every disagreement is a false positive—which means it does not merely add noise but inflates the

committee motive systematically.

The geographic proxy fails differently, and the failure is easier to miss. A firm's headquarters is one address; its operations are many, and a candidate's district may contain a plant, a distribution centre or a call centre while the head office sits a thousand miles away. An indicator built from the headquarters ZIP recovers 18% of a measure that counts operations; a state-level indicator recovers a share close to one in the aggregate and agrees with it on only 80.1% of donations. Section 5.2 takes that apart, because a proxy whose *level* is right and whose *classification* is wrong is the kind an aggregate check would pass.

3.2 Five Reasons, Judged Independently

The measure records five reasons a firm might have for funding a particular candidate, each a separate indicator, none ranked above another:

- **Committee: industry jurisdiction.** At least one of the candidate's sitting committees reaches the donor's own industry through a *sectoral instrument*—a statute, programme or regulatory regime whose scope is defined by what that industry makes, sells or operates.
- **Committee: economy-wide jurisdiction.** A sitting committee reaches the donor's industry with identifiable force, but through a legal function that applies to firms in every sector: antitrust, bankruptcy, ordinary employment law.
- **Committee: tax and spending.** The candidate sits on Ways and Means or Appropriations.

- **Geography.** The firm's headquarters, or an establishment employing 250 or more people, sits in the district being contested.
- **Social tie.** An executive or director of the firm and the candidate attended the same institution within five years of one another.

Nothing forces exactly one of these to hold, and the design would be worse if it did. A candidate who represents the district containing a firm's largest plant *and* chairs the committee that regulates its product is a different case from one who does only the first, and shows up as different only if both can be ticked. Earlier versions of this taxonomy imposed a priority ordering and resolved multi-reason donations by fiat; that convention silently determined the size of the geographic category, and it was dropped. The categories are reported jointly throughout, and where a single ordering is needed for a summary display it is labelled as a reporting convention.

Incumbency is not among the reasons, and its exclusion is a substantive choice. Nearly every candidate in the data is either an incumbent or a contestant for a seat, so a reason that fires on incumbency would explain every donation equally and therefore none of them. An earlier version of this taxonomy offered incumbency as a category; it absorbed 29.2% of a 1,000-donation trial sample and carried no internal gradation whatever, which is what a category assigned by elimination looks like. Incumbency is retained as a candidate characteristic and used as such.

3.3 Jurisdiction Over an Industry, and Jurisdiction Over Everyone

The distinction between the first two reasons is the measure's substantive content, and it is a distinction no committee dummy can draw, because both are committee seats. The operational test is whether a firm in an unrelated industry would describe the same stake in the same words. Clean Air Act Title II certifies an engine as a product before it may be sold and reaches only firms building engine-powered equipment, so Machinery × Energy and Commerce is sectoral. Clean Air Act facility permitting reaches every manufacturing plant in the economy, so where that is all a committee has, the cell is economy-wide. The seam runs through the statute rather than around the committee.

Sectoral and economy-wide are properties of the industry–committee pair, not of the committee. Judiciary is sectoral for Pharmaceuticals, where Hatch-Waxman and the Controlled Substances Act are written for that industry alone, and economy-wide for Electric Utilities, where antitrust and bankruptcy are all there is. Even Agriculture, about as sectoral a committee as exists, reaches eight industries only through commodity-derivatives law, which touches every hedger. Appendix A3 gives a worked example at each code boundary.

The measure is therefore a grid: for each of 76 industries at the GICS six-digit level and each of 19 House committees, a code recording whether that committee reaches that industry through a sectoral instrument (400 cells), reaches it only as it reaches every industry (305 cells), or does not reach it beyond what every firm in the economy faces (739 cells). Every credited cell names the instrument. The line between the second code and the third is precisely the requirement to name one and show it bites this industry particularly, because every committee touches every industry in some remote sense and that is not a reason to

give anybody money. The nineteen committees are the House standing committees less Ways and Means and Appropriations, which are supplied mechanically, plus Permanent Select Intelligence; Appendix A3 records why each inclusion and exclusion was made.

Three counts summarise a donation and are all that the row-level file stores: how many of the candidate's sitting committees are sectoral for this firm's industry, how many are economy-wide, and how many are neither. Everything else derives from them, so nothing can drift out of step when the grid is revised. The third count earns its place by separating two cases a binary confounds: a candidate holding no gradeable committee at all had no opportunity to have jurisdiction, while a candidate holding three that reach nothing had the opportunity and lacked the jurisdiction.

3.4 What the Language Model Is Used For, and What It Is Not

Mullainathan (2025) frames the division of labour between human and algorithmic judgement in terms this measurement problem instantiates almost exactly: the human advantage lies in recognising what is missing from the data, and the algorithmic advantage in extracting information from data the human cannot read at scale. The FEC record contains a firm, a candidate, an amount and a date. The reason is missing from it and always will be. The information that would supply the reason—which statutes a committee writes, and which industries those statutes govern—sits in a corpus of law and institutional practice that no table encodes and that a researcher cannot traverse 1,444 times at a standard a referee can check.

Three features of this particular question make a language model worth more here than

in the applications where such models are usually deployed, and they are worth separating because each answers a different objection. *First, the missing variable is documented rather than latent.* Most unobservables in empirical finance—managerial ability, product quality, the corporate purpose that Nikolov et al. (2025) recover with a language model—are latent constructs that no document records, and a model reading text can only proxy for them. Committee jurisdiction is not latent. It is written in the Rules of the House and in the statutes themselves, and it is missing from the researcher's table for the mundane reason that nobody has ever tabulated it against an industry classification. The task is retrieval and structuring, not inference, which is the task at which the technology has an unambiguous advantage. *Second, the object to be judged is small and closed.* Nineteen committees and seventy-six industries exhaust it. The model's entire output is 1,444 cells, which is short enough to print, so a reader audits the measure rather than trusting it—a property that per-document classification, whose output is as large as the corpus, structurally cannot offer. *Third, the judgement is time-invariant.* Whether Energy and Commerce's remit reaches pharmaceuticals is not a fact about 2012, so a model that has read about 2012 knows nothing about the answer that it should not.

Measuring jurisdiction is an *estimation* problem in Mullainathan's sense rather than a *prediction* problem, and the distinction is worth drawing because the two carry different econometric requirements. In a prediction problem the researcher uses a model to predict an economic variable from associated text, and the model's apparent accuracy is only informative if its training corpus does not overlap the researcher's test sample—the "no training leakage" condition. In an *estimation* problem the researcher has in mind a resource-intensive measurement procedure that would be satisfactory if it could be scaled, and wants

the model to execute it. Reading the jurisdiction of nineteen committees and deciding, industry by industry, whether the remit reaches the industry through a statute written about it is such a procedure. It is not hard for one cell. It is infeasible for all of them.

The model judges the grid, and never a donation. The unit of judgement is the (industry, committee) pair, presented in isolation, with the instrument required on the record for every credited cell. The row-level variable is then a mechanical lookup of the candidate's sitting committees against the donor's industry. Three properties follow that per-row classification cannot have.

Determinism by construction. Mullainathan's caution about language models is that they are brittle—performance varies across tasks in ways that are hard to anticipate—which in a per-row design becomes a worry about whether the same record would be labelled the same way twice. Here identical inputs cannot receive different labels because they are not judged twice: there is one judgement per cell, and 285,072 donations inherit 1,444 of them. Batch order, context length and the labels of neighbouring rows have nothing to act on, because each cell is judged in a context containing no other cell (Appendix A3).

No look-ahead by construction. The binding design constraint in this setting is that the classifier must use only what a donor could have known, and restricting the prompt does not control what a model already knows. The grid sidesteps the problem rather than promising to resist it: the question put to the model contains no firm, no candidate, no amount, no date and no election. Whether Energy and Commerce's remit reaches pharmaceuticals is the same fact whether or not any pharmaceutical firm ever gave anybody money, so there is no outcome for the model's knowledge of the period to leak into.

Auditability at the level of the claim. A grid of 1,444 cells is small enough to publish

whole. A reader who doubts the committee share does not have to accept or reject a number; they can find the cell that produced it, read the instrument named there, and disagree with that. Per-row classification of 285,072 donations offers a reader no such handle, because the object to be checked is as large as the dataset.

Scale follows from the same design rather than being traded against it. Because the judgement is at the industry level, the same 1,444 cells serve every donation in the sample and would serve a larger sample unchanged. The alternative is not merely more expensive; classifying 285,072 records in batches produces labels that depend on how the batches were cut, and that dependence cannot be removed by spending more.

Where the model is deliberately not used. Geography and tax-and-spending membership are settled by the data and are supplied to the analysis already filled: the first from a firm–cycle–district panel built from establishment records and congressional maps, the second from the committee list. An inference the model does not need to make is an inference that cannot be irreproducible. The restriction also keeps the validation honest, because any test of whether the model adds information beyond the mechanical indicators runs on the judged field alone rather than on a set that includes facts the model was handed.

This is a narrower use of a language model than most applications in finance, and the narrowness is the point. One strand puts algorithms to work as predictors, mapping many inputs to an outcome (Gu et al., 2020) or to a decision a human would otherwise make (Erel et al., 2021); another measures a construct document by document (Bertomeu et al., 2025); a third studies what the technology does to the people who use it (Brynjolfsson et al., 2025; Noy and Zhang, 2023). The use here is none of these. The model is asked to structure a body of institutional knowledge once, and the structure is then applied mechanically. It is

used where a mechanical solution does not exist and kept out where one does. Per-row model judgement survives for a single field—the social tie, which depends on two biographies and cannot be reduced to a lookup—and it fires on 3.2% of a 1,000-donation trial sample.

Whether the model earns its place is settled by what the alternative would have delivered. The alternative to the grid is not a slightly noisier measure of the same thing; it is the committee dummy the literature uses, which Section 5.2 shows overstates incidence by a factor of 1.70 and Section 6.3 shows pools two variables with opposite empirical content—one worth a two-thirds increase in the probability of a donation, the other worth nothing. Substituting the dummy for the grid does not attenuate the paper's central finding; it averages the finding away. That is the sense in which the marginal product of the model is large here and would be small in an application where a serviceable proxy already exists.

3.5 What the Design Buys, and What It Does Not

No ground truth exists for which committee has jurisdiction over which industry, and the measure should be judged accordingly. There is no register of correct answers against which 1,444 cells could be scored. Consistency therefore matters more than precision: what a reader is entitled to is one standard applied uniformly, an instrument named wherever a cell is credited, an output small enough to inspect whole, and a measured statement of how much the results move under a defensible alternative coding. The grid was built three times under successively tightened standards and the diffs are retained, so that sensitivity is measurable rather than asserted; Appendix A3 reports the protocol, the prediction written before the run and what the run returned against it.

Three limitations bind and are worth stating before the results rather than after. The measure records whether a reason was *available*, not whether it was *operative*: a firm may employ people in a district and give to that district's member for some other reason entirely, and the grid cannot adjudicate. Validation is external and outstanding—a firm's lobbying disclosures name the bills it worked and therefore the committees it believes have jurisdiction over it, which is revealed preference over exactly the object the grid encodes, and the links between a firm's donations and its lobbying are close enough that the two should agree (Kim et al., 2025; Tripathi et al., 2002). Testing the grid against those disclosures net of industry and committee fixed effects is the check that would settle the question, and lobbying data must stay out of the grid's construction for the test to mean anything. And a language model asked about a period it has read about cannot be instructed into ignorance; the design confines the model to a question whose answer does not depend on the period, which bounds the problem without eliminating the category.

What the design does buy is a measure that separates cases the alternatives conflate. The three-way code distinguishes a committee that does not touch an industry from one that touches it the way it touches everyone, which a binary cannot. Committee portfolios receive different values depending on who is donating—54.2% of candidate–cycles with more than one donor differ in the jurisdiction indicator across their donors, and 57.1% in the count of committees carrying it—so the measure is not a relabelling of candidate quality. Section 6 shows the two committee codes carry entirely different weight in explaining where money goes, which is the strongest evidence available that the distinction is real rather than a taxonomy imposed on indifferent data.

4 Data

4.1 Sources and Sample Construction

The sample is every donation from a Compustat firm's corporate PAC to a U.S. House candidate over the twelve election cycles from 2000 to 2022, one observation per firm–candidate–cycle: **285,072 donations from 1,322 firms to 2,869 candidates**, totalling $895 million. Contribution records are aggregated to the cycle from the FEC itemised files. Candidate identity, party, district and incumbency status come from the Database on Ideology, Money in Politics, and Elections (Bonica, 2014, 2024), from which only the pre-election fields are taken. Committee assignments come from the Stewart–Woon congressional committee series (Stewart and Woon, 2024), lagged to the sitting Congress as Section 2.3 describes. Firms are restricted to the Compustat universe because the second half of this research programme requires accounting data.

Geography is built rather than read off an address, because a firm's headquarters is one location and its operations are many. A firm–cycle–district panel assigns each firm to every congressional district containing its headquarters or an establishment employing 250 or more people, with establishments placed by point-in-polygon against the congressional map the cycle elected. The employment floor is a materiality rule and it does real work: without it a firm's establishments span a median of 57 districts, which is not a geographic tie to anybody, and at 250 employees the median is 2. Appendix A2 gives the construction, including the vintage rule for the boundary files and why headquarters are taken from filing headers rather than from the establishment source.

Donations at or below zero are retained rather than cleaned away—1,582 net refunds

and 2,888 exact zeros—and are flagged and dropped only where a logarithm requires it. Appendix A1 defines every variable the paper uses, with its source, under the name it carries in the replication files.

4.2 The Choice Set

A file of donations records only the gifts that happened, and can therefore say how large a gift was but never why that candidate rather than another. Recovering the second requires the alternatives. For every firm–cycle, the choice set lays out every candidate–cycle in the same cycle, with the motive variables filled in for the candidates the firm passed over as well as those it funded: **4,880,936 firm-cycle × candidate-cycle cells**, on 5.84% of which a donation occurred.

The candidate universe is the corporate PAC market rather than the ballot. A candidate enters a cycle's choice set if at least one sample firm funded them that cycle, which gives a mean of 658 candidates per cycle. Candidates no corporate PAC funded at all are excluded, on the ground that they are not alternatives any of these firms plausibly weighed; including the several hundred such candidates per cycle would inflate every coefficient by adding cells nobody considered. This is the one modelling choice in the construction and it is stated in every table that uses the sample.

The motive variables for the passed-over cells are rebuilt from primitives—the grid, the firm–cycle–district panel and the candidate's committee list—using the code that produced them for the realised cells, and the build halts unless the reconstruction reproduces all 285,072 donated rows exactly on all four. Anything short of exact agreement would mean the two

samples measure different things, and the counterfactual cells would not be comparable to the realised ones. Appendix A2 sets out the step.

4.3 Variables and Summary Statistics

Three features of the donation panel shape everything that follows, and all three are visible in Table 1. The typical gift is $2,000 against a mean of $3,140, and the distribution is concentrated on round numbers as Section 2.1 describes. Relationships are persistent: half of all donations go to a candidate the firm funded in the previous cycle, and 52.4% are followed by another gift to the same candidate two years later. And the money goes overwhelmingly to sitting members—91.9% of donations, against 5.2% to open-seat candidates and 2.9% to challengers—which is why the analysis is largely an analysis of incumbents whatever else it controls for.

The social tie is the one row of Table 1 not measured on the full sample, and it is reported at the trial-sample rate with the count in the table. Establishing a tie requires two biographies rather than a lookup, so it is the one field that cannot be scaled by the grid design; at the time of writing the full-sample pass is incomplete. On a 1,000-donation random subsample, an executive or director of the firm and the candidate attended the same institution within five years of one another on 3.2% of donations. The ties that do appear run through outside directors rather than officers: of the 62 shared-institution pairs in the audited subsample, 60 involve a board member. Social ties are therefore reported and not analysed, and they are absent from the ladder of Section 5.

5 What Corporate PAC Money Is For

Most corporate PAC money has a reason attached to it, and the reason is usually jurisdiction over the donor's own business. This section establishes the composition before Section 6 asks whether the reasons actually move the choice of candidate.

5.1 The Explanation Ladder

Assigning each donation to the most firm-specific reason that applies to it leaves a tenth of corporate PAC giving unaccounted for, and Figure 1 and Panel B of Table 2 report how the rest divides. The ordering is a reporting convention—the reasons are unranked indicators and several fire on the same donation—and its purpose is to make the last band interpretable: what remains is what survives *every* reason rather than what survives one of them.

Committee jurisdiction over the donor's own industry accounts for 53.6% of donations and 54.5% of dollars, and geography for a further 6.6% and 7.6%. Together the two firm-specific reasons cover 60.2% of donations and 62.2% of dollars. The two firm-invariant reasons add 22.3% of donations: a seat on Ways and Means or Appropriations covers 20.0%, and a party leadership post a further 2.3%. Committee reach common to every industry covers 6.7%.

What is left is 10.7% of donations and 8.7% of dollars. That residue is considerably smaller than a two-reason accounting suggests—restricting attention to industry jurisdiction and geography alone leaves 39.8% of donations unexplained—and the difference is entirely the firm-invariant reasons, which a measure built to be firm-specific would discard as uninformative. They are uninformative about *which* firm gives to a candidate and highly informative about *whether* anybody does.

The residue is also concentrated where jurisdiction cannot exist. Of donations in the Others category, 70.9% go to candidates holding no committee seat at all: challengers and open-seat candidates, for whom a bet on a future seat is the only thing on offer. Among the 258,195 donations to committee-holding incumbents, only 3.4% have no reason of any kind. The unexplained share of corporate PAC giving is small, and it is small in a way that makes sense: it sits almost entirely on people who are not yet in Congress.

Others also looks different from the rest in the one dimension that does not enter its definition. Donations in the category are far less likely to be part of a continuing relationship—20.2% follow a gift to the same candidate in the previous cycle, against 50.3% in the sample at large. That is what a speculative gift to a newcomer looks like, and it is evidence against reading the residue as a set of established relationships the measure has failed to explain.

5.2 The Standard Proxies Miss in Opposite Directions

Table 3 sets the proxies used in the literature against the measures used here, and the two panels fail in opposite directions. An indicator for holding any committee seat fires on 90.9% of donations against 53.6% for jurisdiction over the donor's industry, overstating the committee motive by a factor of 1.70. The error runs one way, since jurisdiction implies a seat, so every disagreement is a false positive: of donations to a candidate holding any committee seat, 41.0% reach one with no jurisdiction over the donor's business.

The geographic proxy understates by a similar factor and is more dangerous, because its aggregate is close to right. A headquarters-district indicator fires on 2.3% of donations against 12.8% for a measure counting operations, recovering 18% of it. A home-*state* indicator

fires on 14.5%—within two percentage points of the operations measure—and agrees with it on only 80.1% of donations. The reason is visible in the disagreements: 70.9% of the donations with a qualifying presence in the district come from establishments outside the firm's headquarters state, and 74.4% of same-state donations carry no qualifying presence in the contested district at all. A researcher who checked the state proxy against an operations benchmark in the aggregate would conclude it was adequate, and would be misclassifying one donation in five.

The two errors do not cancel. A study using both proxies overstates the committee motive by 1.70 and understates the geographic motive by 5.6, so the ratio between them—which is the quantity a reader takes away from a decomposition of this kind—is wrong by roughly an order of magnitude.

5.3 Where the Money Sits, by Committee and by Sector

Figure 2 reports, for each of the nineteen graded committees, where the corporate money its members receive comes from. The grid never observes a donation, so the alignment is a consistency check rather than a construction, and it is a demanding one. Energy and Commerce draws 95.8% of its $153 million from industries it reaches through a sectoral instrument. Armed Services draws 78.3% of $100 million, Transportation and Infrastructure 69.4% of $105 million, Financial Services 65.1% of $131 million.

Three committees anchor the other end, and they were named in advance. Budget, Rules and Ethics carry no sectoral cell for any industry in the grid, because none of the three has an industry-specific remit to have one. The prediction on record before the grid was built

was that they would return all zeros, and they do. They also receive $65 million, $23 million and $19 million respectively—money whose reason jurisdiction cannot supply, and which the ladder assigns to the tax-and-spending, leadership or Others categories depending on what else the member holds.

Sectoral composition follows the same logic without being told to. Table 4 orders the eleven GICS sectors by the share of their donations reaching a committee with jurisdiction over their own industry, and the ordering is recognisable: Energy at 62.5% and Communication Services at 59.0% at the top, Consumer Staples at 48.3% and Consumer Discretionary at 48.6% at the bottom. Geography runs almost exactly the other way—30.4% for Consumer Discretionary and 25.4% for Consumer Staples against 5.7% for Communication Services—which is what retail footprints and district-level employment should produce.

Firms give past the odds on both margins. A high jurisdiction share is not by itself evidence of targeting, since a firm in a heavily regulated industry would post one by giving at random; the benchmark is the share of the cycle's own choice set that carried jurisdiction over that firm. Figure 3 plots the two. The average firm-cycle realises a jurisdiction share 17.7 percentage points above what the choice set offered and a geography share 12.7 points above, and the gap is positive in every sector for both reasons.

6 Which Candidate a Firm Funds

6.1 Specification

The paper has no experiment and does not claim one. What the data have is a cross-classified panel: the same candidate receives from firms in many industries, and the same firm gives to many candidates, so two sets of fixed effects can be imposed at once. Estimates take the form

$$Y_{fct} \;=\; \beta_1 \, J_{fct} \;+\; \beta_2 \, G_{fct} \;+\; \alpha_{ft} \;+\; \gamma_{ct} \;+\; \varepsilon_{fct}, \tag{1}$$

where Y_{fct} is an outcome for firm f and candidate c in cycle t, J_{fct} indicates that a sitting committee of c reaches f's industry through a sectoral instrument, G_{fct} that f has a qualifying presence in c's district, α_{ft} is a firm $\times$ cycle effect and γ_{ct} a candidate $\times$ cycle effect. Standard errors are clustered two ways, on firm and on candidate; both dimensions repeat, and clustering on either alone understates the other by roughly a factor of three.

The two fixed-effect sets do different jobs and the paper reports the ladder rather than the final column alone. A firm $\times$ cycle effect absorbs everything about the donor in that cycle—its size, its PAC budget, its politics, its taste for giving. A candidate $\times$ cycle effect absorbs everything about the recipient—seniority, leadership, party, the safety of the seat, whether the race is contested, and membership of Ways and Means. Imposed together they leave the interaction: two firms giving to the same candidate in the same cycle differ in J because their industries differ, and one firm giving to two candidates differs in J because their committees differ. That is variation the grid supplies and no proxy can produce.

Two outcomes are estimated on two samples. On the choice set, Y indicates whether the

firm funded the candidate, and Equation 1 is a linear probability model whose coefficients are read against a base rate of 5.84%. On the donations, Y is log dollars, and the sample loses the 4,470 observations at or below zero that cannot be logged. The first asks whether a reason moves the choice of candidate; the second asks whether it moves the size of the gift conditional on that choice.

What the design does not settle is that committee assignment is not random. A member seeks the committee that matters to the industries in the district and to the donors the member already has, and no fixed effect in Equation 1 removes that. Section 7 bounds how much of the association could run backwards, by holding the firm–candidate pair fixed and using the fact that the assignment is dated before the donation. It bounds the problem; it does not eliminate it.

6.2 Motive Lives on Selection, Not on Size

Table 5 reports the two margins on the same ladder of fixed effects, and the contrast between them is the paper's central quantitative result. On the choice of candidate, jurisdiction over the donor's industry raises the probability of a donation by 3.68 percentage points against a base rate of 5.84%—an increase of roughly two-thirds—and a qualifying presence in the district raises it by 14.28 points, or roughly two and a half times the base rate. Both are estimated with firm × cycle and candidate × cycle effects and both exceed fourteen standard errors.

On the size of the gift the same two reasons are worth 14.6% and 36.5%. Those are not small numbers, but they are an order of magnitude smaller than the selection effects relative

to their bases, and the institutional setting predicted as much: the amount is bounded by statute and sits on a round number three times in five, while the choice of whom to fund is unconstrained. A firm expresses its reasons by picking a candidate, and having picked, gives what it gives.

The movement across columns is itself a result and is the reason the table reports the ladder. Pooled with only cycle effects, jurisdiction is associated with 3.2% more money, significant at the ten percent level and economically slight. Adding firm $\times$ cycle effects leaves 2.4%, indistinguishable from zero. Adding candidate $\times$ cycle effects raises the coefficient to 14.6% and its t-statistic above ten. Raw means say the same thing about the pooled comparison: donations to candidates with jurisdiction average \$3,194 against \$3,077 for those without, a gap of 3.8%, which is the pooled estimate and not the within-candidate one.

The reason the pooled comparison hides the effect is that candidates whose committees reach many industries are systematically different candidates. Broad reach concentrates on committees like Energy and Commerce, whose members are senior, safe and heavily funded from every direction; the pooled coefficient mixes the jurisdiction premium with whatever else makes such a member attractive. Only after the candidate is held fixed, and the comparison runs across donors from different industries funding the *same* person, does jurisdiction separate from candidate quality. That the sign of the finding depends on this is a caution for any study using a committee proxy without the cross-classified structure to condition on the recipient.

6.3 Jurisdiction Over an Industry Against Jurisdiction Over Everyone

The measure's central distinction predicts that the second committee code should carry no weight once the first is controlled, and Table 6 tests it. A committee reaching the donor's industry only as it reaches every industry is associated with a -0.22 percentage point change in the probability of a donation (standard error 0.13), a -1.2% change in the amount (0.87), and a -0.2% change in dollars under Poisson estimation over the whole choice set (2.3). Sectoral jurisdiction in the same specifications is worth $+3.60$ points, $+14.1\%$ and a factor of 1.88.

The contrast is not a mechanical consequence of one variable being rarer. Economy-wide reach fires on 23.5% of donations against 53.6% for sectoral jurisdiction, so the estimates are comparably precise; the economy-wide coefficient is not imprecise, it is close to zero and negative. Nor is it an artefact of comparing committee-holders with non-holders: column 4 restricts to candidates holding at least one seat, so the contrast becomes a seat with jurisdiction against a seat without, and the coefficients barely move.

This is the strongest available evidence that the grid is measuring jurisdiction rather than relabelling committee membership. Both codes describe committee seats; both are assigned by the same model under the same standard in the same isolated judgements; they differ only in whether the statute the model named is written about the donor's industry or about all industries. Firms pay for one and not the other. A committee dummy of the kind used in the literature pools the two and returns their average, which is how a proxy overstating incidence by 1.70 can nonetheless appear to work.

Compounding the two margins, jurisdiction nearly doubles the dollars a firm sends a candidate and district presence more than triples them. Column 6 estimates this by Poisson pseudo-maximum-likelihood on dollars over the whole choice set, which admits the zeros and so combines the decision to give with the decision of how much, and puts the two factors at 1.88 and 3.30. These are the magnitudes to quote for total dollars, and they sit between the two margins as they should.

6.4 More Jurisdiction, More Money

The relation is monotone in the number of committees and concave in it (Figure 4). One committee with jurisdiction over the donor's industry raises the probability of a donation by 3.4 percentage points, two by 4.7, three or more by 6.1; in log dollars the same steps are 13.8%, 17.8% and 20.8%. The first committee does most of the work, which is what an access story implies—a firm needs a hearing somewhere in the process, and a second seat adds less than the first.

A dose response is a weak test on its own and a useful one in combination. It would arise mechanically if the count were proxying for how many committees a candidate holds in total, but the specification also carries the count of committees reaching the donor's industry only economy-wide, and that count is worth 0.29 points against sectoral jurisdiction's 2.58. The count of committees reaching the industry not at all cannot be identified separately: within a candidate $\times$ cycle cell the total number of seats is fixed, so the three counts carry two degrees of freedom.

7 When Jurisdiction Moves

7.1 Money Arrives with Jurisdiction and Does Not Leave with It

Holding the firm–candidate pair fixed and letting the candidate's committee assignment change underneath it gives the sharpest result in the paper. Table 7 reports first differences between consecutive cycles for pairs observed in both. When a candidate gains jurisdiction over a donor's industry—held no committee reaching it last cycle, holds one now—the probability that the firm gives rises by 4.59 percentage points and the amount by 21.9%. When the same candidate loses jurisdiction but remains on committees, the probability changes by 0.24 points (standard error 0.22) and the amount by 1.0% (1.7). Neither is distinguishable from zero.

The asymmetry is not an artefact of pooling two kinds of loss. A member can shed jurisdiction by being reassigned within the House or by leaving it, and only the first holds the relationship fixed. Separating them shows why the pooled loss coefficient is misleading: leaving all committees is associated with a 8.95 point fall in the probability of a donation and a 62.2% fall in the amount, while reassignment within the House is associated with nothing. Money leaves when the member leaves, not when the jurisdiction does.

Two facts make the comparison readable in this direction rather than the other. The committee assignment is dated to the Congress sitting *before* the donation, so the money moves after the assignment rather than before it. And a firm's own industry almost never moves—GICS6 changes in 146 of 7,410 firm–cycles—so a within-pair switch is the candidate's committee change and not a reclassification of the donor. This bounds how much of the cross-sectional association in Section 6 could run backwards from money to assignment. It

does not make the assignment random, and the paper does not claim it does.

Figure 5 traces the gain cycle by cycle. The probability of giving is flat before the change—−0.5 and −0.4 percentage points at three and two cycles prior, neither significant—jumps 4.8 points in the cycle the assignment takes effect, and stays there: 4.9, 4.9 and 5.0 points at one, two and three cycles after. On a base rate of 4.6% in the pre-period this is a doubling, and it is a level shift rather than a drift. The right panel shows the loss, restricted to members who stay on committees, and it is flat throughout the post-period. Its rising pre-period is not a placebo failure: to be at risk of losing jurisdiction a pair must first have gained it, so the run-up into the omitted cycle is the tail of the left panel.

Read together, the two panels say that corporate donation relationships are formed on jurisdiction and maintained on something else. That interpretation also supplies a candidate account of the money the ladder assigns to firm-invariant reasons and to Others: some of it will be relationships whose original jurisdictional reason has since moved on, and which the measure, looking only at the current Congress, cannot see. The data here cannot separate that from a taste for maintaining contacts, and the paper does not try.

7.2 The Composition Has Shifted, and the Two Motives Have Moved Apart

Composition is stable in levels and the pull of the reasons is not, which is why Figure 6 and Figure 7 answer different questions. Table 8 and Figure 6 report the shares. Industry jurisdiction moves from 48.7% of donations in 2000 to 55.4% in 2022, tax and spending from 25.1% to 27.1%, geography from 9.3% to 13.4%, and Others falls from 15.6% to 9.0%. The

mean gift roughly doubles in nominal terms, which the unindexed FECA cap makes a fall in real terms.

Estimating Equation 1 cycle by cycle tells a different story about the same period. Because the unconditional probability of funding any given candidate rises from 5.4% to 6.9% over the window, a constant percentage-point effect would be a shrinking one, so Figure 7 estimates by Poisson and reports a ratio of probabilities. Jurisdiction rises from 1.5 to 1.7 times; district presence falls from 3.4 to 2.6. Corporate PAC giving has become more responsive to what a committee governs and less responsive to where a firm employs people.

The lines do not cross, and the finding is the two trends rather than a reversal. Geography remains the larger effect in every cycle: a firm is still more likely to fund a candidate because it employs people in the district than because a committee reaches its industry. What has changed is the gap between them, and the direction of the change is what a professionalising PAC operation would produce—one that reads committee rosters rather than a map of its own plants.

7.3 Where the Two Motives Bite Hardest

The average effect pools firms whose PAC programmes differ by a factor of thirty in size, and Table 9 separates them. Because base rates range from 0.47% in the smallest quartile of PAC programmes to 17.0% in the largest, the comparable quantity across rows is the coefficient divided by the base, and the table reports both.

The geographic motive belongs to the small PAC. In the smallest quartile—firms funding a median of three candidates—a qualifying presence in the district raises the probability of a

donation by more than eighteen times the base rate, while jurisdiction raises it by 0.74 times. In the largest quartile, funding a median of 88 candidates, geography is worth 0.87 times the base and jurisdiction 0.59. Both reasons matter everywhere and geography's coefficient is never the smaller one; what collapses with scale is its dominance, from roughly twenty-five times jurisdiction's pull in the smallest quartile to one and a half times in the largest—the convergence a fixed cost of learning committee rosters implies.

The candidate's status splits the sample differently. Jurisdiction is worth 0.45 times the base rate for incumbents and is indistinguishable from zero for open-seat candidates, which is mechanical: a candidate who holds no seat has no jurisdiction. The challenger row is worth reading carefully. Its jurisdiction coefficient is 2.97 times the base rate and highly significant, and it rests on 653 donation rows involving eighteen candidates coded as challengers who nonetheless appear with committee assignments—former members seeking to return. The estimate is real and its subject is narrow.

7.4 Who Receives the Money in the Others Category

The residue is a candidate-level fact once aggregated, and it is systematic rather than noise. Table 10 collapses the sample to the candidate–cycle and asks what predicts corporate money and its composition. A committee portfolio reaching the whole corporate economy is worth 1.98 log points of corporate receipts and 1.21 log points of donor count; a seat on Ways and Means or Appropriations is worth 1.74 and 1.00. A committee portfolio that reaches broadly and a seat on the money committees command comparable sums.

Columns 3 to 5 are partly definitional and should be read as magnitudes rather than as

tests. Reach, the money committees and a leadership post are three of the five rungs of the ladder that defines the dependent variable, so each must push the unexplained share down; what the columns add is how far. Reach and the money committees do comparable work, at -0.42 and -0.35 on a share, and a leadership post about a third as much at -0.12. The restriction to incumbents in column 5, where the question has content, leaves the pattern intact.

Column 6 carries no such circularity and returns the least expected result in the table. Holding the number of donors fixed, a candidate whose committees reach more of the corporate economy draws a donor base that is *narrower* by sector, not wider: the coefficient is -0.48 sectors on a mean of 6.7. Broad jurisdiction does not assemble a broad coalition of donors. It assembles a large one drawn from the industries the portfolio actually governs, which is the same selection Figure 2 shows committee by committee, seen from the candidate's side.

8 Conclusion

A corporate PAC donation is among the most completely documented transactions in American public life—donor, recipient, amount and date, all public since the 1970s—and for as long as the records have existed, the one thing they omit is the point of the transaction. This paper measures it, by tabulating a fact that was always public and never tabulated: which congressional committees govern which industries. The answer the measure returns is that firms fund the politicians who govern them. Not powerful politicians as such, not local ones as such, and not committee members as such—members whose committees hold jurisdiction over the donor's own business.

Three findings carry that conclusion. First, the composition: jurisdiction over the donor's industry accounts for half of corporate PAC giving, and adding the district where the firm operates, the tax and spending committees, and party leadership leaves only a tenth of the money with no identifiable reason—most of it backing candidates who hold no committee seat yet, which is to say bets on future members rather than unexplained ties to current ones. Second, the separation the measure was built for: a committee seat that reaches the donor's industry through a statute written about it raises the probability of a donation by nearly two-thirds of the base rate, while a seat that reaches the donor the way it reaches every firm is worth nothing, so a committee dummy that pools them—as the standard proxy does—measures a motive 1.7 times the size of the one that exists. Third, the dynamics: when a candidate gains jurisdiction over a firm's industry the money arrives in that cycle, 4.6 percentage points of donation probability that were absent in the two cycles before; when the jurisdiction later departs and the member stays, the money stays too. Donation relationships are opened on jurisdiction and maintained on something else.

The findings change what can be asked of the returns literature. An estimated return to political spending has until now averaged over donations made to a regulator, to a neighbour, and to a prospect, as if the same asset were being priced; with a motive attached to each donation, returns can be estimated within motive, where they have a meaning. That is the natural second half of this research programme: money given for jurisdiction should pay, if it pays, in regulatory and legislative outcomes for the industry, and money given for geography in decisions that touch the district—predictions the measure now makes testable donation by donation.

The method travels beyond this setting, and it is worth stating what exactly should travel.

The scarce input was not the language model but the design: locating, inside a question about half a million transactions, a small closed mapping—1,444 cells, each answerable from public statutes—that could be judged once, in isolation, published whole, and inherited by every observation mechanically. Documented-but-untabulated mappings of the same shape are common: which regulatory agencies oversee which industries, which sections of the tax code are written about which activities, which committees of state legislatures govern which markets. Wherever one exists, the properties that matter here—determinism, immunity to look-ahead, audit at the level of the claim—come from the design rather than from trust in the model, and they are available at the cost of building the table.

Two extensions are specified and waiting. The grid's external validation runs through lobbying disclosures, in which firms name the bills they work and therefore the committees they believe govern them; the test is whether the grid predicts that contact net of industry and committee fixed effects, and lobbying data has been kept out of the grid's construction so the test stays meaningful. And the social-tie channel—the one motive that resists the grid design because it lives in pairs of biographies rather than in statutes—fires on 3.2 percent of a trial subsample and awaits its full-sample pass. Both belong to the same programme the measure opens: having asked what the money was for, ask next what it bought.

References

Aggarwal, R. K., Meschke, F., and Wang, T. Y. (2012). Corporate Political Donations: Investment or Agency? *Business and Politics*, 14(1):1–38.

Ansolabehere, S., de Figueiredo, J. M., and Snyder Jr, J. M. (2003). Why is There so Little Money in U.S. Politics? *Journal of Economic Perspectives*, 17(1):105–130.

Bertomeu, J., Cheynel, E., Liao, Y., and Milone, M. (2025). Using Machine Learning to Measure Conservatism. *Management Science*, 71(2):1504–1522.

Bertrand, M., Bombardini, M., and Trebbi, F. (2014). Is It Whom You Know or What You Know? An Empirical Assessment of the Lobbying Process. *American Economic Review*, 104(12):3885–3920.

Bombardini, M. and Trebbi, F. (2020). Empirical Models of Lobbying. *Annual Review of Economics*, 12(Volume 12, 2020):391–413.

Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science*, 58(2):367–387.

Bonica, A. (2024). Database on ideology, money in politics, and elections: Public version 4.0. [Computer file]. Stanford, CA: Stanford University Libraries. http://data.stanford.edu/dime.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative AI at Work. *The Quarterly Journal of Economics*, page qjae044.

Cooper, M. J., Gulen, H., and Ovtchinnikov, A. V. (2010). Corporate Political Contributions and Stock Returns. *The Journal of Finance*, 65(2):687–724.

Erel, I., Stern, L. H., Tan, C., and Weisbach, M. S. (2021). Selecting Directors Using Machine Learning. *The Review of Financial Studies*, 34(7):3226–3264.

Fowler, A., Garro, H., and Spenkuch, J. L. (2020). Quid Pro Quo? Corporate Returns to Campaign Contributions. *The Journal of Politics*, 82(3):844–858.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5):2223–2273.

Kang, K. (2016). Policy Influence and Private Returns from Lobbying in the Energy Sector. *The Review of Economic Studies*, 83(1):269–305.

Kellermann, M. and Shepsle, K. A. (2009). Congressional Careers, Committee Assignments, and Seniority Randomization in the U.S. House of Representatives. *Quarterly journal of political science*, 4(2):87–101.

Kim, I. S., Stuckatz, J., and Wolters, L. (2025). Systemic and Sequential Links between Campaign Donations and Lobbying. *The Journal of Politics*.

Kroszner, R. S. and Stratmann, T. (1998). Interest-Group Competition and the Organization of Congress: Theory and Evidence from Financial Services' Political Action Committees. *The American Economic Review*, 88(5):1163–1187.

Mullainathan, S. (2025). Economics in the Age of Algorithms. *AEA Papers and Proceedings*, 115:1–23.

Nikolov, B., Schürhoff, N., and Wagner, S. (2025). AI in corporate governance: Can machines recover corporate purpose? Swiss Finance Institute Research Paper No. 25-23. https://ssrn.com/abstract=5166191.

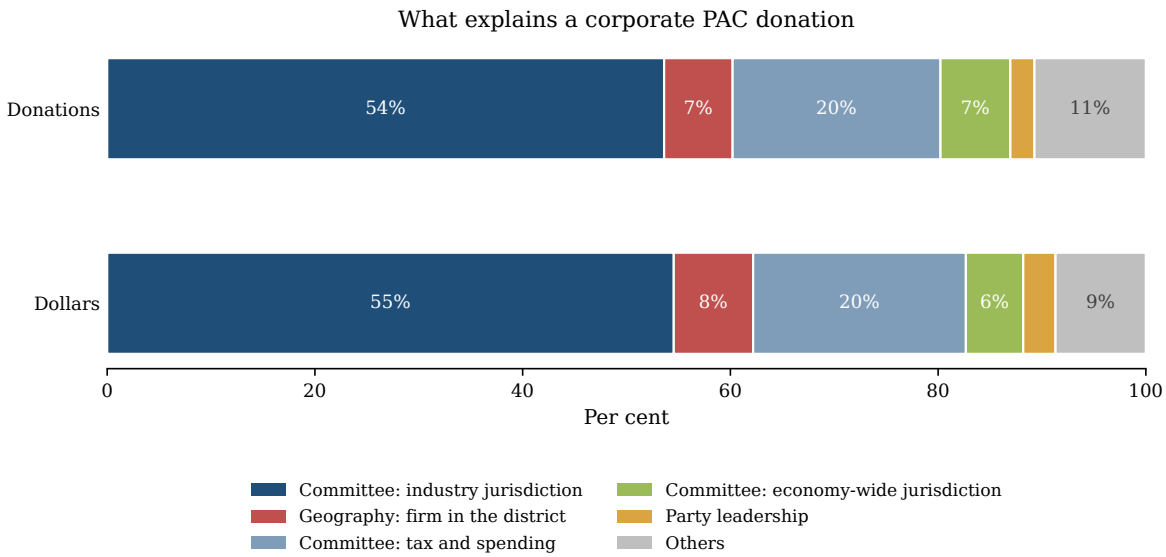
Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Powell, E. N. and Grimmer, J. (2016). Money in Exile: Campaign Contributions and Committee Access. *The Journal of Politics*, 78(4):974–988.

Stewart, C. I. and Woon, J. (2024). Congressional committee assignments, 103rd to 117th congresses, 1993–2022: House. Harvard Dataverse. https://doi.org/10.7910/DVN/XLIHUC.

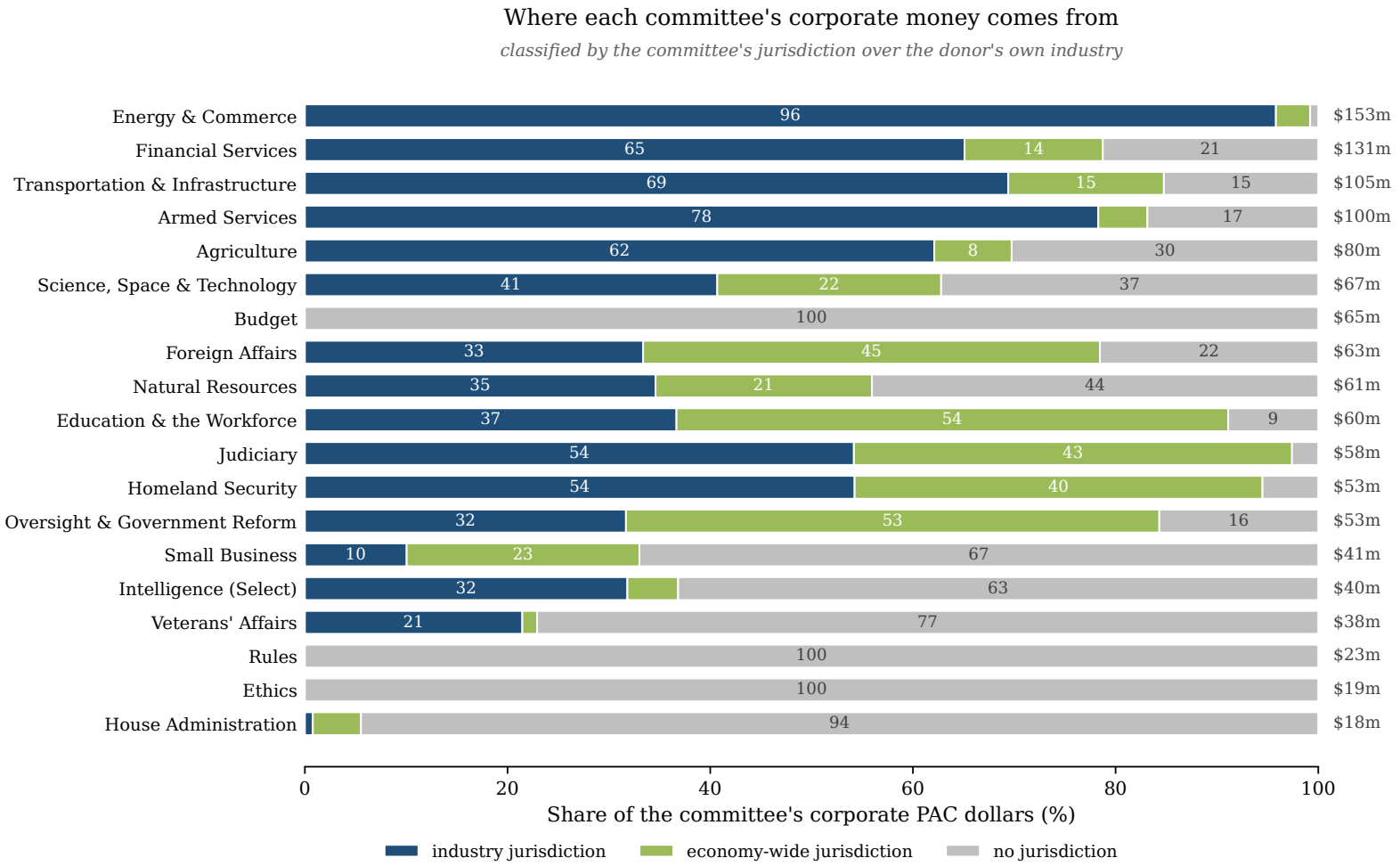
Tripathi, M., Ansolabehere, S., and Jr, J. M. S. (2002). Are PAC Contributions and Lobbying Linked? New Evidence from the 1995 Lobby Disclosure Act. *Business and Politics*, 4(2):131–155.

Figure 1: What explains a corporate PAC donation



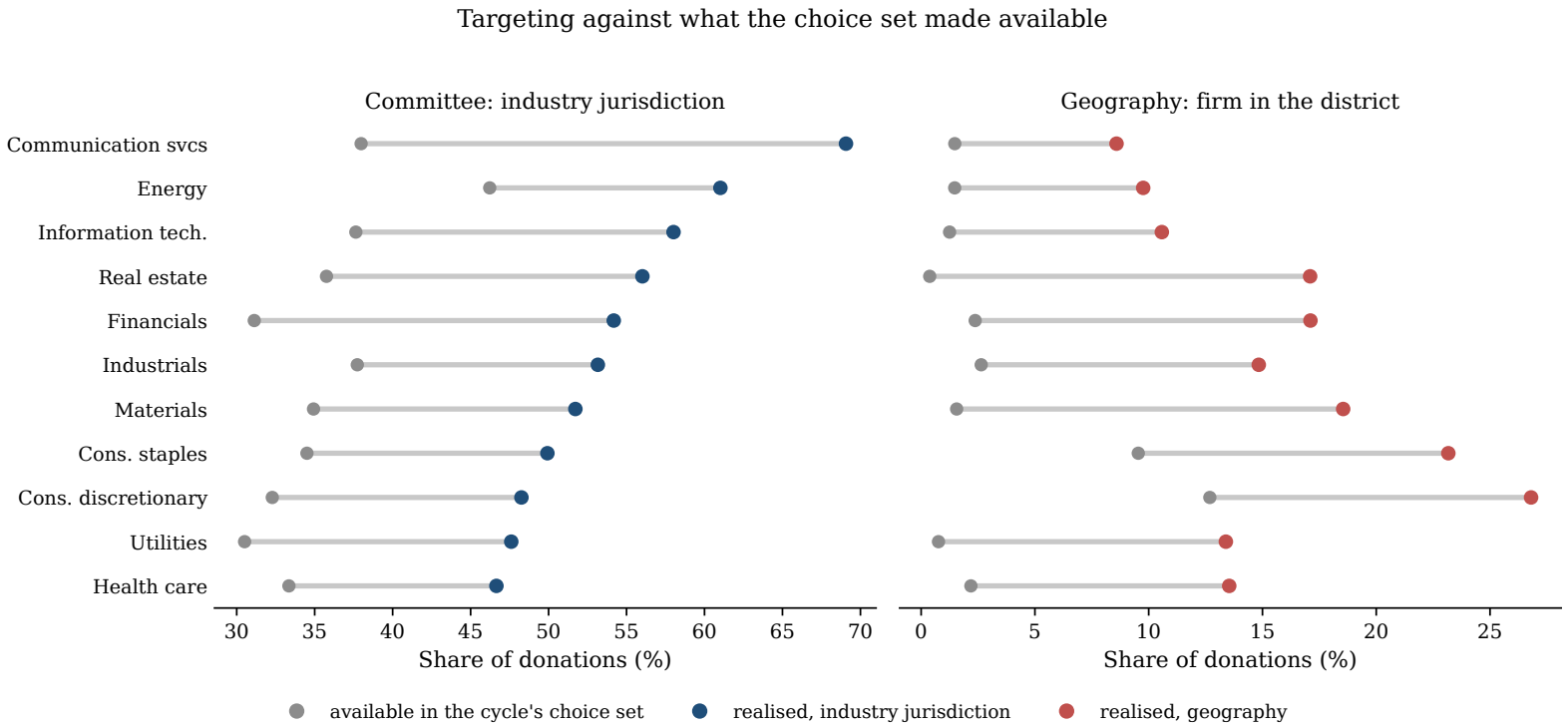
Each donation is assigned to the most firm-specific reason that applies to it, so the last band is what survives every reason rather than what survives one of them. The two bars weight the same 285,072 donations by count and by dollars. Table 2 reports the underlying counts and defines each category.

Figure 2: Where each committee's corporate money comes from



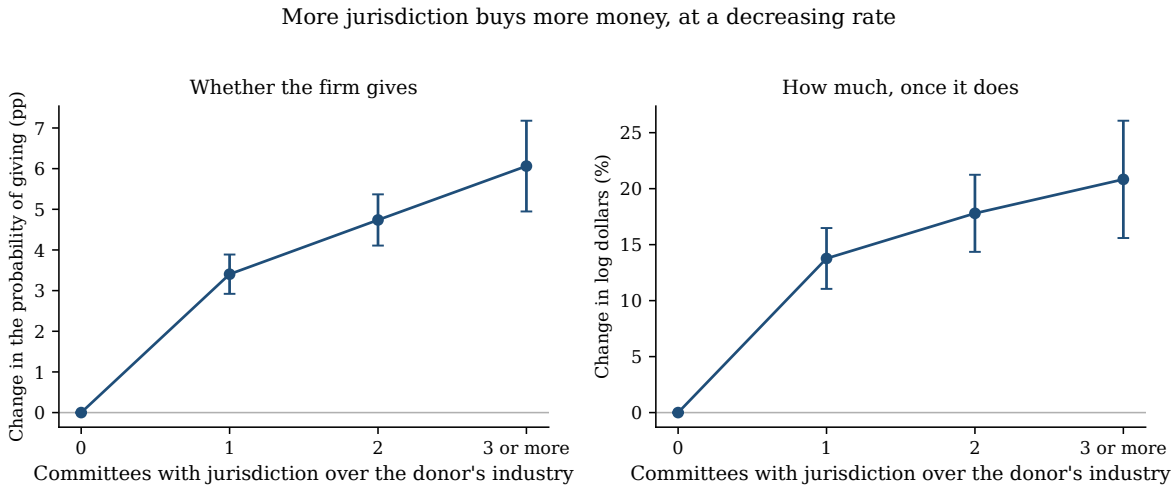
For each of the nineteen graded committees, the share of the corporate PAC dollars received by its members that comes from industries the committee reaches through a sectoral instrument, from industries it reaches only as it reaches every industry, and from industries it does not reach. Dollar totals are at the right. The grid never observes a donation, so the alignment is a consistency check on the grid rather than a construction of it.

Figure 3: Targeting against what the choice set made available



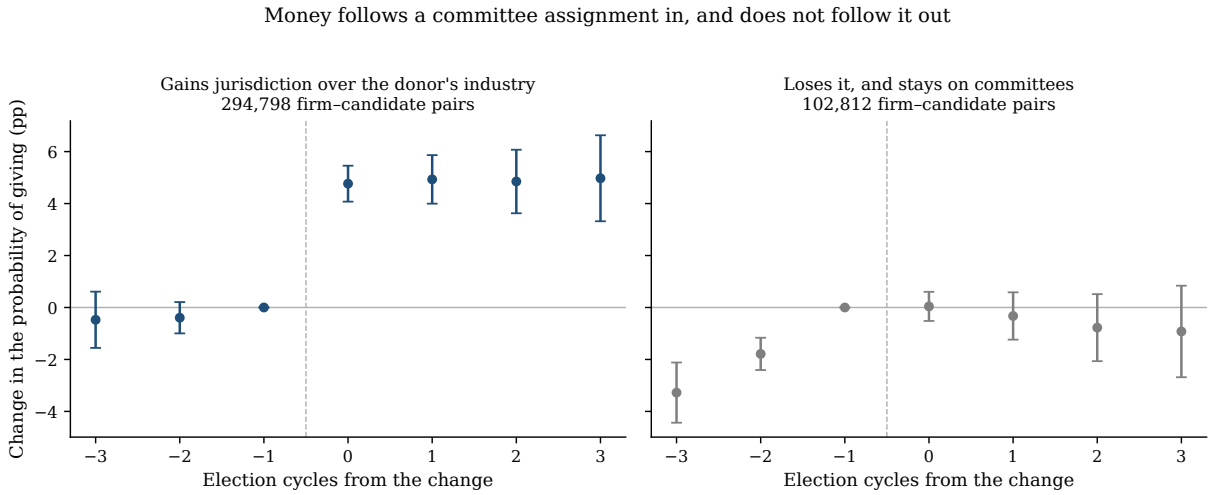
Grey marks the share of the cycle's own choice set that carried the reason for the average firm in the sector; the coloured mark is the share of that firm's realised donations that did. The gap is targeting above what giving at random within the cycle would produce.

Figure 4: More jurisdiction, more money



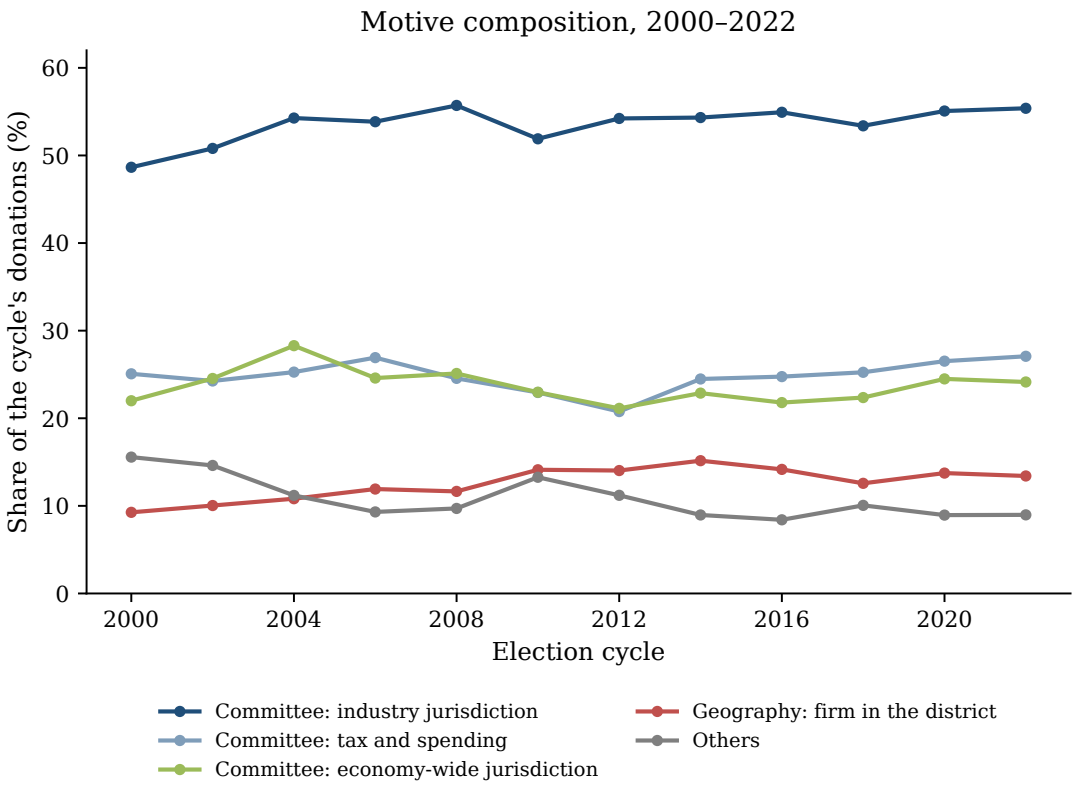
Coefficients on indicators for the number of the candidate's sitting committees that reach the donor's own industry, with zero omitted, from the specifications of Table 5 column 3 (left) and column 6 (right). Bars are 95% confidence intervals from standard errors clustered on firm and on candidate.

Figure 5: The probability of giving around a change in jurisdiction



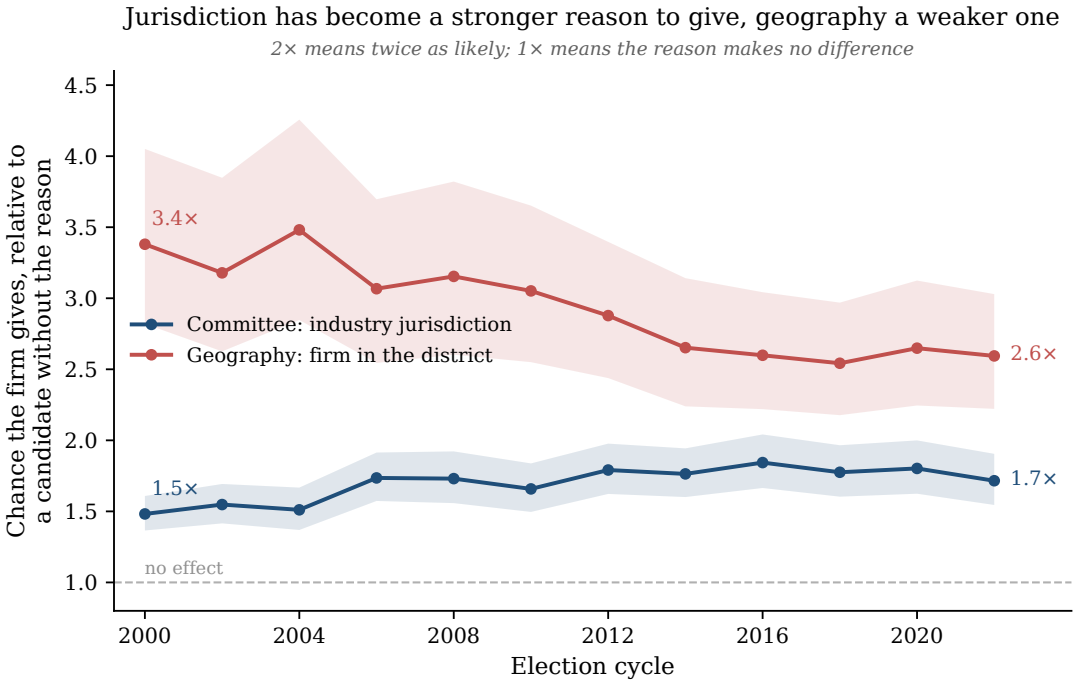
Cycles are indexed relative to the first cycle in which the firm–candidate pair's jurisdiction status changes; the cycle before the change is omitted. The right panel keeps only members who remain on committees, which is the only version of the loss that holds the relationship fixed. Its rising pre-period is the tail of the left panel: to be at risk of losing jurisdiction a pair must first have gained it.

Figure 6: Motive composition, 2000–2022



Share of each cycle's donations on which each reason fires. The reasons are not mutually exclusive and do not sum to one hundred; Others is the share on which none of the five fires. Table 8 reports the same series with dollar totals.

Figure 7: The two motives over time



Poisson estimates of the specification in Table 5 column 3, run cycle by cycle, so that each point is a ratio of probabilities rather than a difference in them. Shaded bands are 95% confidence intervals from standard errors clustered on firm and on candidate. The lines do not cross: geography remains the larger effect throughout.

Table 1: Summary statistics

Panels A–C are the firm–candidate–cycle donation sample: every donation from a Compustat firm's corporate PAC to a U.S. House candidate, 2000–2022 cycles. Panel D collapses that sample to the firm–cycle and to the candidate–cycle, so a firm that funded no candidate in a cycle and a candidate who received nothing are absent by construction. Donations at or below zero are refunds and net-outs and are retained rather than dropped. Committee variables use the committees of the Congress sitting when the donation was made, not the one seated after the election. Gives to this candidate next cycle is missing for the 2022 cycle, which has no successor in the window. The shared-institution tie is the only row not measured on the full sample: it is judged per firm–candidate pair over executive and candidate biographies rather than looked up from a grid, and is available only on a 1,000-donation random subsample, where an executive or director of the firm and the candidate attended the same institution within five years of one another on 3.2% of donations. Its dispersion and quantiles are therefore left blank. Dollars are nominal.

	Mean	SD	P25	Median	P75	N
Panel A. Donation						
Donation ($)	3,140	2,781	1,000	2,000	4,500	285,072
log(donation)	7.603	1.274	6.908	7.601	8.412	285,072
Quarters in which the firm gave	1.838	1.131	1.000	1.000	2.000	285,072
First donation quarter of the cycle	3.979	2.333	2.000	4.000	6.000	285,072
Consecutive cycle with this candidate	2.407	1.836	1.000	2.000	3.000	285,072
Gave to this candidate last cycle	0.503	0.500	0.000	1.000	1.000	285,072
Gives to this candidate next cycle	0.524	0.499	0.000	1.000	1.000	263,627
Net refund (donation < 0)	0.006	0.074	0.000	0.000	0.000	285,072
Panel B. Motive						
Committee with jurisdiction over the firm's industry	0.536	0.499	0.000	1.000	1.000	285,072
number of such committees	0.690	0.741	0.000	1.000	1.000	285,072
Committee with jurisdiction common to every industry	0.235	0.424	0.000	0.000	0.000	285,072
number of such committees	0.275	0.532	0.000	0.000	0.000	285,072
committees not reaching the industry	0.452	0.667	0.000	0.000	1.000	285,072
Committee on tax or spending (Ways and Means, Appropriations)	0.247	0.431	0.000	0.000	0.000	285,072
Geography: firm HQ or a 250+ job plant in the district	0.128	0.334	0.000	0.000	0.000	285,072
headquarters in the district	0.022	0.146	0.000	0.000	0.000	285,072
establishment, 250+ jobs, in the district	0.118	0.323	0.000	0.000	0.000	285,072
Shared-institution tie, trial subsample	0.032					1,000
Others (none of the above)	0.107	0.309	0.000	0.000	0.000	285,072
Panel C. Candidate						
Incumbent	0.919	0.273	1.000	1.000	1.000	285,072
Open seat	0.052	0.221	0.000	0.000	0.000	285,072
Challenger	0.029	0.169	0.000	0.000	0.000	285,072
Republican	0.599	0.490	0.000	1.000	1.000	285,072
Sits on any committee	0.909	0.287	1.000	1.000	1.000	285,072
number of committees	1.777	0.990	1.000	2.000	2.000	285,072
Party leadership post	0.142	0.349	0.000	0.000	0.000	285,072
Panel D. Firm-cycle and candidate-cycle						
Candidates funded per firm-cycle	38.5	56.0	6.0	17.0	47.0	7,410
Dollars given per firm-cycle ($000)	120.8	252.6	10.5	36.4	115.5	7,410
Corporate donors per candidate-cycle	36.1	36.1	5.0	28.0	54.0	7,898
Dollars received per candidate-cycle ($000)	113.3	148.2	8.3	66.6	157.5	7,898

Table 2: The sample, and what explains a corporate PAC donation

Panel A counts the sample: every donation from a Compustat firm's corporate PAC to a U.S. House candidate over the 2000–2022 election cycles, one row per firm–candidate–cycle. Panel B assigns each donation to the most firm-specific reason that applies to it. The five reasons are not mutually exclusive in the data and several fire on the same donation; the ordering here is a reporting convention, and it is what lets the last row be read as what survives every reason rather than what survives one of them. The three committee rows differ in whose business the committee's remit is written about. *Industry jurisdiction* means a committee reaches the donor's own GICS6 industry through a sectoral instrument – a statute, programme or regulatory regime whose scope is defined by what that industry makes, sells or operates – and every credited cell of the committee × industry grid names the instrument. *Economy-wide jurisdiction* means the committee reaches the industry only through a legal function that applies to firms in every sector: antitrust, bankruptcy, ordinary employment law. *Tax and spending* is Ways and Means or Appropriations, whose remits – taxation and tariffs, and the discretionary budget – are identical for every donor and so are supplied mechanically rather than judged. Geography means the firm's headquarters or an establishment with 250 or more employees sits in the district being contested, measured against the congressional map on the ballot. *Others* is the residue: none of the five reasons fires. It is not a claim that no reason exists, only that none of the reasons measured here does; 71% of it goes to candidates holding no committee seat at all. Social ties are absent from the ladder – they are judged per firm–candidate pair rather than looked up, and are available only on a 1,000-donation trial subsample, where they fire on 3.2% of donations. Dollars are nominal.

	Donations	% of donations	$m	% of dollars
Panel A. Sample				
Firm–candidate–cycle donations	285,072		895	
Firms (gvkey)	1,322			
Candidates	2,869			
Election cycles	12			
Firm–cycles	7,410			
Candidate–cycles	7,898			
Panel B. The explanation ladder				
Committee with jurisdiction over the firm's industry	152,852	53.6	488	54.5
Geography: firm HQ or a 250+ job plant in the district	18,803	6.6	68	7.6
Committee on tax or spending (Ways and Means, Appropriations)	57,053	20.0	183	20.5
Committee with jurisdiction common to every industry	19,182	6.7	49	5.5
Party leadership post	6,611	2.3	28	3.1
Others (none of the above)	30,571	10.7	78	8.7

Table 3: The standard proxies against the measures used here

Each row is the share of the 285,072 donations on which the indicator fires; the last column divides that share by the paper's own measure, set in bold. In Panel A the standard proxy – an indicator for sitting on a committee – overstates the committee motive by 1.70 times, and the error runs one way: jurisdiction implies a seat, so every disagreement is a false positive. Of the donations to a candidate holding any committee seat, 59% reach one with jurisdiction over the donor's own industry. In Panel B the two levels are close and the classifications are not. A home-state indicator and an operations-in-the-district measure agree on only 80% of donations, because 71% of the donations with a qualifying presence in the district come from establishments outside the firm's headquarters state, and 74% of same-state donations carry no qualifying presence in the contested district at all. A home- district indicator built from the headquarters ZIP alone recovers 0.18 of the measure.

	% of donations	Donations	Ratio
Panel A. Committee			
Sits on any committee	90.9	259,235	1.70
Sits on one of the 19 graded standing committees	72.8	207,619	1.36
Committee on tax or spending (Ways and Means, Appropriations)	24.7	70,413	0.46
Committee with jurisdiction over the firm's industry	53.6	152,852	1.00
Committee with jurisdiction common to every industry	23.5	66,959	0.44
Panel B. Geography			
Headquarters state equals the candidate's state	14.5	41,428	1.14
Geography: firm HQ or a 250+ job plant in the district	12.8	36,459	1.00
Headquarters district equals the contested district	2.3	6,569	0.18
Headquarters in the district, from the establishment panel	2.2	6,198	0.17

Table 4: Motive composition by GICS sector

Sectors are the eleven GICS level-2 sectors, assigned from the firm's historical GICS code in the donation year and ordered by the share of the sector's donations reaching a committee with jurisdiction over the donor's own industry. Column headings are as in Table 8. Industry jurisdiction is judged at GICS6 rather than at the sector level shown here, so two firms in the same sector can face different committees. Real estate holds 222 donations from a handful of firms and is reported for completeness. Dollars are nominal.

	Firms	Donations	$m	Mean $	Industry	Tax/spending	Geography	Others
Energy	94	18,698	57	3,060	62.5	20.8	9.1	13.4
Communication svcs	56	16,693	60	3,617	59.0	18.8	5.7	14.3
Information tech.	113	16,947	45	2,664	57.0	25.9	5.8	8.5
Industrials	224	61,364	232	3,788	56.4	26.5	10.9	8.9
Real estate	4	222	0	1,466	55.4	17.1	4.5	15.8
Financials	255	43,226	131	3,036	54.8	22.5	9.6	10.2
Utilities	98	25,410	72	2,833	50.6	24.0	7.2	12.1
Health care	182	36,240	111	3,061	49.9	29.5	8.4	11.4
Materials	102	12,970	33	2,565	49.8	25.9	14.7	11.3
Cons. discretionary	175	33,557	99	2,936	48.6	23.3	30.4	8.9
Cons. staples	63	19,745	54	2,718	48.3	25.5	25.4	13.3

Table 5: Which candidate a firm funds, and how much

Columns 1–3 are a linear probability model on the choice set: for every firm-cycle, every U.S. House candidate who received corporate PAC money from any sample firm in that cycle, whether or not this firm gave to them. Columns 4–6 are the donations themselves, dropping the 4,470 rows at or below zero that cannot be logged. Committee jurisdiction is an indicator that at least one of the candidate's sitting committees reaches the donor's own GICS6 industry through a sectoral instrument, looked up from a committee × industry grid; geography is an indicator that the firm's headquarters or an establishment with 250 or more employees sits in the district being contested. Both use the Congress sitting when the donation was made and the congressional map on the ballot. With firm × cycle and candidate × cycle fixed effects together, the estimate is identified only by the match between the donor's industry and the candidate's committees: two firms giving to the same candidate in the same cycle differ in jurisdiction because their industries differ. Standard errors are clustered two ways, on firm and on candidate; * p<0.10, ** p<0.05, *** p<0.01.

	Whether the firm gives			How much, once it does		
	(1)	(2)	(3)	(4)	(5)	(6)
	Gave	Gave	Gave	log($)	log($)	log($)
Committee with jurisdiction over the firm's industry	0.0491***	0.0488***	0.0368***	0.0317*	0.0239	0.1455***
	(0.0035)	(0.0034)	(0.0025)	(0.0165)	(0.0188)	(0.0140)
Geography: firm HQ or a 250+ job plant in the district	0.1448***	0.1437***	0.1428***	0.1703***	0.2892***	0.3645***
	(0.0148)	(0.0088)	(0.0087)	(0.0311)	(0.0169)	(0.0171)
Mean of dependent variable	0.058	0.058	0.058	7.724	7.724	7.724
Cycle FE	Yes	–	–	Yes	–	–
Firm × cycle FE	No	Yes	Yes	No	Yes	Yes
Candidate × cycle FE	No	No	Yes	No	No	Yes
Observations	4,880,936	4,880,936	4,880,936	280,602	280,143	279,120
R^2 (within)	0.024	0.020	0.013	0.005	0.013	0.025

Table 6: Industry jurisdiction against jurisdiction common to every industry

Columns 1–4 are a linear probability model on the choice set; column 5 is log dollars among the donations that occurred; column 6 is Poisson pseudo-maximum-likelihood on dollars over the whole choice set, which combines the two margins and admits the zeros, with net refunds set to zero. Column 4 keeps only candidates holding at least one committee seat, so the contrast is between a seat with jurisdiction and a seat without rather than between a seat and none. Economy-wide jurisdiction is a committee whose remit touches the donor's industry only through a legal function that applies to firms in every sector – antitrust, bankruptcy, ordinary employment law – as against a sectoral instrument written about what the industry makes, sells or operates. The two are recorded separately in the grid and can both be true of the same candidate. Column 3 also enters the number of the candidate's committees reaching the donor's industry not at all; it is absorbed and drops out, because the candidate's total committee count is fixed within a candidate × cycle cell, so the three counts carry two degrees of freedom rather than three. All columns carry firm × cycle and candidate × cycle fixed effects. Standard errors are clustered two ways, on firm and on candidate; * p<0.10, ** p<0.05, *** p<0.01.

	Choice set				Donations	Choice set
	(1)	(2)	(3)	(4)	(5)	(6)
	Gave	Gave	Gave	Gave	log($)	$ (PPML)
Committee with jurisdiction over the firm's industry	0.0369***	0.0360***		0.0361***	0.1406***	0.6321***
	(0.0025)	(0.0025)		(0.0028)	(0.0144)	(0.0490)
Committee with jurisdiction common to every industry		-0.0022*		-0.0023	-0.0118	-0.0023
		(0.0013)		(0.0014)	(0.0087)	(0.0228)
number with industry jurisdiction			0.0258***			
			(0.0018)			
number with economy-wide jurisdiction			0.0029***			
			(0.0011)			
Geography: firm HQ or a 250+ job plant in the district		0.1428***	0.1428***	0.1927***	0.3645***	1.1939***
		(0.0087)	(0.0087)	(0.0116)	(0.0171)	(0.0929)
Mean of dependent variable	0.058	0.058	0.058	0.083	7.724	184
Firm × cycle FE	Yes	Yes	Yes	Yes	Yes	Yes
Candidate × cycle FE	Yes	Yes	Yes	Yes	Yes	Yes
Candidates holding a committee only	No	No	No	Yes	No	No
Observations	4,880,936	4,880,936	4,880,936	3,130,749	279,120	4,707,539
R^2 (within)	0.003	0.013	0.013	0.017	0.025	–

Table 7: Money arrives with jurisdiction and does not leave with it

Columns 1 and 4 hold the firm–candidate pair fixed and estimate in levels; columns 2, 3, 5 and 6 are first differences between consecutive cycles for pairs observed in both, so the pair is differenced out and the coefficients read as the change in giving when the candidate's jurisdiction over the donor's industry changes. Gains jurisdiction means the candidate held no committee reaching the donor's industry in the previous cycle and holds one now; loses it is the reverse. The loss is split by whether the member still holds any committee seat afterwards: reassignment within the House, against leaving the House or losing every seat. Columns 1–3 are on the choice set and the dependent variable is an indicator for giving; columns 4–6 are among donations that occurred and the dependent variable is log dollars, so a pair that stops giving altogether leaves the sample – which is why the exit coefficient is much larger on the extensive margin. Standard errors are clustered two ways, on firm and on candidate; * p<0.10, ** p<0.05, *** p<0.01.

	Choice set			Donations		
	(1) Gave	(2) Δ Gave	(3) Δ Gave	(4) log($)	(5) Δlog($)	(6) Δlog($)
Committee with jurisdiction over the firm's industry	0.0393***			0.0908***		
	(0.0029)			(0.0173)		
Geography: firm HQ or a 250+ job plant in the district	0.0278***			0.0367		
	(0.0053)			(0.0238)		
Gains industry jurisdiction		0.0459***	0.0459***		0.2186***	0.2186***
		(0.0031)	(0.0031)		(0.0213)	(0.0213)
Loses industry jurisdiction		-0.0291***			-0.0120	
		(0.0037)			(0.0177)	
loses it, stays on committees			0.0024			0.0095
			(0.0022)			(0.0169)
loses it by leaving all committees			-0.0895***			-0.6222***
			(0.0073)			(0.0724)
Firm × candidate FE	Yes	differenced	differenced	Yes	differenced	differenced
Cycle FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,676,386	2,650,152	2,650,152	215,398	134,847	134,847
R^2 (within)	0.005	0.004	0.006	0.002	0.006	0.007

Table 8: Motive composition by election cycle

Columns 4–8 are the share of the cycle's donations on which each motive fires. They are not mutually exclusive and do not sum to 100. *Industry* is a committee whose remit reaches the donor's own industry through a sectoral instrument; *Economy-wide* is a committee that reaches the industry only as it reaches every industry; *Tax/spending* is Ways and Means or Appropriations; *Geography* is a firm headquarters or 250+ job establishment in the contested district. *Others* is the share on which none of the five reasons in Table 2 fires, and is the last row of that ladder. Dollars are nominal, which accounts for part of the rise in the mean gift. The jurisdiction share is close to flat while the district share rises by half and the unexplained share falls. Estimating the choice of candidate cycle by cycle shows the same period differently, with the pull of jurisdiction rising throughout.

	Donations	$m	Mean $	Industry	Economy-wide	Tax/spending	Geography	Others	Incumbent
2000	20,160	40	1,967	48.7	22.0	25.1	9.3	15.6	87.5
2002	20,325	44	2,181	50.8	24.5	24.3	10.0	14.6	86.4
2004	18,889	46	2,462	54.3	28.3	25.3	10.8	11.2	90.3
2006	19,833	54	2,735	53.8	24.6	26.9	11.9	9.3	93.6
2008	21,133	60	2,822	55.7	25.1	24.6	11.7	9.7	91.8
2010	25,829	81	3,145	51.9	23.0	22.9	14.1	13.3	87.4
2012	28,279	94	3,313	54.2	21.1	20.8	14.0	11.2	90.5
2014	28,383	97	3,417	54.3	22.9	24.5	15.2	9.0	93.1
2016	27,903	100	3,597	54.9	21.8	24.8	14.2	8.4	94.6
2018	28,253	105	3,721	53.4	22.4	25.2	12.6	10.1	94.3
2020	24,640	92	3,749	55.1	24.5	26.5	13.7	8.9	96.1
2022	21,445	81	3,778	55.4	24.1	27.1	13.4	9.0	95.2

Table 9: Where the two motives bite hardest

Each row is a separate linear probability model on that subsample of the choice set, with firm × cycle and candidate × cycle fixed effects and standard errors clustered two ways on firm and candidate, reported in parentheses. The base rate is the unconditional probability of giving within the subsample; the ÷ base columns divide the coefficient by it, which is the comparable quantity across rows because the base rate ranges from under half a per cent to seventeen. PAC size quartiles are formed on the number of candidates the firm funded in the cycle, and the median count is given in parentheses. Challengers and open-seat candidates hold no committee seat while they run, so the jurisdiction estimate in Panel B's second and third rows rests on the small number of former members seeking to return and is the one coefficient in the table that is not distinguishable from zero for open seats. * p<0.10, ** p<0.05, *** p<0.01.

	Base rate (%)	Industry jurisdiction	÷ base	Geography	÷ base	N
Panel A. By the size of the firm's PAC programme						
Smallest quartile (median 3 candidates)	0.47	0.0035*** (0.0004)	0.74	0.0866*** (0.0094)	18.28	1,306,566
Second quartile (11)	1.74	0.0131*** (0.0011)	0.75	0.1430*** (0.0139)	8.20	1,164,843
Third quartile (28)	4.52	0.0319*** (0.0030)	0.71	0.1660*** (0.0142)	3.68	1,209,783
Largest quartile (88)	17.00	0.0999*** (0.0069)	0.59	0.1480*** (0.0116)	0.87	1,199,744
Panel B. By the candidate's status						
Incumbent	8.19	0.0371*** (0.0028)	0.45	0.1898*** (0.0114)	2.32	3,199,684
Open seat	1.82	0.0054 (0.0035)	0.30	0.0748*** (0.0060)	4.11	808,087
Challenger	0.96	0.0286*** (0.0091)	2.97	0.0371*** (0.0032)	3.85	869,955
Panel C. By period						
2000–2010 cycles	5.12	0.0288*** (0.0022)	0.56	0.1472*** (0.0102)	2.87	2,464,860
2012–2022 cycles	6.58	0.0441*** (0.0034)	0.67	0.1393*** (0.0089)	2.12	2,416,076

Table 10: Who receives money that no firm-specific reason explains

The unit is the candidate–cycle, restricted to the 7,898 House candidate–cycles that received corporate PAC money from at least one sample firm. Reach is the share of the grid's 76 GICS6 industries over which the candidate's sitting committees hold sectoral jurisdiction: zero for a candidate with no seat, 0.92 for a member of Energy and Commerce alone. 'No reason' is a donation on which none of the five reasons in Table 2 fires; columns 3 and 5 weight donations equally and column 4 weights them by dollars. Column 5 keeps incumbents only, which is where the question has content – a challenger holds no committee, so almost every donation to one is unexplained by construction. **Columns 3–5 are partly definitional and should be read as magnitudes rather than as tests**: reach, the tax and spending committees, and a leadership post are three of the five rungs of the ladder that defines the dependent variable, so each must push the unexplained share down. What the columns add is how far, and the answer is that reach and the money committees do comparable work while a leadership post does about a third as much. Columns 1, 2 and 6 carry no such circularity: nothing about how much a candidate raises, from how many firms, or from how many different sectors enters the ladder. Column 6 conditions on the number of donors, so it reads as the sectoral spread of a donor base of given size; the mean is 6.7 sectors, against which the coefficients are small and precisely estimated rather than large. Standard errors are clustered on candidate; * p<0.10, ** p<0.05, *** p<0.01.

	Corporate money raised		Donations with no reason			Donor mix
	(1)	(2)	(3)	(4)	(5)	(6)
	log($ raised)	log(donors)	Share	$ share	Share	Industries
Share of the 76 industries the committees reach	1.9781***	1.2125***	-0.4199***	-0.3943***	-0.3975***	-0.4848***
	(0.2182)	(0.1145)	(0.0206)	(0.0233)	(0.0215)	(0.0872)
Committee on tax or spending (Ways and Means, Appropriations)	1.7354***	0.9952***	-0.3458***	-0.3213***	-0.3282***	-0.2707***
	(0.1600)	(0.0849)	(0.0157)	(0.0182)	(0.0163)	(0.0617)
Party leadership post	0.7766***	0.4902***	-0.1155***	-0.1062***	-0.1116***	-0.4185***
	(0.1134)	(0.0627)	(0.0118)	(0.0115)	(0.0113)	(0.0798)
Number of committees held	0.0280	-0.0385*	-0.0470***	-0.0424***	-0.0421***	0.1393***
	(0.0471)	(0.0232)	(0.0036)	(0.0044)	(0.0036)	(0.0214)
Incumbent	1.6547***	1.5566***	-0.3596***	-0.3790***		
	(0.1874)	(0.0906)	(0.0193)	(0.0243)		
log(corporate donors)						2.1333***
						(0.0122)
Cycle FE	Yes	Yes	Yes	Yes	Yes	Yes
Incumbents only	No	No	No	No	Yes	No
Observations	7,898	7,898	7,898	7,806	5,180	7,898
R^2 (within)	0.319	0.540	0.768	0.720	0.441	0.926

A1 Variable Definitions

Variable names below are the column names in the four analysis files, so a table in the paper, a coefficient in a regression and a column in the replication data carry the same label. Cycle t runs from the January after the previous general election to the general election in November of year t. *Sitting Congress* means the Congress in session when the donation was made, which is the Congress elected at $t-2$; *incoming map* means the district boundaries contested at t. Section 2.3 explains why the two conventions differ.

Variable	Definition and source
Panel A. Identifiers and units	
`gvkey`	Compustat firm identifier.
`cand_id`	FEC candidate identifier, as carried by DIME (Bonica, 2024).
`cycle`	Election cycle, even years 2000–2022.
`firm_cycle`	Group identifier for `gvkey` × `cycle`; 7,410 groups. Absorbed as a fixed effect.
`cand_cycle`	Group identifier for `cand_id` × `cycle`; 7,898 groups. Absorbed as a fixed effect.
`pair`	Group identifier for `gvkey` × `cand_id`; 130,171 groups in the donation sample.
`hist_gics2`, `hist_gics4`, `hist_gics6`	Historical GICS sector, industry group and sub-industry codes for the firm in the donation year. The grid is defined at GICS6, of which 76 appear in the sample.
Panel B. Donation variables	
`amt`	Total corporate PAC dollars given by the firm to the candidate in the cycle, summed over itemised FEC contributions. Nominal. Negative values are net refunds and are retained.
`ln_don`	$\log(\max(\texttt{amt}, 1))$. Defined on the 280,602 donations above zero.
`refund`	$= 1$ if $\texttt{amt} < 0$.
`gave`	Choice set only. $= 1$ if the firm gave to this candidate in this cycle; 0 for a candidate in the same cycle's choice set that the firm passed over.
`ln_don_0`	Choice set only. $\log(\max(\texttt{amt}, 1))$ with $\texttt{amt} = 0$ on the passed-over cells.
`gvkey_cand_don_qtr_count`	Number of calendar quarters of the cycle in which the firm gave to this candidate.
`earliest_gvkey_cand_d_cycle_qtr`	The first of the cycle's eight quarters in which the firm gave to this candidate.
`gvkey_cand_don_nth_cycle`	The number of consecutive cycles, ending at t, in which the firm has given to this candidate.
`repeat_donor`	$= 1$ if the firm gave to this candidate in cycle $t-2$.
`gave_next`	$= 1$ if the firm gives to this candidate in cycle $t+2$. Missing for 2022, which has no successor in the window.

continued on the next page

Variable	Definition and source
Panel C. Motive variables	
`cmte_sector_juris`	= 1 if `n_cmte_sectoral` > 0: at least one committee of the sitting Congress on which the candidate serves reaches the firm's GICS6 industry through a sectoral instrument. From the committee × industry grid, Appendix A3.
`cmte_horiz_reach`	= 1 if `n_cmte_horizontal` > 0: at least one such committee reaches the industry only through a legal function applying to firms in every sector.
`n_cmte_sectoral`	Count of the candidate's sitting committees with grid code `reach` = 2 for the firm's industry.
`n_cmte_horizontal`	Count with `reach` = 1.
`n_cmte_noreach`	Count with `reach` = 0. Not separately identified within a candidate × cycle cell, where the total number of seats is fixed.
`cmte_wm_approp`	= 1 if the candidate sits on Ways and Means or Appropriations in the sitting Congress. String match on `cmte_l1`; equals max(`cmte_wm`, `cmte_approp`).
`geography`	= 1 if the firm's headquarters, or an establishment employing 250 or more people, lies in the district contested by the candidate at t. Equals max(`in_hq`, `in_estab_250`) from the firm–cycle–district panel, Appendix A2.
`in_hq`	= 1 if the firm's headquarters district, from the SEC filing header through a cycle-keyed ZIP crosswalk, is the contested district.
`in_estab_250`, `in_estab_500`	= 1 if the firm operates an establishment of at least 250 (500) employees in the contested district, placed by point-in-polygon against the incoming map. The 500-employee variant is retained for sensitivity and is not used in the reported results.
`ladder`	The reporting convention of Table 2: the most firm-specific reason that applies, coded 1 industry jurisdiction, 2 geography, 3 tax and spending, 4 economy-wide jurisdiction, 5 party leadership, 6 Others.
`no_reason`	= 1 if `ladder` = 6: none of the five reasons fires. Displayed throughout as *Others*.
`firm_specific`	= 1 if `cmte_sector_juris` = 1 or `geography` = 1.
Panel D. Candidate characteristics	
`cmte_l1`	The candidate's committee assignments in the sitting Congress, as a delimited string with the role in parentheses (Stewart and Woon, 2024).
`n_cmte_l1`	Number of committees held in the sitting Congress.
`cmte_any`	= 1 if `n_cmte_l1` > 0. This is the standard committee proxy of Table 3.
`cmte_gradeable`	= 1 if the candidate holds at least one of the 19 committees the grid covers.
`cand_power_l1`	= 1 if the candidate holds a party leadership post in the sitting Congress: Speaker, floor leader or whip. From `seniorpartymember` in Stewart and Woon (2024).

continued on the next page

Variable	Definition and source
cand_ever_power_l1	= 1 if the candidate has held such a post in any prior Congress.
cand_ici	Incumbency status at t from DIME: I incumbent, C challenger, O open seat.
is_incumbent, is_challenger, is_openseat	Indicators from cand_ici.
cand_pty, rep_cand	Party from DIME; rep_cand = 1 for Republicans.
cand_office_st, cand_office_district	State and district number of the seat contested at t.
reach_breadth	Candidate–cycle only. Share of the grid's 76 industries over which the candidate's sitting committees hold sectoral jurisdiction: 0 for a candidate with no seat, 0.92 for a member of Energy and Commerce alone.
Panel E. Firm and collapsed variables	
hq_st, hq_st_match_cand	Headquarters state, and an indicator that it equals the candidate's state. The home-state proxy of Table 3.
hq_cd_match_cand	= 1 if any district the headquarters ZIP maps to equals the contested district. The home-district proxy.
n_cands	Number of candidates the firm funded in the cycle.
pac_q	Quartile of n_cands formed across firm–cycles, not across choice cells; medians of 3, 11, 28 and 88 candidates.
ttl, ln_ttl	Dollars given by the firm, or received by the candidate, in the cycle, and its log.
n_donors, ln_donors	Number of sample firms funding the candidate in the cycle, and its log.
n_industries	Number of distinct GICS sectors among the candidate's donors in the cycle.
share_sectoral, share_geography, share_no_reason	Share of the firm's (or candidate's) donations in the cycle on which the reason fires.
dollar_share_no_reason	Share of the candidate's cycle receipts contributed by donations in the Others category.
avail_sectoral, avail_geography	Share of the cycle's own choice set, for that firm, that carried the reason. The benchmark for targeting.
excess_sectoral, excess_geography	Realised share less available share: targeting above what giving at random within the cycle would produce.

A2 Sample Construction

A2.1 The donation sample

Contributions are itemised FEC records of corporate PAC gifts to federal candidates, aggregated to the firm–candidate–cycle. A cycle runs from the January after the previous general election to the general election in November of the even year, so a gift made in March 2011 belongs to the 2012 cycle. The sample is restricted to House candidates and to donor PACs sponsored by firms with a Compustat identifier, because the second half of this research programme requires accounting data. Neither restriction is imposed after a draw: every merge runs before the sample is fixed, so nothing is selected and then discarded.

Candidate attributes come from DIME (Bonica, 2024), joined on the FEC candidate identifier and cycle. Only fields realised before the election are taken: incumbency status, party, district and gender. Vote shares, winner indicators and general-election status are dropped at the point of the merge rather than filtered later, so no downstream step can reintroduce them. DIME's contributor-ideology scores are also excluded: the static CFscore is fitted over the full 1979–2024 history and the dynamic score at cycle t is fitted partly on cycle-t contributions, including the donation being classified.

Committee assignments come from Stewart and Woon (2024) and are lagged one Congress, so a donation in cycle t carries the committees of the Congress elected at $t-2$ and sitting when the gift was made. Party leadership is taken from the same source's senior party member coding: Speaker, floor leaders and whips.

Three features of the raw data are retained rather than cleaned away. Donations at or below zero occur—1,582 net refunds and 2,888 exact zeros—and are kept and flagged, entering every specification except those requiring a logarithm. A firm's headquarters ZIP straddles more than one congressional district in a substantial share of firm–cycles, so headquarters districts are listed rather than chosen among. And a small number of firm–cycles map to two headquarters states, which is a source defect; it is flagged and not corrected.

A2.2 The firm–cycle–district panel

Geography is built rather than read off an address, because a firm's headquarters is one location and its operations are many. The panel records, for each firm and cycle, every congressional district in which the firm has a qualifying presence.

Establishments come from the Infogroup establishment file, matched to Compustat on ticker and restricted to the cycle's own even year, so a mid-cycle relocation cannot move a firm between districts within a cycle. Each establishment is placed in exactly one district by point-in-polygon against Census TIGER/Line boundary files for the Congress the cycle elected. The vintage rule is the trap in this step and is worth stating: the TIGER release for year Y contains the Congress *sitting* in Y, not the one elected in Y, so the file for cycle Y is the release for $Y + 1$. Getting this wrong in 2012 or 2022 would place firms in boundaries that were never contested.

Headquarters come from SEC filing headers rather than from Infogroup, and the choice is deliberate. Infogroup's coverage of listed firms rises steadily across the window, from under a third in the late 1990s to around three-fifths by 2024, so a geography measure built on it alone would carry a mechanical time trend. The filing-header file has no such gradient. The two sources place a firm in the same district set on 91.3% of firm–years where both are available, and where they differ the median distance is 9.4 kilometres, so the substitution costs little and removes a trend.

The employment floor does real work rather than trimming a tail. Without it, a firm's establishments span a median of 57 congressional districts, which is not a geographic tie to anybody; at 250 employees the median falls to 2. A materiality rule does not fix ubiquity on its own—a national restaurant chain clears any threshold in every district—and no ubiquity screen is applied, on the ground that a chain with a material presence everywhere genuinely has one.

A2.3 The choice set

The choice set pairs every firm–cycle with every candidate–cycle in the same cycle. A candidate enters a cycle's set if at least one sample firm funded them in that cycle, giving a

mean of 658 candidates per cycle and 4,880,936 cells in total, of which 285,072 are realised donations.

Motive variables for the passed-over cells are rebuilt from primitives rather than imputed. The committee counts are recomputed by parsing each candidate's committee string and looking each committee up against the donor's industry in the grid; geography is recomputed by joining the firm–cycle–district panel on the contested district. Both use the code that produced the same variables for the realised cells. The build then verifies that the reconstruction reproduces all 285,072 donated rows exactly on `geography`, `n_cmte_sectoral`, `n_cmte_horizontal` and `n_cmte_noreach`, and halts otherwise. Anything short of exact agreement would mean the counterfactual cells and the realised ones measure different things.

Parsing the committee string requires one piece of care that is easy to miss. Assignments are delimited by commas, and several committee names contain commas—"science, space, and technology"—so a naive split silently drops committees and understates jurisdiction. The split runs on the closing parenthesis of the role suffix instead, and leadership prefixes are stripped before matching.

A2.4 The two collapsed samples

The candidate–cycle sample aggregates donations to the recipient: 7,898 observations recording receipts, donor count, the sectoral spread of the donor base, and the share of donations in each motive category. The firm–cycle sample aggregates to the donor: 7,410 observations recording dollars given, candidates funded, and the share of the firm's donations carrying each reason.

The firm–cycle sample also records what the cycle made available. For each firm–cycle, `avail_sectoral` is the share of that cycle's own choice set carrying jurisdiction over that firm, and `avail_geography` the share carrying a qualifying presence. These are the benchmarks in Figure 3: a high realised share is not evidence of targeting for a firm in a heavily regulated industry, and the difference between realised and available is.

Both collapses are conditional on participation. A candidate no sample firm funded is absent from the first, and a firm that funded nobody from the second, so neither speaks to the decision to participate at all. That question is the choice set's.

A3 The Committee × Industry Grid

A3.1 The coding standard

Every cell of the grid answers one question about one (industry, committee) pair:

> Credit a committee only where its remit reaches this industry through a *sectoral instrument*—a statute, programme or regulatory regime whose scope is defined by what this industry makes, sells or operates. Where the instrument is *horizontal*—defined by a legal function and applying to firms in every sector, with only the intensity varying—it is not sectoral jurisdiction, however large the stake. Regulation and government purchase both count, provided the programme doing the buying is specific to what this industry makes.

The operational test is whether a firm in an unrelated industry would describe the same stake in the same words. If it would, the instrument is horizontal. Cells are coded on a three-way scale:

`reach`	Definition	Instrument named?
2	The committee reaches this industry through a sectoral instrument.	required
1	The committee reaches this industry with identifiable force, but the instrument is horizontal.	required
0	No exposure beyond what every firm in the economy faces.	none

The line between 1 and 0 is the requirement to name an instrument and show that it bites this industry particularly. Every committee touches every industry in some remote sense, and that is not a reason to give anybody money; requiring the instrument on the record is what makes the distinction auditable rather than a matter of taste. Each credited cell stores the instrument (`strand`), the part of the industry it reaches where that is not all of it (`segment`), and a note of at most fifteen words.

A3.2 Committee scope

The grid covers nineteen committees: the twenty House standing committees less Ways and Means and Appropriations, plus the Permanent Select Committee on Intelligence.

Ways and Means and Appropriations are excluded because neither has an industry-specific remit to test a firm against, and both are supplied to the analysis mechanically as `cmte_wm_approp`. Intelligence is included because it holds genuine legislative jurisdiction and is not redundant with the money committees: on the trial sample, 21 of its 23 donation rows are invisible to `cmte_wm_approp`. Joint Taxation is excluded for the opposite reason, its rows being a subset of Ways and Means. Leadership posts are excluded from the grid entirely and recorded as `cand_power_l1`; crediting them would put a firm-invariant fact into the one field whose purpose is firm-specificity.

Industries are the GICS six-digit sub-industries appearing in the sample: 69 in the original build and a further seven added when the sample was extended to the full House population, for 76 in total and $76 \times 19 = 1{,}444$ cells. The increment was judged under the same standard and by the same protocol as the original.

A3.3 Isolation, and why it is the design

Each cell is judged in a context containing no other cell. One agent receives one industry, the standard, and the list of nineteen committees, and returns nineteen rows. It receives nothing about any other industry, nothing about any other agent's output, and nothing about how any earlier version of the grid coded its cells. Three rules enforce this: no agent's answers enter another agent's prompt; no earlier grid version is read before a cell is dispatched; and no industry is re-dispatched because its answer is unwelcome.

The last rule is the one that costs something and it is the one that matters. Re-running a cell until it matches an expected number is tuning, and it would destroy the reproducibility the design exists to produce. A file that parses is final.

Isolation buys determinism by construction rather than by measurement. Identical inputs cannot receive different labels because no cell is judged twice, and no cell's label can depend on the order in which cells were processed, on the length of a context window, or on the labels of neighbouring cells. It also makes the grid re-runnable one cell at a time, so a reader who disputes a single cell can re-judge that cell alone and compare.

A3.4 The ex-ante prediction, and what the run returned

A prediction was written and recorded before the grid was built, so the run would be a test rather than another pass. Four of its components are worth reporting because two were met and two were not.

Budget, Rules, Ethics and House Administration were predicted to return no sectoral cell for any industry, on the ground that none has an industry-specific remit to have one. All four did. That is the prediction that carries information about whether the standard was applied as written, and Figure 2 shows what it implies: the three of them that attract material money attract it entirely from industries their jurisdiction does not reach.

Overall density and Energy and Commerce were predicted to fall further than they did. The prediction was 190–240 sectoral cells of 1,311 against a realised 367, and 35–48 cells for Energy and Commerce against a realised 64. The grid was retained as built. It is judged on whether it applies one standard to every cell and names an instrument wherever it credits one, not on the size of its movement from a previous version coded under a different standard, and adjusting it to hit a recorded range would have been the tuning the isolation rule forbids. The prediction is reported here as a recorded prior rather than as a gate that was applied.

Two earlier grid versions are retained in full, together with the cell-level differences between them. The first applied a loose standard to every committee; the second applied a strict standard to Energy and Commerce and the loose one elsewhere, an asymmetry that could not be defended in print. Retaining the diffs is what makes convention-sensitivity measurable: since no register of correct answers exists, the honest statistic is how far the motive distribution moves under a defensible alternative coding, and that requires the alternative to have been kept.

A3.5 Worked examples at each boundary

Sectoral against horizontal, same statute. Clean Air Act Title II certifies an engine as a product before it may be sold and reaches only firms building engine-powered equipment, so Machinery × Energy and Commerce is a 2. Clean Air Act facility permitting reaches every manufacturing plant in the economy, so where that is all a committee has, the cell is a 1.

The seam runs through the statute, not around the committee.

Sectoral against horizontal, same committee. Judiciary is a 2 for Pharmaceuticals, where Hatch-Waxman and the Controlled Substances Act are written for that industry alone, and a 1 for Electric Utilities, where antitrust and bankruptcy are all there is. Education and the Workforce is a 2 for coal-bearing Metals and Mining under the Mine Act, which covers mines and nothing else, and a 1 for a bank under ordinary employment law.

Horizontal against no reach. Chemicals × Education and the Workforce is a 1: OSHA's process safety management standard is written about chemical processes, so an instrument can be named and shown to bite. Software × Education and the Workforce is a 0: ordinary employment law is all there is, and every firm faces it.

A sectoral committee with horizontal cells. Agriculture, about as sectoral a committee as exists, carries eight cells at `reach` = 1: industries it reaches only through commodity-derivatives law, which touches every hedger. Sectoral and horizontal are properties of the pair, not of the committee.

A3.6 What the grid does not establish

No register of correct answers exists for which committee has jurisdiction over which industry, so the grid cannot be scored against ground truth and is not claimed to be. What a reader is entitled to is a single standard applied uniformly, an instrument named wherever a cell is credited, an output small enough to inspect whole, and a measured statement of how much the results move under an alternative coding.

External validation remains outstanding and its design is fixed. A firm's lobbying disclosures name the bills it worked and therefore the committees it treats as having jurisdiction over it, which is revealed preference over exactly the object the grid encodes. The test is whether the grid predicts that contact net of industry and committee fixed effects. Lobbying data must stay out of the grid's construction for the test to mean anything, which is why the grid is built from statutes alone.