Bureaucratic Information in Congress*

Luca Bellodi†

Most Recent Version

Abstract

Bureaucratic agencies produce a wealth of information that can be used by politicians when making policies. However, little is known about how often politicians rely on bureaucratic information and the factors they consider when they do so. I introduce a novel measurement strategy that applies natural language processing to extract and analyze bureaucratic information cited by US members of Congress (MCs) in 8.5 million congressional speeches (1980-2022). I establish three findings. First, using novel dynamic measures of agency ideology estimated from bureaucrats' campaign contributions, I find that MCs are less likely to cite information from ideologically distant bureaucracies. Second, when multiple bureaucracies operate in the same policy domain, MCs prioritize information from independent agencies. Finally, I leverage the 2020 Supreme Court's decision in *Seila Law v. Consumer Financial Protection Bureau* (CFPB) – which curtailed the independence of the CFPB director – as a shock to the independence of the bureau and find that MCs' citations of CFPB information drops after the Court's decision. These findings contribute to a deeper understanding of the dynamic relationships between branches of government in the US system of separation of power.

*For their helpful comments, I am thankful to Alex Acs, Brandice Canes-Wrone, Gary Cox, George Krause, Mary A. Kroeger, David Lewis, Kenneth Lowande, Massimo Morelli, Christoph Pfeufer, Jennifer Selin, Hye Young You, and seminar participants at Stanford University, the London School of Economics and Political Science, and the University of California, Berkeley Goldman School of Public Policies. The usual disclaimers apply.

†Hoover Institution and Department of Political Science, Stanford University

Information exchange between elected politicians and unelected bureaucracies is key to effective governance. In order to make informed decisions on complex policy issues, politicians need the specialized knowledge and expertise of bureaucratic agencies. In the United States, bureaucratic agencies produce a great wealth of information that can be used by politicians for both policy and electoral goals (Niskanen 1971, Wilson 1989). Bureaucratic information can help members of Congress persuade political opponents, make better policies, or frame the debate around certain policy issues in their favor. At the same time, citing bureaucratic information can represent a form of position-taking, with members of Congress signaling their effort and commitment to policies to their constituents and donors (Mayhew 1974, Maltzman & Sigelman 1996, Grimmer 2013).

For example, in a statement given on July 8, 2015, during a session of the US Senate Committee on Environment and Public Works, Senator Barbara Boxer (D-CA) highlighted some of the positive outcomes of President Obama's Climate Action Plan, an ambitious set of measures aimed at cutting carbon emissions:

> *The recent study by the Environmental Protection Agency shows us 57,000 fewer deaths per year from poor air quality, with economic benefits valued at $930 billion, 12,000 fewer deaths per year from extreme heat and temperature changes, $180 billion per year in avoided damages from water shortages, $3 billion per year avoided damages from poor water quality, $11 billion a year avoided losses in our ag sector, 40 to 59 percent fewer severe and extreme droughts and almost 8 million fewer acres burned each year from wildfires.*

In her speech, the Senator cited evidence produced by the Environmental Protection Agency (EPA) to persuade the Republican-controlled committee and its chairman Senator Jim Inhofe (R-OK), to take action against climate change. These are the concluding words of her speech: "I feel stronger than ever that President [Obama] is on the right path. This Committee is on the wrong path." Far from being an isolated case, my data show that during the past 40 years, the information produced by the EPA has been used in congressional speeches over 3,068 times by a total of 581 different members of Congress.

Information is also a core tenet of bureaucratic legitimacy, and producing expertise is main reason why unelected bureaucratic agencies are delegated significant discretion in administering policy (Gailmard & Patty 2013). However, little is known about the extent to which members of

Congress rely on such information. When do members of Congress cite bureaucratic information? And what do they consider when selecting which information to use? In this paper, I draw on a large literature on bureaucratic expertise and legislators' use of information to study the role of bureaucratic information in Congress, a question that has broad implications for effective governance and bureaucratic accountability.

Bureaucratic agents often possess greater expertise compared to political principals (see, e.g., Epstein & O'Halloran 1999, Miller 2005), and the scarce resources at legislators' disposal limit their ability to directly evaluate information. Therefore, legislators rely on heuristics to decide which information to prioritize. Building on canonical models of information (Crawford & Sobel 1982, Gilligan & Krehbiel 1987), I argue that two important heuristics used by members of Congress are the ideology and the independence of bureaucracies. On the one hand, agency ideology allows legislators to infer whether the information is compatible with legislators' political goals and enterprises. Climate change skeptics are likely to believe that the EPA report mentioned by Senator Boxer is blatant propaganda, at odds with their own preferences and those of hardline conservative voters and donors. On the other hand, when agencies enjoy a high level of statutory independence and are insulated from political pressures, they cultivate a reputation for expertise, they have incentives to specialize, and their outputs are perceived to be more accurate by members of Congress. Agency ideology informs members about the compatibility of the information with members' goals. Agency independence informs members about the quality and accuracy of the information.

To test these predictions, I present the first attempt at studying MCs' use of the information produced by hundreds of US federal agencies over the past 40 years. I introduce a new measurement strategy that uses natural language processing (NLP) to detect when MCs cite policy information produced by bureaucratic bodies in their speeches and apply it to an original corpus of 8.5 million speeches given by members of Congress in the floor and committee hearings. First, I apply dependency parsing to the corpus of speeches and extract legislators' citations of bureaucratic information. Second, I used supervised machine learning and large language models to classify whether the citation contains policy information or not. For every member-agency-congress combination, I am able to quantify how often members mention, cite, and cite policy information produced by the agency. This measure yields a comprehensive picture of the role of bureaucratic information in Congress over a large period of time for a wide set of agencies and legislators with different partisan affiliations. Moreover, I use bureaucrats' campaign contributions to build the largest dataset of dynamic agency ideology ever assembled, covering a period that spans nine presidential administrations from Clinton to Biden.

I present descriptive evidence alongside a rigorous test of the role of ideology and independence in MCs' consumption of bureaucratic information. A key stylized fact I document is that, although Republicans and Democrats cite bureaucracies in their speeches to a similar extent, Republicans make greater use of the expertise of drug- and law-enforcement agencies, whereas the information produced by the Department of Education and the Centers for Disease Control and Prevention is mostly cited by Democrats. This is an indicator of the different priorities of Republicans and Democrats: crime and legality for the Republican party, and healthcare and education for the Democratic party (Egan 2013). Furthermore, I show that MCs' citations of bureaucratic information are decreasing on the floor, whereas they remain stable – and slowly increase in most recent years – in committees. This suggests how the floor is becoming a venue where members score political points, whereas committees remain deliberative arenas where legislators can substantively engage with policy information.

After presenting several stylized facts, I leverage *within*-agency changes in ideological leaning resulting from bureaucratic turnover in leadership positions across presidencies to estimate the effect of ideological distance between members and agencies on members' citations of bureaucratic information. To do so, I compare the frequency of members' citations of policy information produced by agencies and estimate a series of dyadic fixed effects models. I find strong support for the proposed ideology-driven account. To estimate the effect of agency independence, I first compare how members select information from independent and non-independent agencies operating in *the same policy area* and find that independent agencies' information is more likely to be cited by MCs compared to information produced by non-independent agencies. To strengthen these *between*-agencies comparisons, I exploit the 2020 Supreme Court decision curtailing the independence of the director of the Consumer Financial Protection Bureau and compare how members' reliance on information produced by the CFPB changed after the Court's decision. I find that, after the Court's decision, members are significantly less likely to cite information from the CFPB.

This paper makes three contributions to the literature on American political institutions and on the political use of information. First, while a large literature portrays bureaucratic agencies as expert bodies (Gailmard & Patty 2013), there is little empirical evidence on whether, how often, and under what conditions members of Congress use this information. Second, while most of the scholarship on politician-bureaucracy relations focuses on how Congress and presidents control drifting bureaucracies (e.g., Fiorina 1981, McCubbins et al. 1987, Bolton et al. 2016, Lowande 2018, but see Moe 2006), in this paper I show that bureaucracy can play an important role in congressional politics, accounting for a prominent source of information at legislators' disposal. Third, I introduce the first and largest measure of the role of bureaucratic bodies in legislative politics, presenting fine-grained data for 317 agencies and approximately 40 years of floor and committee speeches. The proposed measures allows research to study other questions about information provision and usage across different fields in political science. The findings also have implications for the institutional design of bureaucratic agencies, suggesting that statutory features granting independence to agencies can counteract the ideological polarization underlying legislators' decision to use bureaucratic expertise in Congress.

Bureaucratic Information in Legislative Politics

Since Weber (1946), bureaucratic expertise has been portrayed both as a resource for politicians and, at the same time, as a threat. A resource because legislators can learn from bureaucrats about the expected consequences of policies. A threat because the informational advantage of bureaucrats *vis-à-vis* their political principals often creates a problem of control (McCubbins & Schwartz 1984, Aberbach 1990).Due to this informational advantage, bureaucratic agencies can exert considerable influence over policy, sometimes extending beyond their statutorily defined boundaries (Moe 2012).

Krause (1996), for instance, describes agency-political relations as a "two-way street", where agency performance can affect politicians' budgetary preferences, and Carpenter (2001) shows that, during the US Progressive Era, bureaucratic agencies enjoying a good reputation across multiple audiences were able to secure their desired policies despite the opposition of elected politicians. Zooming in on legislative politics, Nicholson-Crotty & Miller (2012) find a positive relationship between an agency's perceived effectiveness and politicians' perceptions of bureaucratic influence on legislative outcomes, while Blom-Hansen et al. (2021) find that local politicians in four different countries are likely to rely on bureaucrats' expertise and information when forming their opinions about policy proposals.

Using observational data, Shobe (2017) shows how federal agencies play an important role as reviewers and editors of legislative texts, either by request from Congress or as a result of the agency's own monitoring of legislation. Kroeger (2022) reports similar findings on state legislation and finds that bureaucracy-sponsored bills are more likely to be approved by the legislature when there is unified government and when the capacity of the legislature is weaker compared to that of the bureaucracy. A similar demand-side approach to studying politicians seeking information is explored by Ban et al. (2023), who map the universe of witnesses testifying before congressional committees. They show that bureaucrats—in addition to being the largest category of witnesses—are important providers of analytical information, and they are invited to testify most often when legislators are exploring a legislative issue and are open to acquiring new information. Similarly, Ban et al. (2024) look at bureaucrats' incentives to supply information and find that, when appearing before committees, bureaucrats supply more analytical information to legislators who are presidential co-partisans, suggesting that ideology is key not only to legislators' decisions to use the information but also to bureaucrats' decisions to supply it.

An important question that remains unanswered is when members of Congress utilize bureaucratic information and the factors they consider when deciding to use it. In what follows, I build on a large literature on strategic communication and bureaucratic expertise and test the predictions with new data on MCs' reliance on information produced by the US federal bureaucracy in Congress.

When Do MCs' Cite Bureaucratic Information?

Members of Congress are moved by a combination of re-election and policy goals, and giving speeches – arguably one of the main activities of elected officials – can be a powerful tool to frame a policy problem in their favor, strengthen a political argument, or communicate effectively to constituents and interest groups (Mayhew 1974, Grimmer 2013, Grimmer et al. 2014, Lee 2016). Bureaucracies are a one-stop-shop for MCs seeking to acquire policy information (Ban et al. 2023). Members of Congress might report what an independent agency or a department says to claim credit for the success of a program or to direct voters' attention to a specific issue. Agency expertise can be mobilized to increase the salience of a political debate, influence the political agenda, or persuade other members through credible information. However, not all information produced by bureaucracy is the same, and members of Congress know that bureaucracies can supply information to advance their own preferred policy or that of a president from the other party, and they might be hesitant to use information from ideologically distant bureaus.

Politicians may decide to use information produced by the bureaucracy if they believe it is accurate and compatible with their own political goals. Therefore, when selecting information from bureaucracies, MCs consider two dimensions: quality and ideology. Quality refers to whether the information faithfully describes or meaningfully synthesizes reality, whereas ideology refers to the extent to which the information is consistent with MCs' goals and preferences. For instance, information that – more or less explicitly – highlights the social benefits of higher taxes is likely incompatible with the preferences of a libertarian politician. Perfectly informed politicians could scrutinize every piece of information and select the one that maximizes quality and ideological compatibility. However, agencies possess greater domain-specific expertise than most legislators, and members of Congress have limited time and resources to review each report or study produced by multiple bureaucratic agencies. In fact, the entire principal-agent literature on information asymmetries and politicians' delegation of authority to bureaucracy assumes that expert bureaucracies have an informational advantage over their own political principals (Miller 2005, Gailmard & Patty 2012), and a large literature on Congress suggests that members are under-staffed and spend most of their time raising money and campaigning (Currinder 2018, Lee 2016). Therefore, members of Congress resort to agency ideology and independence as heuristics when deciding whether to use the information or not, prioritizing information produced by ideologically similar and more independent agencies. I shall now consider the role of ideology and independence separately.

Let us consider a conservative member of Congress exposed to some information from the EPA under the Obama administration. The preferences of both actors toward environmental regulations are known, and the conservative legislator is likely to think that the EPA is distorting the information it produces to advance its liberal agenda. Notice that the information advantage of the bureaucracy does not mean that MCs are always incapable of assessing the accuracy and ideological slant of agency reports and studies. Even if the conservative member had expertise in environmental policy and could isolate accurate from biased information, she would quickly realize that the political implications of the information are at odds with her own policy preferences and would ultimately decide to disregard it. Therefore, MCs are less likely to cite information produced by ideologically distant bureaucracies, believing it to be incompatible with their own political priorities. This prediction is at the core of canonical models of strategic communication, which show that information exchange is greater when both the sender and the receiver have similar preferences over policy (Crawford & Sobel 1982, Gilligan & Krehbiel 1989, Austen-Smith 1990, Gailmard & Patty 2012, Denisenko et al. 2024).

PREDICTION 1: *MCs are less likely to cite information produced by ideologically distant bureaucracies.*

In practice, changes to the ideological leaning of agencies occur as a cascade: a new president is elected, and through appointments and turnover among agency leaders, she influences the policy priorities and ideological slant of agency communications and outputs.[1] As members of Congress observe the new ideological leaning of agencies, they update their beliefs about the expected compatibility of the information with their own political goals. However, ideology is not the only agency attribute observed by members of Congress.

Some agencies are designed to ensure a stronger independence from political and business influences. In certain policy areas, politicians delegate authority and grant statutory indepen-

[1] An example of these dynamics limited to the Environmental Protection Agency is reported in Section E1 in the Appendix.

dence to bureaucratic bodies to ensure consistency and credibility of agency policy over time and irrespective of changing governments (see e.g., Barro & Gordon 1983). By delegating independence to agencies, elected politicians shrink the degree of control that the president and Congress exert on bureaucracy, thus cultivating bureaucrats' incentives to make costly investments in expertise and agency specialization (Gailmard & Patty 2013). As outlined by Gilligan & Krehbiel (1987), principals might decide to restrict the array of procedures aimed at controlling the agent to promote the agent's incentive to specialize and acquire information, especially when the agent is ideologically apart from the parent body and when the policy issue is complex (Callander 2008). A clear example of such commitment is the independence of central banks and regulatory agencies for the credibility of monetary policies, controlling inflationary tendencies, and ensuring a level playing field for public and private companies (Cukierman et al. 1992, Keefer & Stasavage 2003). For instance, the Federal Reserve Board is governed by a multi-member body with proven expertise and fixed terms, whose members can only be dismissed for neglect of duty or malfeasance in office. These statutory features ensure that the Federal Reserve operates autonomously without responding to pressures from other political actors (Selin 2015). Recent empirical evidence suggests that independence and protection from discretionary appointments and removals increase the performance of public organizations (Aneja & Xu 2024).

Agency independence has important consequences for the way members of Congress perceive agency output, including information. By citing information from agencies that do not respond to the will of the government of the day, legislators can strengthen the perceived neutrality of information. For instance, independence improves the perceived and objective quality of regulation (Bertelli & Whitford 2009, Koop & Hanretty 2018), and independent agencies enjoy better reputations among political elites compared to more politicized agencies and departments (Bellodi 2023). Similarly, legislators exposed to information produced by agencies with valued reputations are induced to believe that certain types of lies (biased information) are extremely costly for the agency's reputation, increasing the perceived accuracy of information (Lupia & McCubbins 1998). Therefore, when exposed to information from independent agencies – even though politicians might disagree with the bureaucracy's ideology – legislators are more likely to believe that the information is accurate, given the strong commitment of the agency to its policy area.

When members of Congress know that the source of information is an independent body, they receive a signal about the information agencies produce, which increases the expected quality of information. Statutory independence, acting as a quality-enhancing device, increases members' reliance on bureaucratic information.

PREDICTION 2: *All else equal, MCs are more likely to cite information produced by independent bureaucracies.*

This account shows how agency ideology and independence are important shortcuts in legislators' process of prioritizing information. Ideology informs members about the compatibility of information with members' policy goals, while independence informs members about the quality and accuracy of information.

A New Measure of Legislators' Use of Information

In this section, I present a new, quantitative measurement strategy that captures the extent to which policy information produced by the bureaucracy is cited in Congress. I apply natural language processing techniques to a large corpus of floor and committee speeches by the universe of US members of Congress over the past 40 years that detect when MCs cite agencies' information and extract the type of information they use. This measurement strategy has two main advantages over existing methods. By looking at individual speeches, I am able to produce high-frequency, granular estimates for a large set of MC-agency dyads over a long period. This level of cross-sectional and temporal variation is particularly suitable for statistical analysis. Moreover, the proposed measure does not rely on respondents' perceptions and addresses the issues of social desirability bias.

Information Extraction

Standard text-analysis methods that rely on word frequencies are not suitable for detecting when a federal agency is simply mentioned (e.g., "I admire the work of the Fed") or cited (e.g., "The Fed reported that [...]"). One solution is offered by recent advancements in computational linguistics, which allow researchers to extract dependency relationships between words and identify when a (sequence) of words are used as sources of information. These techniques, often referred to as syntactic analysis or dependency parsing, can identify the action of saying something, the subject performing the action, and the object of the action. Let us consider the following statement: "The Fed *[subject]* said *[action]* that higher interest rates will slow inflation *[object]*." By creating extraction rules that detect certain dependency relationships, I can match every instance in which a bureaucracy is used as a source of information in a speech (i.e., a citation) and then analyze the type of information used. [2]

The measurement strategy I propose consists of three steps. First, I extract from each speech all sentences mentioning the name of an agency, tag parts of speech (e.g., subject, verb, direct object, etc.), and detect dependency relationships. Second, I extract clauses that match pre-defined syntactic frames, capturing different ways members can cite agencies. Third, I isolate the citation, namely the actual information used in the text, and use supervised machine learning to classify the content into policy and non-policy information. For each sentence containing the name of an agency, I observe whether the member mentioned (e.g., "I admire the work of the Fed"), cited (e.g., "The Fed said that in the long run, we're all dead"), and used policy information from the agency (e.g., "The Fed said that higher interest rates will slow inflation.").

Step 1: Parts-of-Speech Tagging and Dependency Parsing

I tag and parse the sentences mentioning agencies with SpaCy, a supervised learning algorithm that achieves state-of-the-art performance on several NLP tasks such as part-of-speech tagging and dependency parsing (Choi et al. 2015, Honnibal & Johnson 2015). Tokens – i.e., single words – within sentences have syntactic properties and follow specific dependency relationships. Except for the root of the sentence ("said"), each word is dependent on another word. In the example provided above, "The" is dependent on the word "Fed," which in turn is the nominal subject of

[2] Syntactic analysis and dependency parsing are new frontiers in political science research, but a few promising applications show the benefit of retaining dependency relationships between words when analyzing text. Atteveldt et al. (2017), for instance, shows how US and Chinese media portrayed the role of Hamas and Israel in the 2008-9 Gaza war differently; and Vannoni et al. (2021) apply syntactic analysis to a corpus of US state laws to estimate delegation of powers to governors of US states.

Table 1: Dependency Parsing.

Token ID	Token	Part-of-Speech	Head Token ID	Dependency Relation
1	The	DETERMINER	2	determiner
2	Fed	PROPER NOUN	3	nominal subject
3	said	VERB	3	ROOT
4	that	ADPOSITION	9	marker
5	higher	ADJECTIVE	**7**	adjectival modifier
6	interest	NOUN	**7**	compound
7	rates	NOUN	9	nominal subject
8	will	VERB	9	auxiliary
9	slow	VERB	3	clausal complement
10	inflation	NOUN	9	direct object
11	.	PUNCT	3	punctuation

Notes: Output of dependency parsing. Each token is assigned an ID, which describes syntactic dependency relationships between tokens. Token IDs in bold are used as examples in the text.

the verb "to say." The result of syntactic parsing is displayed in Table 1, which reports unique identifiers for each token, the token (i.e., the word), the part-of-speech, the unique ID of the head token (i.e., the "parent" token), and the type of dependency relation. For instance, the head token ID of the words "higher" and "interest" is the token ID 7: "rates", which means that "higher" and "interests" (one an adjective and another a noun) are dependent on the word "rates."

Step 2: Extraction Rules

Once the parser has tagged each token of the sentence, I annotate the sentence based on extraction rules that detect citations, namely instances where someone is reporting (*i*) something said, written, or released by someone, (*ii*) the source of information contained in the citation, and (*iii*) the content of the citation. I create two comprehensive sets of extraction rules: the first captures direct and indirect statements of agencies ("the Fed said", "as said by the Fed") and "according-to" structures ("according to the Fed"); the second captures direct or indirect outputs of agencies ("the Fed's proposal is", "the Fed's proposal to", the "Fed's study suggests").

To match direct and indirect statements, I specify a vector of "say verbs", so the parser marks the lemmatized version of the verb – thereby capturing verbs declined in every form (active or passive) or tense – and its respective subject or in case of an indirect statement, the agent.[3] For

[3]Importantly, I exclude questions and instances where a negation is syntactically dependent on one of these

Table 2: Syntactic Frames.

Extraction Rule	Syntactic Structure	Sentence Example
Statements		
Direct Statement	subject + say verbs	The Fed [say verb] *that higher interest rates will slow inflation.*
Indirect Statement	agent + say verbs	As [say verb] by the Fed, *higher interest rates will slow inflation.*
According-to Structure	accord + object of preposition	According to the Fed, *higher interest rates will slow inflation.*
Outputs		
Direct Nominal Output	output + possession modifier	The Fed's [output] is *to increase interest rates.*
Indirect Nominal Output	output + possession modifier	I fully endorse the Fed's [output] *to increase interest rates.*
Direct Output	output + say verbs	A [output] from the Fed indicates *to increase interest rates.*

Notes: Syntactic frames designed to extract citations from sentences with examples of sentences matching each frame.

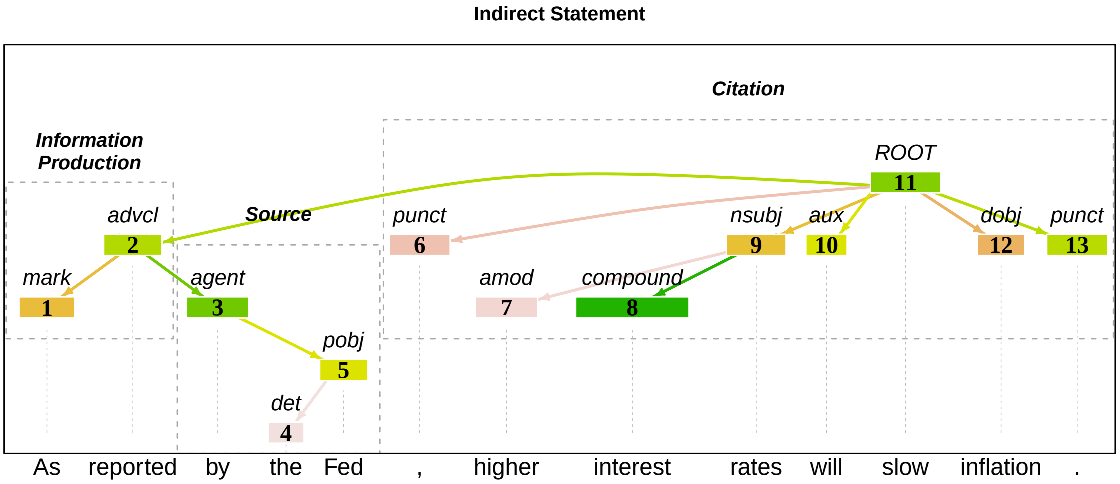
"according-to" structures, the parser detects the token "accord" and the object of the preposition, which will be the source of the information. For direct and indirect nominal outputs, I specify a vector of output-related words for the parser to detect (e.g., "study", "proposal," "recommendation," "suggestion"). Their possessive determiner or the object of prepositions such as "of," "by," or "from" – which mark the owner of the output – is labeled as the source of the output.[4] When labeling the source of the information, I also include cases where individuals affiliated with the agency are producing information. For instance, the algorithm can mark the following direct statement by Representative Proxmire (D-WI) as a citation from the EPA: *"EPA's Deputy Assistant Administrator for Radiation Programs has stated that if all Americans reduced the air infiltration in their homes by 50 percent, the resulting buildup of radon gas could eventually lead to an additional 10,000 to 20,000 cases of lung cancer a year."* Finally, all the tokens that depend on say-verbs, output-related verbs, or according-to structures are labeled as citations. Table 2 reports the precise tokens and syntactic structures used to compile the extraction rules and the sentences in which a legislator could use the information produced by the Fed, with the citation in italics.

I then apply the extraction rules to the tagged sentences. Figure 1 shows the dependency tree

"say verbs" (e.g., "The Fed did not respond to my request" will not be marked).

[4] Say-verbs and output-type words are reported in Section B in the Appendix.

Figure 1: Dependency Tree.



Notes: Dependency tree of an illustrative example where the Fed's information is used in a speech.

of the final output of the syntactic analysis for the example of the indirect statement, one that might seem particularly challenging to extract. Figure C.1 in the Appendix shows the dependency trees of other extraction rules.

To validate the crucial step of citation extraction and ensure the parser can successfully detect instances in which MCs are using a bureaucratic agency as a source of information, I compare the performance of the automatic extraction of citations to human judgment. I extracted 250 random sentences classified as citations by the parser and 250 that are not classified as citations. I then asked an independent coder to decide whether the 500 sentences mentioning the name of an agency were using that agency as a source of information. I find that the human coder and the automatic extraction agree on 82% of the time.

Step 3: Classifying Citations

The final step consists of classifying whether the citation contains policy information. Politicians might use a negative tone to comment on what an agency said (e.g., "The Fed said something completely wrong!"), or they could cite an agency without making any reference to policy (e.g., "The Fed said that in the long term we're all dead.").[5]

[5]In Section F9 in the Appendix, I show that only 5% of citations citing policy information have a negative stance towards the agency, compared to a baseline probability of 15% for all sentences mentioning agencies. The main results are robust to omitting these statements (see Table F.15).

To classify citations based on whether they use policy information, I train and validate several machine learning classifiers that compute the predicted probability of a sentence containing policy information. I follow the standard steps of classification tasks. First, I produce a numerical representation of sentences quoting agencies, which will serve as a matrix of predictors. I use sentence transformers to compute fixed-length sentence vectors (i.e., embeddings), capturing the semantic properties of the text. Second, I build and annotate a training dataset. Given the large number of agencies, policy language can differ for agencies operating in different domains. The classifier's performance, hence, hinges on a large and carefully annotated training dataset. An optimal trade-off between the size of the training data size and the annotation's reliability is now offered by large language models for zero-shot classification. I extract a random sample of 30,000 sentences quoting agencies and annotate them with the assistance of GPT models developed by Open AI. Among the many capabilities of these large language models, a growing literature shows that GPT outperforms crowd-workers for standard annotation tasks (see e.g., Gilardi et al. 2023). The model takes as input a prompt and returns an answer as in an ordinary chat. Finally, I train several machine learning classifiers on the vector representation of 75% of the sentences in the training dataset to achieve good predictions on the 25% of held-out tweets. I select a linear discriminant analysis classifier based on its good performance metrics.[6] Sentences with a predicted probability of policy information ($> .5$) are assigned a score equal to 1 and 0 otherwise. In Section D in the Appendix, I show the prompt and discuss in more detail the classification task and the performance of the classifiers.[7]

As an example of the nuanced classification, the following statement given by Senator Jim Inhofe (R-OK) is assigned a probability of containing policy information of .99: "According to the US Department of Transportation, every 1 billion invested in highway construction creates 47,500 jobs and generates more than $2 billion in economic activity." Conversely, the following statement given by Senator David Vitter (R-LA) is assigned a probability of .22 "Now, I know in a lot of these meetings and conversations, the EPA and Peace Corps say, oh, no, no, no." In fact, even

[6]All performance metrics are above .8 (see Table D.1).

[7]In Section F8 in the Appendix, I present additional validation tests on the reliability of GPT for the annotation task at hand, and I show that the main results of the analysis are robust to using a simple dictionary method that captures quantitative evidence and statistical facts in members' citations of agencies.

though both statements are quoting something said by an agency, only the former includes policy information.

In addition to the validation exercises reported above, in the Appendix, I present two additional tests in support of the predictive validity of the measure. First, show that agencies that are perceived as more skilled are on average cited more often by members of both parties (see Section D.3. Second, I show that there exists a strong correlation between members' public and private demand for agency policy information. Using data on members' informal requests to agencies from Lowande (2018), I show that when members send policy requests privately to agencies, they are also 4-7 percentage points more likely to cite policy information from the same agency in Congress (see Section D.2).

Despite the numerous validation tests presented above, there are two important limitations of the measure that deserve attention. First, this measurement strategy does not capture implicit ways legislators could use the information produced by bureaucracies. By anchoring the citation to the name of the agency (or individuals whose affiliation with the agency appears in the text), the proposed method is only able to capture explicit ways of using bureaucratic information. Second, this strategy is silent about the reasons why members of Congress use bureaucratic information. MCs have different motives for using bureaucratic information, and this paper represents the first general attempt at detecting the main conditions under which politicians decide to do so.

Data & Descriptive Facts

The empirical strategy is based on various sources of unstructured data related to federal agencies, congressional speeches, and bureaucrats' campaign donations. In this section, I describe each source in detail.

Agencies and Speeches

First, I compile a comprehensive list of agencies, including executive departments, bureaus, independent agencies, boards and commissions from government websites and existing datasets on

Table 3: Speeches Data: Descriptive Statistics.

Descriptive Statistics	**Floor**	**Committee**	**Total**
Years covered	1994-2022	1980-2022	
Speeches			
Total	1,634,602	6,868,156	8,502,758
Mentioning agencies	258,924 (15.8%)	657,050 (9.6%)	915,974 (10.8%)
Sentences			
Mentioning agencies	696,096	1,191,154	1,887,250
Citing agencies	46,512 (6.7%)	50,209 (4.2%)	96,721 (5.1%)
Citing policy information	23,533 (3.4%)	17,862 (1.5%)	41,395 (2.2%)

Notes: Descriptive statistics of speeches, sentences, and citations of bureaucratic agencies.

agencies' attributes, totalling 317 bureaucratic bodies.[8]

Second, I assemble a corpus of 8.5 million speeches by 2,098 unique legislators: 1,6 million given in the floor and 6,8 million in committees. I scraped floor (1994-2022) and committee (2010-2022) speeches from the digitized version of the *Congressional Record* and obtained transcripts of congressional committee sessions for the period 1980-2009 from ProQuest.[9] After replacing the various ways in which agencies are mentioned with a standardized name, I subset all speeches mentioning at least one agency.[10]

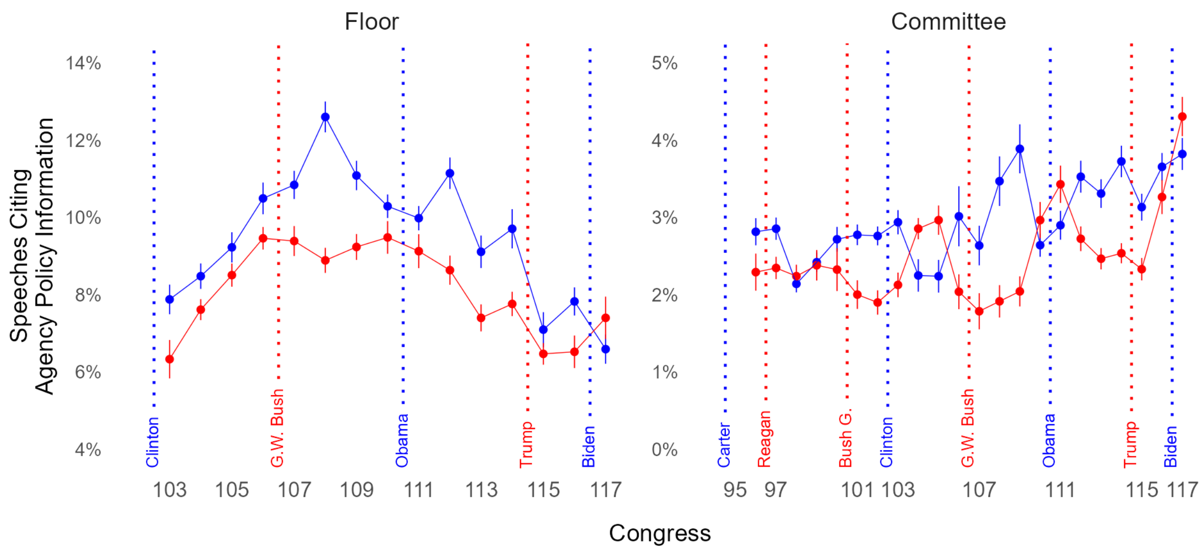
A total of 915,974 speeches mention the name of at least one agency (10.8%). 55% of speeches are given by Democrats and 44.6% by Republicans. I split these speeches into sentences and keep only the sentences containing the name of an agency. I detect 1,877,250 sentences mentioning bureaucracies, on which I apply the three measurement steps described above: 5.1% of the sentences are citations of agencies, and 2.2% are citations containing policy information. The conditional probability of reporting policy information when quoting agencies is .47, and it is minimally higher for Democrats (.48) than Republicans (.46). Descriptive statistics concerning the sample of speeches, sentences, and citations are reported in Table 3.

[8]The datasets I used to compile agency lists are those produced by Selin (2015), Richardson et al. (2018), Bertelli et al. (2013), Chen & Johnson (2015), which I integrate with lists of agencies reported on institutional government website: `https://www.usa.gov/branches-of-government`.

[9]Information about the quality of the transcripts and the speech parsing steps is reported in Section A of the Appendix.

[10]When processing the text of the speech, I ensure that state agencies and statutes carrying the name of agencies are not included.

Figure 2: Share of Speeches Using Agency Policy Information.



Notes: Share of speeches by members of the Democratic and Republican parties citing policy information produced by agencies in floor and committee speeches separately. Blue for Democrats, red for Republicans. Each dot represents the share of speeches in any given Congress. Dashed vertical lines mark presidential transitions.

Stylized Facts on Members Citations of Agency Policy Information

The granularity of the data allows the monitoring of individual members' reliance on information produced by every agency in the dataset. For instance, of the 2,081 unique legislators who mention bureaucratic agencies in the period under study, Christopher H. Shays (R-CT) cites agency policy information the most – 2,108 times – during the 15 Congresses he served. Among these mentions, Representative Shays cited information produced by 128 unique agencies. The agency with the largest number of policy citations is the EPA, with 177 mentions of policy information.

At a more aggregate level, two sets of stylized facts about partisan differences and agency-specific differences in citing bureaucratic information in Congress are worth noting.[11] Figure 2 shows the share of speeches citing policy information on the floor and committee and over time.

First, while Democrats cite agency policy information more than Republicans, the difference – though statistically significant – is relatively small (+1.4% in the floor and +0.3% in committees). Second, president co-partisans are not more likely to cite bureaucratic information than MCs out-partisans of the president, which suggests that MCs utilize evidence from bureaucratic agencies

[11] Table D.2 in the Appendix reports descriptive statistics for each agency.

to scrutinize the administration's actions. Finally, the most striking trend is the decline in MCs' reliance on bureaucratic information in floor speeches, with the share of speeches citing agency information sharply decreasing from the second Bush presidency. This downward trend might be the result of the high level of partisan polarization on the floor, where MCs increasingly care about scoring political points rather than engaging with policy information. This interpretation is consistent with the contrasting trend in committees, where MCs' reliance on bureaucratic information is more stable and, though fluctuating over time, did not experience as drastic a decrease as was observed on the floor.

As for agency-specific differences, there are noticeable partisan distinctions in the use of information produced by specific agencies, consistent with findings from the literature on partisan issue ownership (Egan 2013). Perhaps unsurprisingly, Democrats cite policy information from the EPA 52% more often than Republicans. Similarly, since 2019, Democrats have been more likely than Republicans to use the information produced by the Centers for Disease Control and Prevention and the FDA – agencies that played pivotal roles in the government's response to the COVID-19 pandemic. Conversely, Republicans exhibited a 54% higher use of policy evidence produced by enforcement agencies compared to Democrats.[12]

Finally, to benchmark how often MCs rely on the bureaucracy as a source of information in their speeches, I compare the number of times they cite bureaucracies as a source of information to the number of times they use alternative sources. Specifically, I assemble a comprehensive list of the country's major newspapers and think tanks (45 newspapers and 120 think tanks). While the average agency is cited 130 times (41,395 policy citations / 317 agencies), the average number of policy citations from newspapers is 73 (3,329 / 45), and from think tanks is 25 (3,023 / 120). Bureaucracies are cited on average 2-5 times more often than newspapers and major US think tanks, suggesting that, even comparatively, bureaucratic bodies are key players in providing information to Congress.[13]

[12]These bureaucracies are the Drug Enforcement Administration, Office of National Drug Control Policy, Immigration and Customs Enforcement, Customs and Border Protection.

[13]The lists of newspapers and think tanks were assembled from Wikipedia pages "/List_of_think_tanks_in_the_United_States" and "/List_of_newspapers_in_the_United_States" and are available in the replication package accompanying this paper.

Agency Ideology

Despite agency ideology playing a prominent role in most theories of bureaucratic behavior, measurement remains challenging. There are two approaches to measuring agency ideology, with important conceptual differences. First, scholars have produced estimates of the ideological leaning of the "mission" of the agency. Richardson et al. (2018) offer the most recent example of this approach. The authors use a survey of bureaucrats to estimate perceptions of policy views of agencies across Democratic and Republican administrations. While this approach taps the expertise of experienced bureaucrats, it does not allow agency ideology to change over time.[14]

The second approach is centered on a dynamic concept of agency ideology. Presidents have significant influence over appointing top-tier bureaucrats, and turnover in leadership positions has downstream effects on the actions and goals of agencies, too (Spenkuch et al. 2023). Estimates of agency ideology are computed by aggregating individual bureaucrats' preferences and behaviors. The most commonly used data on dynamic agency ideology is the one produced by Chen & Johnson (2015), who aggregate campaign donations of bureaucrats across 79 agencies and five presidential terms, from the first term of Clinton to the first term of Obama.[15] The resulting estimates are weighted averages of the DW-NOMINATE scores of the recipient of the donation.

All measures have limitations. However, given the main goal of this paper to estimate the effect of agency ideology on members' reliance on policy information produced by agencies, dynamic measures of ideology are preferable, as they allow the observation of how much the consumption of policy information increases or decreases with changes in within-agency ideology. Therefore, I compute dynamic estimates of agency ideology following an approach similar to that of Chen & Johnson (2015), yet covering a more extensive period and a larger sample of agencies. I extract all campaign contributions from bureaucrats and aggregate the CF-score of bureaucratic contributors, giving greater weight to bureaucrats who donate more. The underlying assumption of this weighting strategy – the same used by Chen & Johnson (2015) – is that more influential bureaucrats, who exert more control over the agency's activities, have higher salaries and are more likely to

[14] A similar approach is used by Clinton & Lewis (2008) and Clinton et al. (2012).

[15] This dataset has been widely used in political science to study the political control of the bureaucracy (Lowande 2018), strategic appointments (Moore 2018), career paths of bureaucrats (Bolton et al. 2020), and rule-making (Ellig & Conover 2014, Potter 2019).

make larger donations. To support this assumption with evidence, I matched data on appointees' donations from Bonica et al. (2015) to the database of federal employees' donations and found that the average donation of appointees is \$932, whereas the average donation in the whole dataset is \$53.

I construct the estimates in five steps, which are described and validated in greater detail in Section E in the Appendix. First, I download donation data from the Database on Ideology, Money in Politics, and Elections (DIME) (Bonica 2023) for every election cycle from 1990 to 2022. The DIME contains the universe of itemized donations recorded by the Federal Election Commission, with information on the amount donated, the contributor's name and employer, the contributor's CF-score, and several identifiers for the recipient of the donation. Second, since the goal is to extract donations made by federal employees, I extract all the employer entries for a total of more than 9.7 million unique employers. I use a deep learning model for record linkage to assign to each employer the agency that best matches the employer string (Arora & Dell 2023). I successfully match 13,596 unique employers to 302 agencies. Third, I extract the 6,637,563 contributions from 569,621 unique individuals employed by a federal agency. I finally obtain a repeated cross-section of bureaucrats for every agency-presidency pair. Let x_{ij} be the fixed CF-score of bureaucrat i employed by agency j, and N the set of bureaucrats employed by agency j donating money during presidency t. The goal is to create a dynamic measure of agency ideology x_{jt} by aggregating the preferences of the N individual bureaucrats at the agency-presidency level. To aggregate the CF-score of individual bureaucrats, I follow the approach of Chen & Johnson (2015), who assign greater weight to bureaucrats contributing more money, and in the Appendix I show that the main results are robust to alternative aggregation strategies (Section F7). I group contributors by presidency and take the total amount ϕ_{ijt} donated by bureaucrat i working in agency j during presidency t. The dynamic ideology of each agency is equal to the average of bureaucrats' CF-score weighted by the total amount they donated:

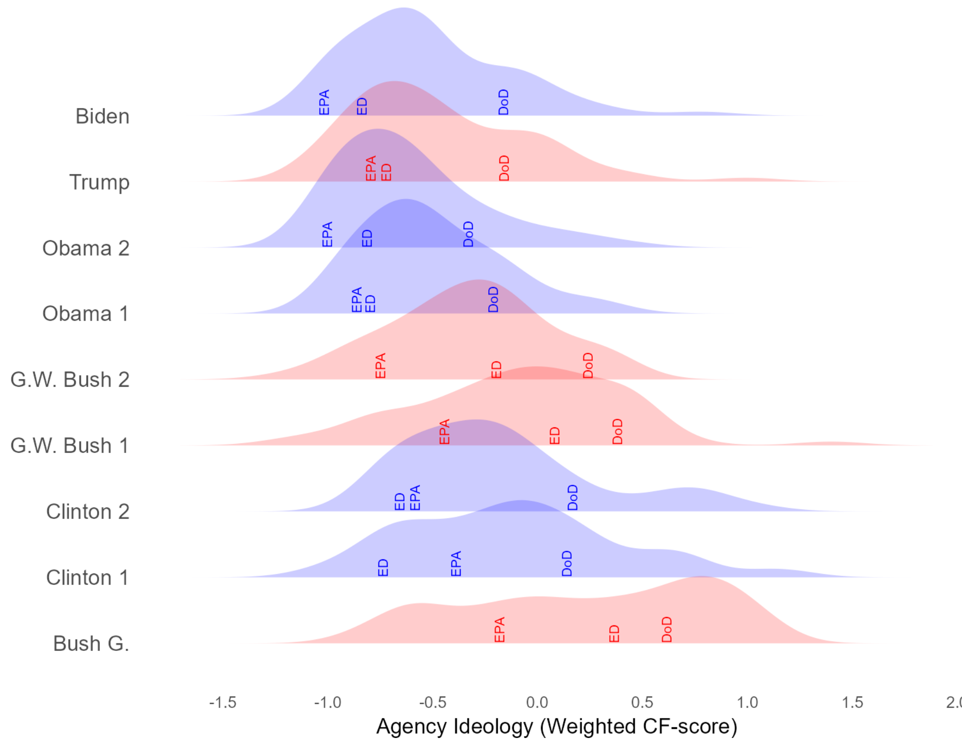
$$x_{jt} = \frac{\sum_{i=1}^{n} x_{ijt}\phi_{ijt}}{\sum_{i=1}^{n} \phi_{ijt}} \tag{1}$$

The final dataset includes 2,230 observations covering 302 agencies over a maximum of 9 presidential administrations. For each agency, I compute the average number of donors across presidencies. The median number of donors by agency-presidency is 29, and the mean is 296. This high variance in the number of donors is because the list of agencies includes very small organizations, all executive departments, and large independent agencies. When comparing the number of donors with the total number of employees in each agency, I find that, on average, 9.3% of employees are donors (the median is 3.3%). In the empirical analysis, I focus on the 64 agencies with an average of at least 200 donors per presidency and show that the results are robust to any alternative thresholds from 50 to 500 (see Section F3). Furthermore, to further validate the measure, I show that the estimates positively correlate with those produced by Chen & Johnson (2015) (Figure E.3) and presidential transitions have a sizable effect on the ideology of agencies (see Table E.6).

In Figure 3, I display the distribution of agency ideology across presidential administrations. The figure shows that agency ideology shifts when a new president takes over the White House. Interestingly, the distribution has higher variance until the first G.W. Bush presidency, and it becomes increasingly concentrated around the median from the first Obama presidency. The figure also reports how the ideology scores of four representative agencies generally considered very liberal or conservative change across presidencies. The Department of Education and the Environmental Protection Agency, often portrayed as a liberal agency, are always to the left of the Department of Defense and the Department of Justice, generally perceived to be conservative departments. However, not all agencies shift their ideology to the same degree. The EPA, for instance, moves to the right under Trump, whereas the ED does not.

These new estimates have four key advantages over existing ones. First, they cover the largest number of agencies and presidential terms possible. Second, they allow to observe how the ideology of agencies changes across presidential administrations. Third, using the CF-scores of both contributors and recipients of donations, the ideology of agencies is estimated on the same scale as that of members of Congress, allowing to produce valid distance metrics. Finally, I can show how the empirical results change when sequentially excluding agencies with a few donors to check whether the findings are sensitive to the sample size of the pool of bureaucrats who make

Figure 3: Distribution of Agency Ideology Across Presidential Terms.

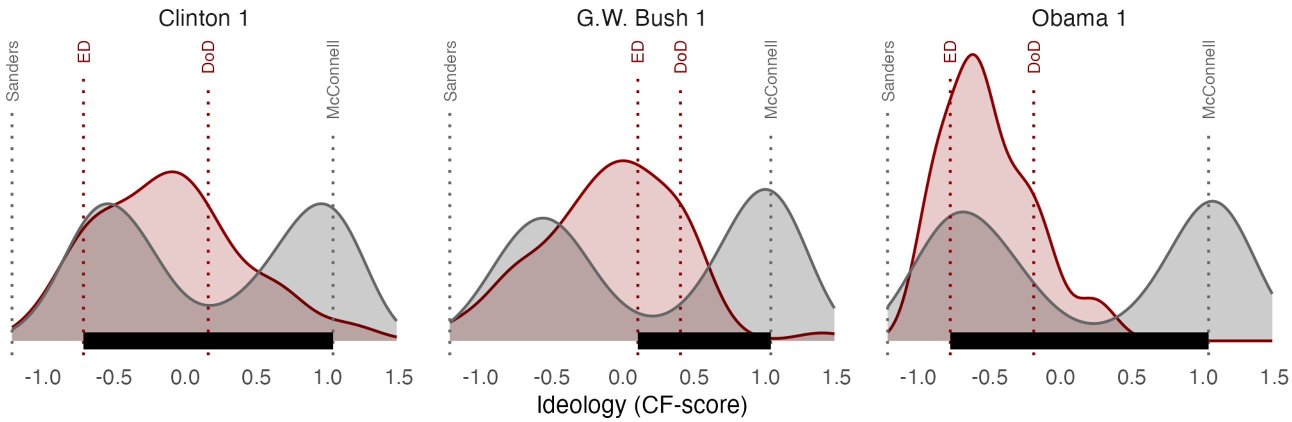


Notes: Distribution of the ideology scores of bureaucratic agencies across different presidential administrations. The plot includes the 64 agencies with an average of at least 200 donors by the presidency and locates the ideology of four agencies: the Department of Education (ED), the Environmental Protection Agency (EPA), the Department of Justice (DOJ), and the Department of Defense (DOD).

political donations.

To measure the ideological distance, I compute the absolute value of the difference between agency and members' estimates. To measure members' ideology, I use the candidate/recipients' CF-score. Figure 4 shows the ideological distribution of agencies and members of Congress across the first Clinton, G.W. Bush, and Obama presidencies, with the thick dark bar on the horizontal axis showing how I construct the measure of ideological distance between MCs and bureaucracies. Dashed vertical lines locate two agencies and two legislators on the ideological spectrum. The Department of Defense (DOD) is on the right-hand side of the distribution. In contrast, the Department of Education (ED) – an agency generally perceived to be very liberal – is on the left-hand side. However, as shown in Figure 3, the ED under Bush shifts to the right, whereas both the DOD and ED move to the left under Obama. As a result, the ideological distance between Senator Mitch McConnel (R-KY) and the ED is shrinking under the Bush presidency compared to the previous and following democratic presidencies.

Figure 4: Agency and MCs Ideology.



Notes: Distribution of the ideology scores of agencies (in red) and members of Congress (in gray) across the first presidencies of Clinton, G.W. Bush, and Obama. Thick bars represent the ideological distance between Senator Mitch McConnell and the Department of Education.

Empirical Strategy

I perform two main analyses to test the predictions. First, I estimate a set of dyadic fixed-effects models leveraging within-agency-by-member variation in ideological distance. Second, to test the effect of agency independence, I compare legislators' citations of policy information from independent and non-independent agencies operating within the same policy domain. To strengthen the analysis of agency independence, I exploit a 2020 Supreme Court decision that curtailed the independence of the Consumer Financial Protection Bureau to estimate how the decision affected members' reliance on the CFPB's information.

Research Design: Agency Ideology

To estimate the effect of MC-agency ideological distance on members' reliance on policy information produced by agencies, I build a dyadic dataset at the MC-agency-congress-venue level, with the venue representing the floor and committees. For instance, I can track how often a member cites policy information produced by the EPA on the floor and committees over time. To minimize selection effects, I include all MC-congress-venue tuples with at least one available speech. This definition of the sample allows for building a dataset of 16,853 MC-congress-venue combinations. Each member is observed on the floor and on committees for any given congress, provided she gave at least one speech in either venue. I then cross-join the MC-congress-venue matrix with a dataset of 4,817 agency-congress observations and obtain a dataset of 4,081,895 MC-congress-venue-agency observations. When subsetting the dataset to agencies with an average of ≥ 200 donors by presidential term, the size of the dataset moves from 3,377,013 to 882,775 observations, including 64 unique agencies, 1,781 unique members of Congress, and 110,670 unique member-agency dyads. Notably, the dyadic structure of the data allows the estimation of the effect of ideology on members' reliance on agency expertise at the extensive margins. To do so, I report zeros when the member does not cite policy information from any given agency in any congress-venue combination. While this approach makes the data relatively sparse, it is suitable for analyzing a scenario where even not citing agencies might be the product of ideological differences.

My most stringent specification leverages change within legislator-agency pairs while holding fixed agency characteristics that change across Congresses and conditioning on a set of MC-level time-changing covariates. Specifically, I estimate regressions of the form:

$$y_{ijtv} = \gamma_{ij} + \delta_{jt} + \beta \text{Ideological Distance}_{ijt} + \zeta v + \psi' x_{it} + u_{ijt} \tag{2}$$

where y_{ijtv} is the number of sentences citing agency policy information for member i agency j in congress t and venue v (floor or committees), γ_{it} are MC-agency fixed effects, δ_{jt} are agency-by-congress fixed effects, v is an indicator for the venue (floor or committees), and x_{it} a vector of MC-covariates, including whether the member is chair of a committee or subcommittee, how effective she is at passing legislation, whether she is majority or minority leader, and the number of speeches given in any given congress-venue.[16] Ideological Distance$_{ijt}$ is the primary variable of interest, capturing the time-changing absolute value of the distance between the ideology of agency j and member i in congress t. Identifying variation originates from changes in agency ideology over time (since members' ideology is fixed over time) and from the cross-sectional variation within any MC-congress-venue tuple. Crucially, γ_{ij} enables accounting for legislators' and parties' fixed characteristics, such as ability, attention to different agencies or policy domains, and issue ownership. This fixed effect also accounts for agency characteristics such as statutory features and policy areas. δ_{jt} captures changes affecting agencies over time, allowing to hold constant time-changing

[16] Data for member-specific covariates are from Volden & Wiseman (2014, 2018).

characteristics of agencies and other shocks that could affect the use of information produced by agencies for every legislator in any given congress. Finally, the vector x_{it} accounts for member-specific characteristics that can be correlated with ideology and members' reliance on bureaucratic expertise. β identifies the effect of ideological distance on the number of times members cite policy information produced by bureaucracy.

Results

Table 4 presents the main regression results. Ideological distance reduces members' reliance on policy information produced by agencies. The columns vary the specification to probe the robustness of the results. Columns (1) reports three sets of member, congress, and agency fixed effects, Column (2) adds MC covariates, Column (3) adds dyads (MC-agency) fixed effects, and Column (5) estimates Equation 2 and represents the preferred and most conservative specification with MC-agency and agency-congress fixed effects. Reassuringly, the inclusion of these fixed effects does not significantly affect the estimates in terms of sign and precision, although the size of the effect doubles from the simple specification of Column (1) to the most conservative of Column (5).

The dependent variable is the number of times each member cites policy information from any given agency over time. Members do not cite information from every agency, hence the average number of policy citations is rather low and equal to .028. The maximum number of policy citations is from the Democratic Representative Gary L. Ackerman (D-NY), who cites policy information produced by the EPA 58 times in the 112th Congress.

On average, ideological distance decreases the use of policy information produced by the bureaucracy, even when holding everything constant at the MC-agency and MC-congress level. A one-unit increase in ideological distance, which is approximately equal to 1.6 standard deviations of ideological distance in the analysis dataset, decreases members' reliance on policy information by $-.005$, which is as large as 18% compared to the sample mean. The effect of ideological distance is precisely estimated across specifications.

These results are obtained by pooling speeches given on the floor and committees while including the respective indicator variable as a control. In Table F.9 in the Appendix, I show

Table 4: Ideological Distance and MCs' Use of Bureaucratic Information.

	# Sentences Citing Policy Information			
	(1)	(2)	(3)	(4)
Ideological Distance	-0.0028*	-0.0045***	-0.0036*	-0.0049**
	(0.0011)	(0.0011)	(0.0015)	(0.0015)
Floor/Committee	✓	✓	✓	✓
MC Covariates		✓	✓	✓
Mean DV	0.03	0.03	0.03	0.03
R^2	0.069	0.101	0.293	0.297
Observations	870,419	870,419	870,419	870,419
Agency FE	✓	✓		
MC FE	✓	✓		
Congress FE	✓	✓	✓	
MC-Agency FE			✓	✓
Agency-Congress FE				✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Sample of 64 agencies with an average number of donors by presidency greater than 200. The outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Columns (1) to (3) use data from speeches on the floor and in committees. Columns (4) and (5) use data from floor and committee speeches separately. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

that the effect of ideological distance is two times larger on the floor compared to committees. This difference might be due to the fact that floor speeches are more public-facing and theatrical, where members play by party rules. In this scenario, where members' exposure to constituents and donors is greater, electoral dynamics might augment the effect of preference disagreement.

Robustness Tests

The robustness of the results crucially hinges on the validity of the measurement of the main variables. In Section F in the Appendix, I present several robustness tests. First, I show that the results are robust when using two different transformations of the dependent variable. I replace the absolute frequency of policy citations with a dummy variable equal to 1 if the $y_{ijtv} > 0$ and 0 otherwise. Similarly, to reduce the influence of outliers, I use the $\log(1 + y_{ijtv})$ transformation of the number of sentences citing agency policy information (see Table F.7). Second, I show that the results are not sensitive to the number of agencies implicitly included or excluded from the analysis based on the average number of donors used to compute agency ideology. In Figure F.5, I show that the effect of ideological distance is negative and statistically significant when using agencies with an average number of donors by presidency ranging from 50 to 500. Third, I find that the ideological distance between members and agencies decreases members' reliance on information both on the floor and committees (see Table F.9) and in the House and Senate (see Table F.10).

Fourth, I show that the results are not driven by the inclusion/exclusion of one single agency by sequentially removing from the analysis dataset each MC-agency$_j$ observations with j being sequentially each one of the 64 agencies in the dataset (Figure F.6). Fifth, I show that the results are robust to using alternative approaches to aggregating bureaucrats' donations (see Table F.11). Sixth, I find that statements quoting agencies strongly correlate with a positive or neutral stance, and the results hold when excluding statements that cite agencies while criticizing them (see Table F.15). Seventh, in light of some recent criticisms of using proprietary language models for classifying political text (see, e.g., Barrie et al. 2024), I find similar results when using a dictionary method to scale citations based on the presence of quantitative evidence and statistical facts. Finally, to show that ideological distance affects members' reliance on policy information and not members' attention to bureaucracies in general, I show that there is no precisely estimated association between ideological distance and the number of times members mention agencies (see Table F.16). These results prove that members have the same probability of reporting what is said by agencies that are both ideologically close and distant and that ideology matters only for their decision to cite *policy information* produced by agencies.

Research Design: Agency Independence

In the second analysis, I test whether agency independence affects members' reliance on bureaucratic policy information. Identifying the effect of fixed characteristics of agencies is challenging, for statutory features of agency designs seldom change over time and are, by design, endogenous. Agency independence is plausibly highly correlated with other characteristics of the agency, such as the policy area in which the agency operates. Therefore, I perform two tests. First, I exploit cross-sectional variation in agency design and compare the policy citation rate of independent and non-independent agencies operating in the same policy area. Second, I leverage *within-agency* variation in independence as a result of the 2020 Supreme Court ruling that lifted a key source of independence of the Consumer Financial Protection Bureau.

Between-agency Variation in Independence

Agency independence has been defined and measured in multiple ways. The simplest definition of independence indicates whether the agency is located outside the Executive Office of the President (EOP) and executive departments. However, scholars have proposed more substantive definitions of independence, such as the extent to which specific statutory features of agencies make them more or less responsive to political principals. Along this tradition of work, the most frequently used measures of independence are those produced by Selin (2015), who introduces two indicators of agency structural independence based on the limits imposed on political principals with respect to the appointment of key agency decision-makers and the political review of agency policy.

I use these two alternative measures to estimate the relationship between agency independence and members' citation of policy information. I build a panel dataset at the agency j and congress t level. The outcome variable is the total number of sentences citing policy information that agency j received in Congress t. Legislators might cite independent agencies more than non-independent agencies because of their specific policy area. To hold the policy area fixed, I assign each agency to one of the 20 policy areas from the Comparative Agendas Project (CAP). I perform this classification with a supervised classifier specifically pre-trained on CAP policy documents by Dickson & Hobolt (2024).[17] The model takes as input the name of the agency and predicts the policy area. For example, the Agricultural Marketing Service is assigned the policy area "agriculture", and the Commodities Futures Trading Commission is assigned the policy area "domestic commerce". Fifteen of the twenty policy areas have at least one independent agency. I then estimate equations of the form:

$$y_{jpt} = \mu_{pt} + \eta \text{Independence}_j + \theta x_{jt} + \epsilon_{jpt} \tag{3}$$

where y_{jpt} is the number of sentences citing policy information from agency j assigned to

[17]The model is publicly available at `https://huggingface.co/z-dickson/CAP_multilingual`.

Table 5: Effect of Agency Independence.

	# Senteces Citing Policy Information					
Measure of Independence:	Dummy		Agency Policy		Decision Makers	
	(1)	(2)	(3)	(4)	(5)	(6)
Independence	3.043***	2.385*	1.458***	1.648***	0.533**	-0.110
	(0.781)	(1.063)	(0.291)	(0.427)	(0.174)	(0.264)
Agency Ideology		-1.441***		-1.833***		-1.709***
		(0.428)		(0.539)		(0.515)
Mean DV	6.790	2.570	1.992	2.603	1.992	2.603
R^2	0.073	0.095	0.075	0.097	0.071	0.093
Observations	5,967	3,930	5,371	3,533	5,371	3,533
Congress-CAP Policy Area FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates. Robust standard errors in parenthesis. The outcome variable is the number of sentences agencies' policy information was cited in Congress. Columns 1-2: independence is a dichotomous variable equal to 1 if the agency is located outside the EOP or departments. Columns 3-4: agency independence is a continuous and standardized measure based on the limitations on political control of agency policy. Columns 5-6: agency independence is a continuous and standardized measure based on the limitations on political control of agency decision-makers. Agency ideology is standardized. Changes in the number of observations are due to the fact that not all agencies have a measure of ideology or independence. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

policy area p during Congress t, μ_{pt} are policy area-by-congress fixed effects, capturing the dynamic salience of policy areas and presidential priorities, and x_{jt} is the time-changing measure of agency ideology. Independence is one of the three measures discussed above: a dichotomous variable and the two continuous measures produced by Selin (2015). All these measures are fixed over time. By comparing agencies within the same policy area and Congress, the specification allows for the comparison of agencies that are "competitors" in the provision of information in their specific policy area.

Table 5 presents the results. On average, policy information produced by independent agencies is cited more often by members of Congress. The effects are robust to different measures of independence, except for the decision-makers indicator, which becomes very noisy when conditioning the estimate on agency ideology. The continuous indicators are standardized, hence representing the increase in the number of policy citations resulting from a +1 standard deviation increase in agency independence. The effect size is quite large if compared to the mean in the data, ranging from 60-90% the mean, depending on the measure of independence utilized.

To strengthen these correlational results, in the following section, I analyze a case when one agency experienced a sudden shock to its level of independence. This allows me to compare how members of Congress reacted to the decrease in agency citations after the shock compared to a counterfactual of other unaffected agencies.

Within-Agency Variation in Independence: ***Seila Law v CFPB***

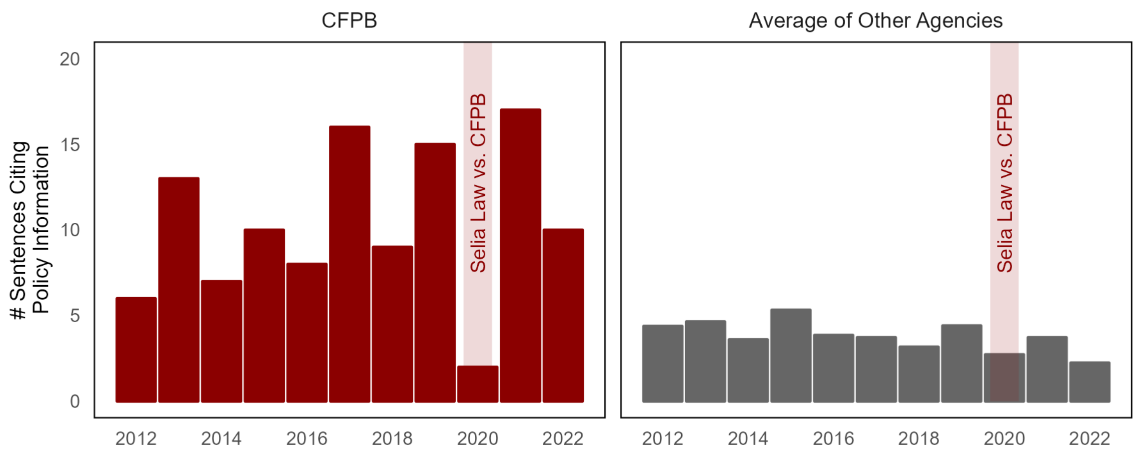
In response to the 2008 financial crisis, Congress and former President Barack Obama established the CFPB, an independent federal agency whose goal was to watch over predatory financial services practices. As provided by its statute – and unlike other independent agencies – the CFPB was governed by one director who could be removed by the president from her position only for cause, that is "inefficiency, neglect of duty, or malfeasance in office." Political appointees serving as directors were protected from at-will dismissal and enjoyed a significant degree of independence in managing the agency. This guarantee of independence was nonetheless removed with the 2020 *Seila Law v. Consumer Financial Protection Bureau*, where the Supreme Court ruled that the restrictions on the removal of the CFPB director are unconstitutional.[18]

This sudden shock to the CFPB's independence had a striking effect on members' citations of policy information produced by the bureau. Figure 5 shows on the left panel how often members cited policy information produced by the CFPB over time. The right panel shows the average for all the other agencies. Members' reliance on the CFPB's information plummeted in 2020, the year of the Court's decision. The decrease is short-lived and does not extend to the first years of the Biden presidency.

This sharp decrease might hide partisan differences in members' reactions to the decision. When the CFPB suddenly lost its shield against presidential control, members from the party different from that of the president might have experienced a greater adjustment compared to members co-partisan with the president. A closer look at former President Trump's approach to the CFPB reveals that Trump was often accused of using a bureau initially tasked with a

[18]The dispute began when Seila Law, a law firm that provides debt-relief services to consumers, was under investigation by the CFPB for possible violations of telemarketing sales rules. Seila Law challenged the CFPB's powers to obtain documents from the firm, arguing that the bureau's organization was unconstitutional due to its configuration of one director removable only "for cause". Instead, Seila Law argued, the director should be removable "at will" by the president – that is, for any reason.

Figure 5: Members' Citation of CFPB Policy Information.



Notes: Members' citations of policy information produced by the CFPB (left panel) and members' average use of information produced by other agencies over time. Red vertical bar marks the year of the Court's ruling.

pro-consumer mission "to serve the wishes of the most powerful financial companies in America" (Forbes 2020).

To strengthen the evidence provided visually, I leverage the timing of the Court's decision as a shock to the independence of the CFPB and compare how members' reliance on the CFPB's information changed as a result. I compare members' citations of the CFPB's policy information over three time periods: under Trump and before the decision, under Trump but after the decision, and under Biden and after the decision. The structure of the analysis dataset is the same as the one used in the previous test for agency ideology, with one exception: instead of using Congresses as time indicators, I use the three time periods described above, namely Trump-pre-decision, Trump-post-decision, and Biden-post-decision. The flavor of this specification is that of a differences-in-differences design, where treated units (member-CFPB pairs) are compared to control units (members-other agencies pairs) across the three time periods. Each observation captures the number of times member i uses policy information from agency j in period t. For this test, I pool all speeches given on the floor and on committees and estimate equations of the form:

$$y_{ijt} = \gamma_i + \theta_j + \delta_t + \tau \mathbf{1}\{t \in \text{Post}, j = \text{CFPB}\} + u_{ijt} \quad (4)$$

where $\mathbf{1}\{t \in \text{Post}, j = \text{CFPB}\}$ is an indicator equal to one for member-agency dyads where

Table 6: Effect of *Seila v CFPB* Case on Members' Citation of Policy Information.

	# Sentences Citing Policy Information		
Members:	All (1)	REP (2)	DEM (3)
CFPB × Post (Trump)	-0.017*	-0.010	-0.025*
	(0.008)	(0.010)	(0.011)
CFPB × Post (Biden)	0.027	-0.008	0.061*
	(0.017)	(0.013)	(0.030)
Mean DV	0.006	0.005	0.007
R^2	0.034	0.030	0.039
Observations	186,669	89,284	96,273
MC FE	✓	✓	✓
Time FE	✓	✓	✓
Agency FE	✓	✓	✓

Notes: OLS estimates. SE clustered by member-agency dyad. The outcome variable is the number of sentences citing agency policy information for each member-agency-time period combination. CFPB is equal to 1 for member-CFPB pairs. Post (Trump) equals 1 for observations during the Trump presidency and after the ruling. Post (Biden) equals 1 for observations during the Biden presidency after the ruling. The reference time period is Pre (Trump), for observations during the Trump presidency before the ruling. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

agency = CFPB, and 0 otherwise. γ_i, θ_j, and δ_t are member, agency, and time fixed effects. τ estimates the change in members' reliance on information produced by the CFPB in two time periods: under the Trump-post decision and under the Biden-post decision, compared to the period under the Trump-pre decision. The key identifying assumption is that, absent the Supreme Court ruling, members' reliance on the CFPB's information would have experienced parallel trends to members' reliance on information produced by other agencies.[19]

Table 6 presents the results. Column (1) shows the results of a regression that includes all members. Columns (2) and (3) focus on Republicans and Democrats separately. On average, after the Court's ruling and in the last months of the Trump presidency, members are less likely to use information produced by the CFPB. The effect is precisely estimated and statistically significant

[19] In Section G4 in the Appendix, I indirectly test the parallel trends assumption by estimating Equation 4 with placebo post-treatment indicators. Furthermore, Table G.19 shows the results are robust to replacing the count outcome variable with a dichotomous and log-transformed measure. Finally, as evidence that the CFPB did not reduce the supply of information during the year of the ruling, Figure G.8 shows that the CFPB's enforcement activities are even more vigorous in the year of the ruling.

at the 95% level. The number of sentences citing agency policy information from the bureau decreases by 0.017, which represents a reduction of approximately three times the sample mean. The estimates in Columns (2) and (3) indicate that these effects are more precisely estimated in the sub-sample of Democrats and are limited to the Trump term. Democrats' reliance on the CFPB's policy expertise rebounds to a level higher than in the pre-decision period when Biden is president. This suggests that when agencies are stripped of their independence, members who are out-partisan with the president react the most. There appears to be an interaction effect between ideology and independence. With the CFPB director now exposed to "at-will" dismissal, Democrats became more skeptical of an agency that could now be pressured by presidential directives. However, when this risk is mitigated with the election of a Democratic president, Democrats again use the expertise of the agency.[20]

Discussion and Conclusions

For almost a century, scholars have stressed the foundational role of expertise in the study of bureaucratic-legislative interactions, which has been described as the "lifeblood of executive branch action" (Gailmard & Patty 2013, 1). Federal agencies in the United States produce a tremendous amount of information, from policy reports to press releases and regulatory documents. The Environmental Protection Agency alone posted more than 6,000 datasets on `data.gov`, the official government repository of open data produced by federal agencies. The flow of information between elected officials and unelected bureaucrats is of paramount importance not only for effective governance but also for the accountability of unelected bureaucrats, who hold considerable influence over policy. Despite the importance of bureaucratic information in the US system of government, empirical investigations of whether members of Congress make use of such a vast amount of information are scarce. I connect novel empirical evidence to a question that has important implications for the scholarly understanding of inter-branch interactions in the US system of separation of power.

This paper represents the first attempt to bring observational evidence to a key question in the scholarship on the bureaucracy, Congress, and policymaking. Building on canonical models

[20] I find a similar interaction effect between independence and ideology when estimating the effect of ideological distance among independent and non-independent agencies separately (see Table G.17 in the Appendix).

of information and expertise acquisition, I find that members of Congress prioritize information produced by ideologically similar agencies. Moreover, agency independence, by fostering specialization and a reputation for expertise, increases members' citations of bureaucratic information. While most of the empirical literature on inter- and intra-branch relationships in the United States focuses on the extent to which the President and Congress can control drifting agencies, in this paper, I shed light on an empirically neglected dynamic: this paper highlights an often-overlooked dynamic: legislators frequently depend on the information and expertise of bureaucratic agencies when engaging in various political activities.

This paper also opens new research avenues within the study of American political institutions. First, while the findings highlight general patterns in members' decisions to cite bureaucratic information, future work could examine how members' strategic selection of information affects legislative outcomes. Second, the novel measurement strategy introduced in this paper allows scholars to study the market for policy-relevant information and the competitive nature that exists between different sources of information. Other influential organizations, such as think tanks or interest groups, often compete with bureaucracies in providing credible and timely policy information and can bring very different perspectives inside Congress. Third, scholars of the internal organization of Congress could examine how changes in congressional staff capacity and composition affect legislators' reliance on policy expertise and the diversity of policy information invoked. Finally, the new flexible measurement strategy proposed in this paper can be used in other subfields of political science to observe how information is used by a multitude of actors in different venues.

The key findings of this research also connect to a broader debate on the consequences of polarization. If members of Congress increasingly shift to the extreme of the ideological spectrum, and moderate candidates are increasingly less likely to run for office (see e.g., Hall 2019), bureaucratic information in Congress is destined to decrease. As members become increasingly ideologically distant from agencies, bureaucracy may be replaced by alternative sources of information, which could be more politicized and biased, such as partisan think tanks or interest groups. This pattern might be exacerbated even further by recent events that undermined agency autonomy, such as the striking down of the *Chevron* doctrine by the Supreme Court and Republicans' heightening animosity towards the so-called deep state.

References

Aberbach, J. D. (1990), *Keeping a Watchful Eye: The Politics of Congressional Oversight*, Brookings Institution Press, Washington, DC.

Aneja, A. & Xu, G. (2024), 'Strengthening state capacity: Civil service reform and public sector performance during the gilded age', *American Economic Review* **114**(8), 2352–87.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20230019*

Arora, A. & Dell, M. (2023), 'Linktransformer: A unified package for record linkage with transformer language models'.

Atteveldt, W. V., Sheafer, T., Shenhav, S. R. & Fogel-Dror, Y. (2017), 'Clause analysis: Using syntactic information to automatically extract source, subject, and predicate from texts with an application to the 2008-2009 gaza war', *Political Analysis* **25**, 207–222.

Austen-Smith, D. (1990), 'Information transmission in debate', *American Journal of Political Science* **34**(1), 124–152.
URL: *http://www.jstor.org/stable/2111513*

Ban, P., Park, J. Y. & You, H. Y. (2023), 'Bureaucrats in congress: Strategic information sharing in policymaking'.
URL: *https://www.dropbox.com/s/28idwmhic06bht9/bureaucrats.pdf?dl=0*

Ban, P., Park, J. Y. & You, H. Y. (2024), 'Bureaucrats in congress: Strategic information sharing in policymaking'.
URL: *https://www.dropbox.com/s/28idwmhic06bht9/bureaucrats.pdf?dl=0*

Barrie, C., Palmer, A. & Spirling, A. (2024), 'Replication for language models: Problems, principles, and best practice for political science'.
URL: *https://arthurspirling.org/documents/BarriePalmerSpirling_TrustMeBro.pdf*

Barro, R. J. & Gordon, D. B. (1983), 'A positive theory of monetary policy in a natural rate model', *Journal of political economy* **91**(4), 589–610.

Bellodi, L. (2023), 'A dynamic measure of bureaucratic reputation: New data for new theory', *American Journal of Political Science* **67**(4), 880–897.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12695*

Bertelli, A. M., Mason, D. P., Connolly, J. M. & Gastwirth, D. A. (2013), 'Measuring agency attributes with attitudes across time : A method and examples using large-scale federal surveys', *Journal of Public Administration Research and Theory* **25**, 513–544.

Bertelli, A. M. & Whitford, A. B. (2009), 'Perceiving credible commitments: How independent regulators shape elite perceptions of regulatory quality', *British Journal of Political Science* **39**, 517–537.

Bestvater, S. E. & Monroe, B. L. (2023), 'Sentiment is not stance: Target-aware opinion classification for political text analysis', *Political Analysis* **31**(2), 235–256.

Blom-Hansen, J., Baekgaard, M. & Serritzlew, S. (2021), 'How bureaucrats shape political decisions: The role of policy information', *Public Administration* **99**(4), 658–678.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/padm.12709*

Bolton, A., de Figueiredo, J. M. & Lewis, D. E. (2020), 'Elections, Ideology, and Turnover in the US Federal Government', *Journal of Public Administration Research and Theory* **31**(2), 451–466.
URL: *https://doi.org/10.1093/jopart/muaa051*

Bolton, A., Potter, R. A. & Thrower, S. (2016), 'Organizational capacity, regulatory review, and the limits of political control', *Journal of Law, Economics, and Organization* **32**, 242–271.

Bonica, A. (2023), 'Database on ideology, money in politics, and elections: Public version 3.1', Computer file.
URL: *https://data.stanford.edu/dime*

Bonica, A., Chen, J. & Johnson, T. (2015), 'Senate gate-keeping, presidential staffing of "inferior offices," and the ideological composition of appointments to the public bureaucracy', *Quarterly Journal of Political Science* **10**, 5–40.

Burnham, M., Kahn, K., Wang, R. Y. & Peng, R. X. (2024), 'Political debate: Efficient zero-shot and few-shot classifiers for political text'.
URL: *https://arxiv.org/abs/2409.02078*

Callander, S. (2008), 'A theory of policy expertise', *Quarterly Journal of Political Science* **3**(2), 123–140.
URL: *http://dx.doi.org/10.1561/100.00007024*

Carpenter, D. P. (2001), *The Forging of Bureaucratic Autonomy*, Princeton University Press.

Chen, J. & Johnson, T. (2015), 'Federal employee unionization and presidential control of the bureaucracy: Estimating and explaining ideological change in executive agencies', *Journal of Theoretical Politics* **27**, 151–174.

Choi, J. D., Tetreault, J. & Stent, A. (2015), 'It depends: Dependency parser comparison using a web-based evaluation tool', *ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference* **1**, 387–396.

Clinton, J. D., Bertelli, A., Grose, C. R., Lewis, D. E. & Nixon, D. C. (2012), 'Separated powers in the united states: The ideology of agencies, presidents, and congress', *American Journal of Political Science* **56**, 341–354.

Clinton, J. D. & Lewis, D. E. (2008), 'Expert opinion, agency characteristics, and agency preferences', *Political Analysis* **16**, 3–20.

Crawford, V. & Sobel, J. (1982), 'Strategic information transformation', *Econometrica* **50**, 1431–1451.

Cukierman, A., Webb, S. B. & Neyapti, B. (1992), 'Measuring the independence of central banks and its effects on policy outcomes', *The World Bank Economic Review* **6**, 353–398.

Currinder, M. (2018), *Money in the House: Campaign Funds and Congressional Party Politics*, first edition edn, Routledge.
URL: *https://doi.org/10.4324/9780429498862*

Denisenko, A., Hafer, C. & Landa, D. (2024), 'Competence and advice', *American Journal of Political Science* .
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12905*

Dickson, Z. P. & Hobolt, S. B. (2024), 'Going against the grain: Climate change as a wedge issue for the radical right', *Comparative Political Studies* .

Egan, P. J. (2013), *Partisan Priorities: How Issue Ownership Drives and Distorts American Politics*, Cambridge University Press.

Ellig, J. & Conover, C. J. (2014), 'Presidential priorities, congressional control, and the quality of regulatory analysis: an application to healthcare and homeland security', *Public Choice* **161**, 305–320.

Epstein, D. & O'Halloran, S. (1999), *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers*, Cambridge University Press.

Fiorina, M. P. (1981), *Congressional Control of the Bureaucracy: A Mismatch of Incentives and Capabilities*, 2nd edn, Congressional Quarterly Press.

Forbes (2020), 'Supreme court says trump can fire consumer watchdog director, but cfpb here to stay'. Accessed: 2024-09-15.
URL: *https://www.forbes.com/sites/advisor/2020/06/30/supreme-court-says-trump-can-fire-consumer-watchdog-director-but-cfpb-here-to-stay/?sh=582cd3218c6a*

Gailmard, S. & Patty, J. (2012), 'Formal models of bureaucracy', *Annual Review of Political Science* **15**, 353–377.

Gailmard, S. & Patty, J. (2013), *Learning While Governing: Expertise and Accountability in the Executive Branch*, University of Chicago Press.

Gilardi, F., Alizadeh, M. & Kubli, M. (2023), 'Chatgpt outperforms crowd-workers for text-annotation tasks', *Proceedings of the National Academy of Sciences* pp. 1–14.
URL: *http://arxiv.org/abs/2303.15056*

Gilligan, T. W. & Krehbiel, K. (1987), 'Collective decisionmaking and standing committees: An informational rationale for restrictive amendment procedures', *Journal of Law, Economics, and Organization* **3**, 287–335.

Gilligan, T. W. & Krehbiel, K. (1989), 'Asymmetric information and legislative rules with a heterogeneous committee', *American Journal of Political Science* **33**(2), 459–490.

Grimmer, J. (2013), *Representational style in Congress : What Legislators Say and Why It Matters*, Cambridge University Press, Cambridge.

Grimmer, J., Westwood, S. & Messing, S. (2014), *The Impression of Influence: Legislator Communication, Representation, and Democratic Accountability*, Princeton University Press.

Hall, A. B. (2019), *Who Wants to Run?: How the Devaluing of Political Office Drives Polarization*, University of Chicago Press.
URL: *https://doi.org/10.7208/chicago/9780226609607.001.0001*

Honnibal, M. & Johnson, M. (2015), 'An improved non-monotonic transition system for dependency parsing', *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing* pp. 1373–1378.

Keefer, P. & Stasavage, D. (2003), 'The limits of delegation: Veto players, central bank independence, and the credibility of monetary policy', *American Political Science Review* **97**, 407–423.

Koop, C. & Hanretty, C. (2018), 'Political independence, accountability, and the quality of regulatory decision-making', *Comparative Political Studies* **51**, 38–75.

Krause, G. A. (1996), 'The institutional dynamics of policy administration : Bureaucratic influence over securities regulation', *American Journal of Political Science* **40**, 1083–1121.

Kroeger, M. A. (2022), 'Bureaucrats as lawmakers', *Legislative Studies Quarterly* **47**, 257–289.

Lee, F. E. (2016), *Insecure majorities : Congress and the Perpetual Campaign*, The University of Chicago Press, Chicago;.

Lowande, K. (2018), 'Who polices the administrative state?', *American Political Science Review* **112**, 874–890.

Lupia, A. & McCubbins, M. D. (1998), *The Democratic Dilemma: Can Citizens Learn What They Need to Know?*, Cambridge University Press, New York.

Maltzman, F. & Sigelman, L. (1996), 'The politics of talk: Unconstrained floor time in the u.s. house of representatives', *The Journal of Politics* **58**(3), 819–830.
URL: *https://doi.org/10.2307/2960448*

Mayhew, D. (1974), *Congress: The Electoral Connection.*, Yale University Press.

McCubbins, M., Noll, R. G. & Weingast, B. R. (1987), 'Administrative procedures as instruments of political control', *Journal of Law, Economics, & Organization* **3**, 243–277.

McCubbins, M. & Schwartz, T. (1984), 'Congressional oversight overlooked: Police patrols versus fire alarms', *American Journal of Political Science* **28**, 165–179.

Miller, G. J. (2005), 'The political evolution of principal-agent models', *Annual Review of Political Science* **8**, 203–225.
URL: *http://www.annualreviews.org/doi/10.1146/annurev.polisci.8.082103.104840*

Moe, T. M. (2006), 'Political control and the power of the agent', *Journal of Law, Economics, and Organization* **22**, 1–29.

Moe, T. M. (2012), 'Delegation, control, and the study of public bureaucracy', *The Forum* **10**.

Moore, E. H. (2018), 'Polarization, excepted appointments, and the administrative presidency', *Presidential Studies Quarterly* **48**, 72–92.

Nicholson-Crotty, J. & Miller, S. M. (2012), 'Bureaucratic effectiveness and influence in the legislature', *Journal of Public Administration Research and Theory* **22**, 347–371.

Niskanen, W. (1971), *Bureaucracy and Public Economics*, Aldine-Atherton.

Pennebaker, J. W., Boyd, R. L., Jordan, K. & Blackburn, K. (2015), *The Development and Psychometric Properties of LIWC2015*, University of Texas at Austin.

Potter, R. A. (2019), *Bending the Rules. Pocedural Politicking in the Bureaucracy*, University of Chicago Press.

Richardson, M. D., Clinton, J. D. & Lewis, D. E. (2018), 'Elite perceptions of agency ideology and workforce skill', *Journal of Politics* **80**, 303–308.

Selin, J. L. (2015), 'What makes an agency independent?', *American Journal of Political Science* **59**, 971–987.

Shobe, J. (2017), 'Agencies as legislators: An empirical study of the role of agencies in the legislative process', *The George Washington Law Review* **85**, 451–535.

Spenkuch, J. L., Teso, E. & Xu, G. (2023), 'Ideology and performance in public organizations', *Econometrica* **91**(4), 1171–1203.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA20355*

Vannoni, M., Ash, E. & Morelli, M. (2021), 'Measuring discretion and delegation in legislative texts: Methods and application to us states', *Political Analysis* **29**, 43–57.

Volden, C. & Wiseman, A. E. (2014), *Legislative Effectiveness in the United States Congress: The Lawmakers*, Cambridge University Press.
URL: *Updated at https://thelawmakers.org*

Volden, C. & Wiseman, A. E. (2018), 'Legislative effectiveness in the united states senate', *Journal of Politics* **80**, 731–735.
URL: *Updated at https://thelawmakers.org*

Weber, M. (1946), *Essays in Sociology*, Oxford University Press, New York. From Max Weber: Essays in Sociology.

Wilson, J. (1989), *Bureaucracy: What government agencies do and why they do it.*, Basic Books.

Appendix

A Speeches: Data Quality

I access transcripts of speeches from two sources. For floor speeches (1994-2022) and for committee speeches (2010-2020), I scraped the digitalized version of the Congressional Record. For committee speeches (1980-2009) I relied on transcripts available through ProQuest.

Online Version of Congressional Record

To access speeches on the online version of the Congressional Record, I obtain the universe of available links to congressional hearings material and to floor speeches via the website `https://www.govinfo.gov/sitemaps`. The sitemaps contain one main URL for each year, and each URL contains as many URLs as the number of packages in which the record has been grouped. For instance, the material for the Senate Hearing with ID 76804 can be accessed via the following link: `https://www.govinfo.gov/content/pkg/CHRG-107shrg76804/html/CHRG-107shrg76804.htm`, which re-directs to the text of the entire committee session.

Thanks to metadata listing the members of Congress who gave speeches in the session, it is possible to extract the speech with a set of flexible regular expressions that capture the structure "title + surname + period + white space + start of sentence".

Pro-Quest Data

For older committee sessions, I accessed transcripts directly from ProQuest. I obtained 42,277 transcripts of congressional committee sessions, each one consisting of one large html file, and no metadata exists to facilitate the extraction of single speeches. Speeches are nonetheless identifiable thanks to the way they appear in the text. The title and SURNAME of the speaker precede the speech and are reported in capital cases. "Mr. FORD", for instance, marks a new speech. Many individuals give speeches or statements, and to extract speeches by politicians alone, I exploit the fact that at the beginning of each transcript, the names of all members of Congress are reported, followed by their home state. From every transcript, I, therefore, extract all the names of politicians with a regular expression that matches the name and surname of individuals followed by the name of their respective state. Only speeches by any of the extracted names are parsed from the transcript.

Despite some typos in the full text, a careful look at a random sample of parsed speeches suggests the quality of the parsing procedure is sufficiently high to confidently attribute speeches to legislators. By merging the surname, date of congressional session, and state of the legislators, I am then able to match data on committee speeches with each legislator's DW-NOMINATE score.

B Extraction Rules

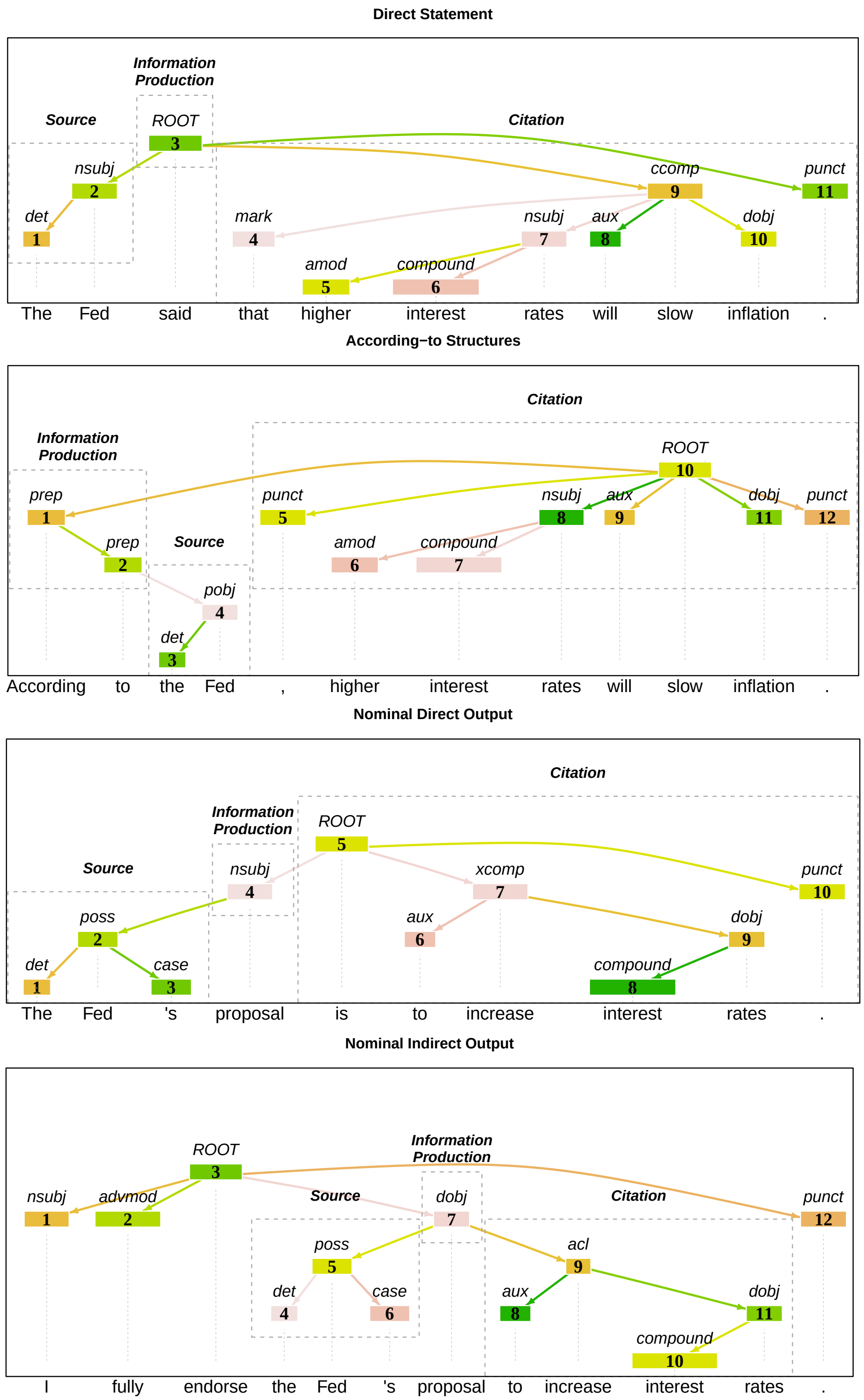
Below I report the lists of say-verbs and output-type words used to extract legislators' citations of bureaucratic information.

Say verbs used to match syntactic rules: acknowledge, admit, advance, advise, advise, affirm, agree, argue, assert, assume, assume, assure, believe, claim, clarify, complain, concede, conclude, confirm, consider, contend, convince, decide, decide, define, demonstrate, document, encourage, estimate, evaluate, explain, find, identify, indicate, inform, predict, present, presume, project, propose, propose, prove, realise, realize, recommend, refer, remind, report, respond, reveal, say, see, set out, show, state, suggest, tell, testify, think, warn.

Output-type words are: advice, advise, analysis, argument, article, assessment, brief, comment, complaint, conclusion, copy, data, datum, decision, directive, document, estimate, evidence, figure, forecast, guidance, guideline, idea, indication, information, input, inquiry, instruction, memo, observation, opinion, paper, plan, position, prediction, prevision, program, programme, project, projection, proposal, rationale, reasoning, recommendation, report, statement, statistic, statistics, strategy, study, suggestion, survey, testimony, thesis, view.

C Dependency Parsing: Examples

Figure C.1: Dependency Trees for Additional Syntactic Frames.



Notes: Parsed dependency trees of the four remaining illustrative examples where the FED is used in speech. Implemented through the *rsyntax* package in R.

D Classifying Policy Content

To classify citations into policy and non-policy content, I first represent each citation numerically, building on recent advancements in deep learning, and use a pre-trained language model that is suitable for downstream NLP tasks on short texts. The model takes as input the text of the citation and produces a fixed-length, dense vector that encodes its meaning and semantic properties in a numerical form. I use the SentenceTransformers framework, and specifically the model `paraphrase-MiniLM-L6-v2` available at `https://huggingface.co/sentence-transformers`.

Once I have an embedding for every citation, I assemble the training dataset on which to train the classifier. Because the language used to convey policy information produced by agencies can be very different from one domain to another, I extract a large random sample of 30,000 citations and annotate them with large language models. To ensure a high-quality annotation, I select the model that, at the time of writing, achieves the best performance on difficult tasks (`gpt-4o`) and ensure replicability of the annotation by setting the temperature of the model — a parameter that governs the randomness and thus the creativity of the responses — to 0, which means that the responses will be almost deterministic, yielding the same response over and over again to a given prompt. We leave all other parameters to their default settings. After careful design, I assign to GPT the following prompt:

> I will show you a statement given by a US member of Congress citing the `agency_name`. Determine whether the member of Congress is citing policy information produced by the `agency_name`, for example, statistical evidence, policy reports, expected outcomes of a program, and so on. Simply print 'Yes' if the statement contains policy information and 'No' if it doesn't. If the member is not citing the `agency_name`, then print 'No'. This is the statement: <statement>`sentence_quoting_agency`</statement>.

On the annotated dataset, I then train 4 different classifiers, available through the `sklearn` library in Python:

1. Gradient Boosting Classifier (GB)
2. AdaBoost Classifier (ABC)
3. Random Forest (RF)
4. Linear Discriminant Analysis (LDA)

I set the model parameters to their default setting and train the model on 75% of the sentences in the training dataset, and I leave the remaining 25% as a held-out test set on which to assess the performance of the classifier.

In Table D.1 below, I report the performance metrics of each of the four classifiers. While all classifiers work well on the held-out test set, I select the LDA, for it is the one performing best.I then train the LDA model on the entire training dataset (this time including the test set too). Finally, I apply the classifier to the total sample of sentences quoting agencies.

Table D.1: Performance metrics of classifiers.

Classifier	F1	Accuracy	Recall	Precision
GB	0.786	0.805	0.775	0.798
RF	0.777	0.803	0.741	0.816
ABC	0.765	0.784	0.761	0.770
LDA	0.826	0.840	0.820	0.831

Notes: Performance metrics for each classifier. Performance metrics calculated from comparing the model predictions to the annotated labels of held-out test sets (25% of the tweets in a 30,000-sentence training dataset).

D1 List of Agencies

Table D.2: List of Agencies and Descriptive Statistics.

	Mentions			Citations			Policy Citations		
Agency Name	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**
Environmental Protection Agency	102,213	54	45	6,899	56	44	3,063	59	40
Department of Defense	110,078	56	44	4,824	58	41	1,968	59	41
Centers for Disease Control and Prevention	13,493	55	45	2,325	57	42	1,867	59	41
Department of Energy	62,654	55	45	3,612	53	46	1,695	52	47
Department of Agriculture	39,808	56	44	2,495	55	45	1,495	55	45
Department of State	55,657	52	48	3,582	50	50	1,252	48	52
Department of the Treasury	65,563	52	47	3,706	53	47	1,220	54	46
Department of Justice	69,948	53	47	4,193	52	47	1,215	52	47
Bureau of the Census	7,363	63	37	1,430	59	40	1,174	60	40
Office of Management and Budget	41,431	60	40	2,741	58	42	1,133	55	45
Department of Veterans Affairs	89,511	53	46	2,534	55	44	1,132	58	41
Department of Labor	18,413	55	45	1,657	59	41	1,043	61	39
Federal Reserve	47,728	55	44	2,883	56	43	1,001	57	43
Department of Transportation	20,834	57	42	1,682	56	43	927	54	45
Federal Bureau of Investigation	48,693	49	51	2,758	51	48	916	57	43
Bureau of Labor Statistics	2,424	58	42	973	55	45	906	55	45
Air Force	70,705	52	48	2,454	61	39	892	63	37
Department of the Navy	79,350	54	46	2,573	60	40	871	63	37
Department of Health and Human Services	21,402	51	49	1,538	52	48	855	54	46
Department of Commerce	15,113	55	45	1,235	55	45	811	53	46
Internal Revenue Service	48,176	42	58	2,098	46	54	769	52	47
Department of Education	17,724	52	48	980	54	45	683	57	41
Small Business Administration	25,288	60	40	1,258	51	49	654	42	58
Forest Service	28,858	48	52	1,559	49	51	652	54	46
Food and Drug Administration	52,949	56	43	2,215	59	41	595	59	40
Energy Information Administration	1,507	49	50	652	49	50	583	50	50
Department of Housing and Urban Development	28,013	58	42	1,321	57	43	557	62	37
Coast Guard	47,139	56	44	1,319	58	41	525	57	42
Department of Homeland Security	32,827	49	51	1,306	50	50	448	50	49
Pentagon	21,723	55	44	1,311	58	41	426	58	41
Federal Trade Commission	15,382	60	40	924	60	39	415	64	35
National Oceanic and Atmospheric Administration	11,530	56	44	673	60	39	390	63	37
Federal Emergency Management Agency	23,461	58	42	852	59	41	310	58	41
National Aeronautics and Space Administration	31,270	56	44	1,072	57	42	303	62	37
Federal Communications Commission	26,087	55	45	1,163	56	43	291	56	44
International Trade Commission	3,531	52	48	449	49	51	273	51	49
Drug Enforcement Administration	10,483	48	52	620	40	60	260	41	59
Federal Deposit Insurance Corporation	10,311	56	44	610	55	45	260	58	41
US Postal Service	25,819	57	43	794	65	34	255	68	29
Federal Highway Administration	2,560	53	47	320	59	41	253	58	42
Social Security Administration	6,530	52	47	397	50	48	245	50	48
Central Intelligence Agency	17,516	55	45	1,216	54	45	230	56	44
Office of Personnel Management	8,619	59	40	443	62	37	229	59	40
Geological Survey	1,927	60	40	310	57	43	226	59	40
General Services Administration	18,554	56	43	735	59	41	223	59	41
Nuclear Regulatory Commission	15,218	62	37	827	62	37	213	55	43
Fish and Wildlife Service	7,449	45	55	602	44	56	203	43	57
Securities and Exchange Commission	28,735	54	45	1,083	53	46	203	56	43
Department of the Interior	13,415	53	47	680	53	47	188	50	50
National Science Foundation	13,497	60	40	451	63	37	176	69	31
Occupational Safety and Health Administration	16,997	50	50	551	45	55	176	44	56
National Institutes of Health	7,789	52	47	314	55	44	167	56	43
Council of Economic Advisers	1,248	56	43	265	57	43	158	62	38
National Park Service	10,943	51	49	541	47	53	158	44	56
Customs and Border Protection	5,755	49	51	273	37	63	142	32	68
National Highway Traffic Safety Administration	1,126	52	48	171	56	44	140	57	43
Centers for Medicare and Medicaid Services	1,146	50	50	170	46	54	131	47	53
Consumer Product Safety Commission	4,557	72	27	196	76	24	131	76	24

Table D.2: List of Agencies and Descriptive Statistics. *(continued)*

	Mentions			Citations			Policy Citations		
Agency Name	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**
Federal Energy Regulatory Commission	13,980	56	44	557	61	38	131	56	43
Agency for International Development	8,692	55	45	251	51	49	128	51	49
Consumer Financial Protection Bureau	8,176	40	59	294	43	57	113	65	35
Office of National Drug Control Policy	2,393	38	62	188	43	57	113	42	58
National Transportation Safety Board	2,020	60	40	255	60	40	108	57	43
Transportation Security Administration	11,848	51	48	315	48	50	98	48	51
Immigration and Customs Enforcement	6,725	44	55	288	40	60	94	32	68
Bureau of Land Management	3,978	45	55	196	46	54	84	50	50
Bureau of Prisons	2,288	60	40	133	56	44	82	51	49
Federal Aviation Administration	3,394	54	45	131	52	47	74	51	47
Commodities Futures Trading Commission	8,477	56	44	328	55	44	66	62	36
Economic Research Service	347	55	45	70	64	36	65	65	35
Bureau of Economic Analysis	228	55	45	64	47	53	59	47	53
Substance Abuse and Mental Health Services Administration	412	51	49	63	51	49	57	51	49
Bureau of Reclamation	4,327	48	52	178	56	44	55	62	38
Federal Election Commission	4,482	49	51	189	48	51	55	62	35
Indian Health Service	2,908	59	41	118	54	46	48	54	46
Bureau of Indian Affairs	2,419	56	44	105	50	50	47	55	45
Public Health Service	1,961	60	40	101	61	39	47	57	43
Maritime Administration	1,081	57	43	72	53	47	40	48	52
Veterans Health Administration	1,905	56	44	58	64	36	40	72	28
Citizen and Immigration Services	1,454	52	47	80	51	48	37	54	46
Federal Transit Administration	695	47	53	42	57	40	33	58	39
Library of Congress	4,400	55	44	166	58	42	33	55	45
National Labor Relations Board	6,302	47	53	245	44	56	33	36	64
Federal Housing Administration	10,125	57	43	123	60	40	31	55	45
Defense Intelligence Agency	770	45	54	88	35	64	29	38	59
Health Resources and Services Administration	643	33	67	42	50	50	29	48	52
Food and Nutrition Service	924	68	32	49	63	37	28	64	36
Federal Railroad Administration	909	64	36	46	61	39	27	63	37
Pension Benefit Guaranty Corporation	755	62	38	40	48	52	27	48	52
Council on Environmental Quality	1,594	59	40	76	53	47	24	58	42
Agency for Healthcare Research and Quality	135	52	47	24	75	25	23	74	26
Bureau of the Public Debt	169	64	36	28	75	25	23	91	9
Export-Import Bank of the United States	4,128	55	44	54	57	43	22	68	32
Equal Employment Opportunity Commission	988	69	30	51	75	25	21	86	14
Federal Motor Carrier Safety Administration	614	43	57	34	44	56	21	33	67
Bureau of Alcohol, Tobacco, Firearms, and Explosives	281	53	47	31	71	29	20	85	15
Board of Veterans Appeals	696	67	33	37	76	24	19	84	16
Commission on Civil Rights	662	63	37	52	77	23	19	89	11
International Trade Administration	442	52	48	27	15	85	19	16	84
National Guard Bureau	763	57	43	54	59	41	19	68	32
Farm Service Agency	687	42	58	28	39	61	18	33	67
Merit Systems Protection Board	1,203	57	43	58	62	38	17	53	47
National Institute of Standards and Technology	1,125	55	45	36	64	36	17	76	24
Office of the Secretary of Defense	1,255	58	42	38	79	21	17	82	18
Agency for Toxic Substances and Disease Registry	146	68	32	16	75	25	16	75	25
Tennessee Valley Authority	1,219	49	51	44	50	48	16	50	50
Department of the Army	1,381	56	44	46	65	35	15	73	27
Government Printing Office	1,137	58	42	30	67	33	15	60	40
National Institute of Building Sciences	101	62	38	16	75	25	15	73	27
Federal Maritime Commission	1,698	61	39	40	62	38	14	29	71
Patent and Trademark Office	1,586	56	44	57	54	46	14	79	21
Peace Corps	7,907	65	35	106	54	46	14	36	64
Postal Rate Commission	588	54	46	34	53	47	14	57	43
Defense Contract Audit Agency	295	76	24	27	81	19	13	85	15
Defense Logistics Agency	599	62	38	25	52	44	13	46	46
Missile Defense Agency	506	53	47	41	73	27	13	69	31
Election Assistance Commission	468	67	33	21	52	48	12	50	50

Table D.2: List of Agencies and Descriptive Statistics. *(continued)*

	Mentions			Citations			Policy Citations		
Agency Name	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**
National Security Agency	1,257	55	45	82	60	39	12	50	50
Postal Regulatory Commission	236	65	35	23	61	39	12	50	50
Veterans Benefits Administration	1,222	61	39	24	50	50	12	50	50
National Archives and Records Administration	234	56	44	15	47	53	11	55	45
Natural Resources Conservation Service	381	37	62	20	50	50	11	64	36
Pipeline and Hazardous Materials Safety Administration	117	50	50	16	38	62	11	27	73
Railroad Retirement Board	279	53	47	20	60	40	11	82	18
Agricultural Research Service	625	54	46	18	56	44	10	60	40
Bonneville Power Administration	850	52	48	25	64	36	10	60	40
Foreign Agricultural Service	565	53	47	19	63	37	10	50	50
National Agricultural Statistics Service	72	44	56	10	40	60	10	40	60
Office of Federal Procurement Policy	371	61	39	20	65	35	10	70	30
Office of the Comptroller of the Currency	420	61	38	21	52	48	10	70	30
Commodity Credit Corporation	1,037	57	43	13	38	62	9	44	56
Food Safety and Inspection Service	391	47	53	15	60	40	9	56	44
Mine Safety and Health Administration	329	62	38	20	50	50	9	22	78
National Credit Union Administration	1,079	52	47	35	49	51	9	44	56
Appalachian Regional Commission	1,021	58	42	19	37	63	8	38	62
Office of Science	1,229	49	50	34	65	35	8	50	50
Office of Special Counsel	909	61	39	56	52	48	8	38	62
Corporation for National and Community Service	420	64	36	17	65	35	7	86	14
Interstate Commerce Commission	764	62	38	26	62	38	7	43	57
National Cemetery Administration	138	60	40	11	73	27	7	71	29
National Indian Gaming Commission	256	55	45	10	40	60	7	29	71
Agricultural Marketing Service	269	59	41	10	80	20	6	67	33
Community Development Financial Institutions Fund	257	51	49	6	100	0	6	100	0
Financial Crimes Enforcement Network	232	50	50	10	70	30	6	67	33
Legal Services Corporation	2,512	54	46	29	55	45	6	67	33
National Nuclear Security Administration	431	45	55	33	48	52	6	67	33
National Telecommunications and Information Administration	299	64	36	9	78	22	6	67	33
Office of the Director of National Intelligence	228	54	43	22	36	59	6	17	83
Administration for Children and Families	155	55	45	8	62	38	5	60	40
Advisory Council on Historic Preservation	180	58	42	10	80	20	5	60	40
Animal and Plant Health Inspection Service	361	48	52	10	70	30	5	40	60
Office of Government Ethics	908	49	51	72	47	53	5	40	60
Surface Transportation Board	665	52	48	18	28	72	5	60	40
Defense Contract Management Agency	68	47	53	5	40	60	4	25	75
Economic Development Administration	1,234	71	29	20	75	25	4	75	25
Farm Credit Administration	805	54	46	20	50	50	4	50	50
Federal Housing Finance Agency	233	50	50	7	29	71	4	50	50
Marshals Service	1,470	51	49	37	54	46	4	75	25
National Institute on Disability and Rehabilitation Research	20	80	20	4	75	25	4	75	25
Parole Commission	306	72	28	12	75	25	4	50	50
Rehabilitation Services Administration	90	70	30	4	100	0	4	100	0
Administrative Conference of the United States	571	40	60	41	49	51	3	0	100
Broadcasting Board of Governors	186	46	53	12	83	17	3	100	0
Bureau of Ocean Energy Management	82	51	49	5	0	100	3	0	100
Civil Rights Division	1,879	66	34	34	59	41	3	33	67
Federal Agricultural Mortgage Corporation	502	42	58	6	17	83	3	33	67
Federal National Mortgage Association	113	65	34	7	57	43	3	0	100
Federal Prison Industries	774	52	48	14	21	79	3	0	100
Federal Retirement Thrift Investment Board	95	65	35	3	33	67	3	33	67
Minority Business Development Agency	224	78	22	5	80	20	3	67	33
Office of Energy Efficiency and Renewable Energy	138	59	41	4	75	25	3	100	0
Office of Thrift Supervision	429	66	34	25	76	24	3	33	67
Social Security Advisory Board	51	24	75	6	33	67	3	33	67

Table D.2: List of Agencies and Descriptive Statistics. *(continued)*

	Mentions			Citations			Policy Citations		
Agency Name	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**
Western Area Power Administration	169	49	51	5	60	40	3	67	33
American Battle Monuments Commission	200	50	50	8	50	50	2	0	100
Bureau of Engraving and Printing	179	58	42	6	50	50	2	50	50
Corporation for Public Broadcasting	863	57	43	12	42	58	2	50	50
Defense Finance and Accounting Service	101	37	62	4	25	75	2	50	50
Defense Nuclear Facilities Safety Board	134	63	37	9	56	44	2	100	0
Defense Security Cooperation Agency	13	54	46	2	50	50	2	50	50
Defense Technical Information Center	26	69	31	2	50	50	2	50	50
Federal Law Enforcement Training Center	333	47	52	4	50	50	2	50	50
Financial Management Service	163	31	69	3	33	67	2	0	100
Government National Mortgage Association	631	56	44	6	33	67	2	50	50
Institute of Peace	328	57	43	6	67	33	2	50	50
National Capital Planning Commission	231	61	38	8	50	50	2	50	50
National Council on Disability	109	60	40	10	60	40	2	50	50
National Mediation Board	365	39	61	12	25	75	2	50	50
Office of Administration	307	48	52	10	50	50	2	0	100
Office of Fiscal Service	9	67	33	2	100	0	2	100	0
Office of Health, Safety, and Security	44	73	27	2	100	0	2	100	0
Office of Justice Programs	316	45	54	4	25	75	2	50	50
Privacy and Civil Liberties Oversight Board	204	51	46	10	40	60	2	0	100
Wage and Hour Division	275	63	37	3	100	0	2	100	0
Benefits Review Board	43	60	40	1	0	100	1	0	100
Bureau of Consular Affairs	92	57	43	1	100	0	1	100	0
Bureau of Economics	82	73	27	6	83	17	1	100	0
Bureau of Indian Education	108	52	48	1	0	100	1	0	100
Bureau of Industry and Security	49	57	43	2	100	0	1	100	0
Bureau of Safety and Environmental Enforcement	16	56	44	2	100	0	1	100	0
Chemical Safety and Hazard Investigation Board	66	62	38	3	67	33	1	100	0
Defense Commissary Agency	66	53	47	2	50	50	1	100	0
Defense Health Agency	35	46	51	1	0	100	1	0	100
Defense Information Systems Agency	52	50	50	2	50	50	1	100	0
Defense Threat Reduction Agency	80	52	48	3	67	33	1	0	100
Domestic Nuclear Detection Office	142	49	51	5	40	60	1	100	0
Employment and Training Administration	188	55	45	1	100	0	1	100	0
Executive Office for United States Attorneys	26	62	38	1	0	100	1	0	100
Federal Home Loan Mortgage Corporation	65	55	45	1	100	0	1	100	0
Federal Labor Relations Authority	216	54	46	6	67	33	1	100	0
Financial Stability Oversight Council	379	42	58	26	50	50	1	100	0
Housing Finance Agency	12	83	17	1	100	0	1	100	0
Independent Payment Advisory Board	408	20	80	5	40	60	1	100	0
Justice Management Division	38	50	50	2	100	0	1	100	0
Millennium Challenge Corporation	282	46	54	7	29	71	1	0	100
Missing Personnel Office	27	33	67	2	0	100	1	0	100
National Technical Information Service	140	69	31	1	100	0	1	100	0
Office of Environmental Management	97	45	55	1	0	100	1	0	100
Office of Legal Counsel	1,191	54	46	109	57	43	1	0	100
Office of Minority Economic Impact	37	81	19	1	100	0	1	100	0
Office of Navajo and Hopi Indian Relocation	16	25	75	1	0	100	1	0	100
Office of Nuclear Reactor Regulation	31	61	39	2	50	50	1	0	100
Office of Surface Mining, Reclamation and Enforcement	78	56	44	3	67	33	1	100	0
Office of Vocational and Adult Education	11	82	18	1	100	0	1	100	0
Office of the United States Trade Representative	93	51	49	4	50	50	1	0	100
Overseas Private Investment Corporation	472	50	49	7	14	86	1	0	100
Panama Canal Commission	343	62	38	7	43	57	1	0	100
Public and Indian Housing	136	60	40	2	50	50	1	100	0
Research and Innovative Technology Administration	7	71	29	1	100	0	1	100	0
Risk Management Agency	138	29	70	5	40	60	1	100	0
Rural Housing Service	110	32	68	1	100	0	1	100	0
Rural Utilities Service	217	47	53	1	0	100	1	0	100

Table D.2: List of Agencies and Descriptive Statistics. *(continued)*

Agency Name	Mentions			Citations			Policy Citations		
	Total	D (%)	R (%)	Total	D (%)	R (%)	Total	D (%)	R (%)
Saint Lawrence Seaway Development Corporation	52	56	44	2	100	0	1	100	0
Veterans Employment and Training Service	45	31	69	2	50	50	1	100	0
Women's Bureau	36	72	28	1	100	0	1	100	0
Administration for Community Living	24	75	25	1	100	0	0	0	0
Alcohol and Tobacco Tax and Trade Bureau	23	70	30	0	0	0	0	0	0
Arms Control and International Security	92	41	59	6	33	67	0	0	0
Board of Directors of the Hope for Homeowners Program	0	0	0	0	0	0	0	0	0
Border and Transportation Security Directorate	36	14	86	1	0	100	0	0	0
Bureau of Competition	143	67	33	4	50	50	0	0	0
Bureau of Diplomatic Security	91	38	62	1	100	0	0	0	0
Bureau of Educational and Cultural Affairs	120	62	38	2	50	50	0	0	0
Bureau of International Labor Affairs	17	59	41	1	0	100	0	0	0
Bureau of International Narcotics and Law Enforcement Affairs	34	32	68	0	0	0	0	0	0
Bureau of Political-Military Affairs	13	54	46	0	0	0	0	0	0
Civilian Security, Democracy, and Human Rights	6	50	50	0	0	0	0	0	0
Court Services and Offender Supervision Agency	68	57	40	0	0	0	0	0	0
Defense Acquisition Regulations System	2	0	100	0	0	0	0	0	0
Defense Advanced Research Projects Agency	191	57	43	4	50	25	0	0	0
Defense Human Resources Activity	6	50	50	0	0	0	0	0	0
Defense Legal Services Agency	3	0	100	0	0	0	0	0	0
Defense Media Activity	5	60	40	0	0	0	0	0	0
Defense Security Service	25	40	60	1	0	100	0	0	0
Defense Technology Security Administration	27	78	22	1	100	0	0	0	0
Delta Regional Authority	49	53	47	0	0	0	0	0	0
Department of Defense Education Activity	0	0	0	0	0	0	0	0	0
Directorate of Defense Trade Controls	5	40	60	0	0	0	0	0	0
Economic Growth, Energy, and the Environment	13	46	54	0	0	0	0	0	0
Economics and Statistics Administration	14	50	50	0	0	0	0	0	0
Employee Benefits Security Administration	19	32	63	0	0	0	0	0	0
Employees Compensation Appeals Board	3	67	33	0	0	0	0	0	0
Employment Standards Administration	65	54	46	0	0	0	0	0	0
European and Eurasian Affairs	72	43	57	3	33	67	0	0	0
Executive Office for Immigration Review	80	42	57	2	50	50	0	0	0
Federal Mediation and Conciliation Service	115	55	45	2	50	50	0	0	0
Federal Mine Safety and Health Review Commission	67	61	39	1	0	100	0	0	0
Federal Student Aid	135	61	39	2	100	0	0	0	0
Field Policy and Management	0	0	0	0	0	0	0	0	0
Foreign Claims Settlement Commission	129	64	36	2	0	100	0	0	0
Grain Inspection, Packers, and Stockyards Administration	57	46	54	0	0	0	0	0	0
Institute of Education Sciences	34	56	44	0	0	0	0	0	0
Inter-American Foundation	6	67	33	1	0	100	0	0	0
International Boundary and Water Commission	120	52	48	4	50	50	0	0	0
Metropolitan Washington Airport Authority	30	60	40	0	0	0	0	0	0
Mississippi River Commission	101	46	54	8	100	0	0	0	0
National Consumer Cooperative Bank	54	52	48	0	0	0	0	0	0
National Foundation on the Arts and the Humanities	6	67	33	0	0	0	0	0	0
National Geospatial-Intelligence Agency	36	39	61	1	0	100	0	0	0
National Infrastructure Protection Center	36	31	69	1	0	100	0	0	0
National Institute of Food and Agriculture	87	17	83	0	0	0	0	0	0
National Reconnaissance Office	245	36	64	4	75	25	0	0	0
National Security Education Board	8	62	38	0	0	0	0	0	0
Occupational Safety and Health Review Commission	141	25	75	2	50	50	0	0	0
Office of Acquisition Policy	8	75	25	1	0	100	0	0	0
Office of Economic Adjustment	73	78	22	0	0	0	0	0	0
Office of Electricity Delivery and Energy Reliability	74	58	41	2	100	0	0	0	0

Table D.2: List of Agencies and Descriptive Statistics. *(continued)*

	Mentions			Citations			Policy Citations		
Agency Name	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**	**Total**	**D (%)**	**R (%)**
Office of Elementary and Secondary Education	18	72	28	1	0	100	0	0	0
Office of Energy Policy and New Uses	0	0	0	0	0	0	0	0	0
Office of Federal Contract Compliance Programs	82	70	30	0	0	0	0	0	0
Office of Foreign Assets Control	210	52	48	6	67	33	0	0	0
Office of Fossil Energy	87	45	55	2	50	50	0	0	0
Office of Healthy Homes and Lead Hazard Control	5	100	0	0	0	0	0	0	0
Office of Labor-Management Standards	37	46	54	2	50	50	0	0	0
Office of Nuclear Energy	173	40	60	0	0	0	0	0	0
Office of Postsecondary Education	26	77	23	1	100	0	0	0	0
Office of Rural Development	13	62	38	2	100	0	0	0	0
Office of Safe and Healthy Students	0	0	0	0	0	0	0	0	0
Office of Special Education and Rehabilitative Services	34	59	41	1	0	100	0	0	0
Office of Special Trustee for American Indians	10	70	30	0	0	0	0	0	0
Office of Surety Guarantees	3	33	67	0	0	0	0	0	0
Office of Workers' Compensation Programs	70	30	70	2	0	100	0	0	0
Office of the Federal Coordinator for Alaska Natural Gas Transportation Projects	0	0	0	0	0	0	0	0	0
Public Buildings Service	210	51	49	4	50	50	0	0	0
Public International Organization	2	0	100	0	0	0	0	0	0
Rural Business and Cooperative Development Service	12	17	83	0	0	0	0	0	0
Securities Investor Protection Corporation	68	47	53	2	50	50	0	0	0
Selective Service System	389	57	43	4	50	50	0	0	0
Smithsonian Institution	1,307	57	43	25	36	64	0	0	0
State Justice Institute	142	54	46	3	33	67	0	0	0
Test Resource Management Center	6	33	67	0	0	0	0	0	0
Trade Deficit Review Commission	7	29	71	0	0	0	0	0	0
Trade and Development Agency	120	60	39	0	0	0	0	0	0
Tricare Management Activity	2,097	58	42	11	55	45	0	0	0
US Information Agency	143	50	50	5	80	20	0	0	0
US Soldiers' and Airmen's Home	7	100	0	0	0	0	0	0	0
Washington Headquarters Services	15	60	40	0	0	0	0	0	0

Notes: Number of times members of Congress mentioned, cited, or used policy information produced by each bureaucratic agency. Percentages for Democrats and Republicans are reported next to each column indicating the total for each agency. Agencies are displayed in descending order based on the total number of mentions by members regarding policy information.

D2 Predictive Validity of Members' Use of Agency Policy Information

As an additional test in support of the validity of the measure, I show that members from both parties are more likely to cite information from agencies whose workforce is perceived as more skilled. I use data from Richardson et al. (2018) on perceptions of workforce skill for 139 agencies and use this indicator to predict the number of times members cite policy information produced by agencies. Richardson et al. (2018) produce time-fixed estimates of agencies' workforce skills by applying a statistical model to a large survey of federal. I build an unbalanced panel at the agency-congress level and count how many times members of Congress cite policy information from any given agency over time. For this test, I pool speeches given on the floor and on committees.

In particular, I estimate the following model via OLS:

$$y_{jt}^{P} = \alpha_t + \pi_c + \beta \text{Skill}_j + \epsilon_{jt}$$

where y_{jt}^{P} is the number of times members from party P use policy information produced by agency j in congress t. Skill is the standardized indicator of workforce skill produced by Richardson et al. (2018). Because this measure does not change over time, I am not able to estimate agency fixed effects, hence I include agency-type dummies, accounting for differences between executive department, independent agencies, executive sub-agencies, and boards and commissions. I estimate the equation on members from all parties, as well as subsetting the data to members of the Republican and Democratic parties separately. Table D.3 reports the results. Column (1) shows a positive correlation between agencies' perceived skillfulness and the number of times members cite policy information from agencies. Columns (2) and (3) report similar effects for Democrats and Republicans. On average, a +1 standard deviation in the agency's perceived skillfulness is associated with a +2-5 citation with policy information.

Table D.3: Agency Perceived Workforce Skill and Members' Reliance on Policy Information.

	# Sentences Citing Policy Information		
Among Members:	All	DEM	REP
	(1)	(2)	(3)
Workforce Skill	3.53*	2.11*	1.42*
	(1.46)	(0.84)	(0.63)
Mean DV	13.96	7.81	6.08
R^2	0.25	0.22	0.26
Observations	2,792	2,792	2,792
Congress FE	✓	✓	✓
Agency Type FE	✓	✓	✓

Notes: OLS estimates. SE clustered by agency. The outcome variable is the number of times members of Congress cite policy information from agencies. The unit of observation is each agency-congress pair. Column (1) uses data from all members of Congress. Columns (2) and (3) subset the data to Democrats and Republicans.

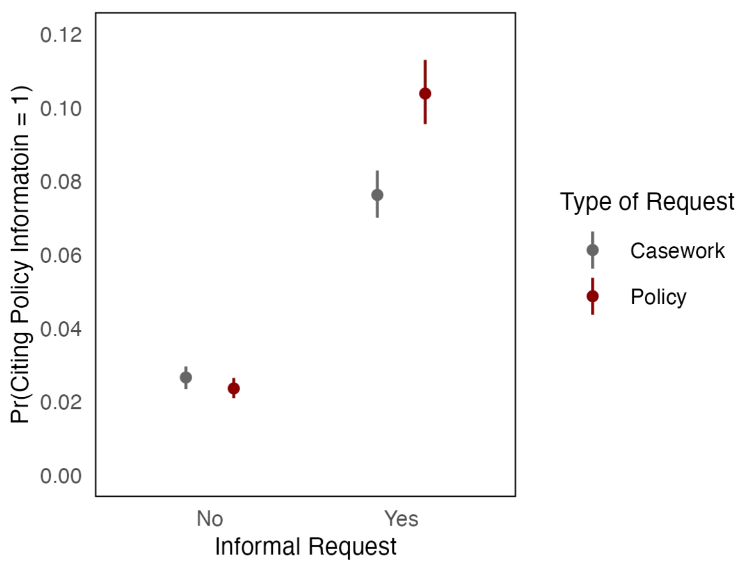
D3 Members' Use of Agency Policy Information and Informal Requests to Agencies

To show that the public nature of congressional speeches correlates with informal, private communication between members of Congress and agencies, I show that the probability that a member cites policy information produced by an agency is highly correlated with whether the same member of Congress sent an informal request to the agency by means of private correspondence.

Data on informal requests from members is from Lowande (2018), who assembles a dataset of members' requests to 16 federal agencies through various FOIA requests, covering the 110th and 111th Congresses (i.e., the last Congress of the second Obama presidency and the first Congress of the Trump presidency). For each member-agency-congress tuple, the data reports whether the member contacted the agency and whether the type of contact concerned policy or casework.

I match the data on MC-agency private correspondence with a subset of my dataset on members' citation of agency policy information, including the 16 agencies covered in Lowande (2018), for a total of 17,408 MC-agency-congress tuples. To display the correlation between members' public and private demand for agency information, I report the conditional probability of citing agency information in Congress based on whether the member privately requested policy information from the agency. Figure D.2 shows that the probability of citing policy information increases from .02 to more than .10 when members privately contacted agencies.

Figure D.2: Members' Public and Private Interaction with Agencies



Notes: Conditional probability of members' citing policy information from agencies based on whether they sent an informal request to the agency concerning policy or casework.

To strengthen the correlation displayed visually, I relate the probability of citing policy information produced by agencies to whether members privately contacted the agency by estimating the following equation:

$$Pr(y_{ijt} = 1) = \gamma_i + \alpha_j + \delta_t + \beta \text{Request}^T_{ijt} + \epsilon_{ijt}$$

where γ_i are MC fixed effects, α_j agency fixed effects, δ_t Congress fixed effects and Request^T_{ijt} is a dummy equal to 1 if member i sent an informal request of type $T =\in \{\text{Casework}, \text{Policy}\}$ to agency j in Congress t. Standard errors are clustered by MC. Table D.4 reports a strong correlation between policy requests made informally through private correspondence between the MC and the agency and whether the member in the same Congress cited policy information from the same agency.

Table D.4: Correlation Between Members' Public and Private Demand of Information

	Pr(# Sentences Citing Policy Information > 0)			
	(1)	(2)	(3)	(4)
Casework Request	0.018***	0.016***	0.017***	0.0002
	(0.004)	(0.004)	(0.004)	(0.005)
Policy Request	0.071***	0.060***	0.060***	0.035***
	(0.007)	(0.006)	(0.006)	(0.006)
Mean DV	0.04	0.04	0.04	0.04
R^2	0.031	0.151	0.151	0.178
Observations	16,275	16,275	16,275	16,275
MC FE		✓	✓	✓
Congress FE			✓	✓
Agency FE				✓

Notes: OLS estimates. SE clustered by MC. Outcome variable is whether the number of sentences citing agency policy information for each member-agency-congress combination is greater than 0. Policy and Casework Requests are dummy variables equal to 1 if for contact and 0 otherwise. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

E Agency Ideology and Campaign Contributions

To produce estimates of agency ideology, I perform five steps. First, I obtain donation data for each election cycle from 1990 to 2022 from the Database on Ideology, Money in Politics, and Elections (DIME) (Bonica 2023). The DIME contains the universe of itemized donations recorded by the Federal Election Commission, with information on the amount donated, the name and employer of the contributor, unique identifiers for the recipients and contributors of the donation, and ideology scores for each contributor and recipient.

Second, since the goal is to extract donations made by federal employees, I extract all the employer entries for a total of more than 9.7 million unique employers.

Third, I use `LinkTransformer` (Arora & Dell 2023) – a deep learning model for record linkage that achieves state-of-the-art performance in several data manipulation tasks – to assign to each employer the agency that best matches the employer string. The advantage of using this model, rather than standard fuzzy matching algorithms, is that it represents strings as numeric vectors (embeddings), thus optimally handling typos, acronyms, and multiple ways in which donors enter their employer's information. For instance, I am able to extract 164 different ways in which EPA bureaucrats report the name of the EPA as their employer.

Each unique employer is assigned an agency name and the model returns a similarity score from 0 to 1. To showcase the strength of this record linkage approach, consider the following employer "fima fema dhs". American government experts would recognize this organization to be the Federal Insurance and Mitigation Administration, housed in the Federal Emergency Management Agency (FEMA), an executive sub-agency of the Department of Homeland Security. The model matches this entry with FEMA, with a similarity of .70.

Fourth, I manually verify and correct the matching for every match with a similarity score greater than .9 (4,546 unique employers), whereas I check matches with a similarity score between .7 and .9 with the assistance of GPT (89,979). I use the OpenAI API and pass each pair of employer-agency matches, and I simply ask the model to decide whether the match is correct or not. Matches whose similarity is below .7 are excluded since they most likely are spurious matches. Eventually, I am able to assign 13,569 employers denominations to 302 agencies in my sample of 317 agencies.[21]

Fifth, I extract the 6,637,563 contributions made by 569,621 unique individuals whose employers match one of the 13,596 agencies-employers. I group all the donations by contributor-presidency level and take the total amount ϕ_{ijt} donated by bureaucrat i working in agency j during presidency t. The choice to group donations at the presidential term level is due to the fact that presidential transitions, by triggering turnover in bureaucracies, are the main source of variation in agency ideology. In addition, by grouping donations at the presidential term level, the final estimates are less sensitive to sampling variability, for more bureaucrats/donors contribute to the agency-presidential term average. The ideology each bureaucrat is given by her CF-score.

To aggregate individual bureaucrats' ideology at the agency level, I follow the approach used by Chen & Johnson (2015) and assign weights to individual bureaucrats' CF-score proportional to the total amount donated by the bureaucrat. This weighting strategy assigns greater importance

[21]The 15 agencies for which I do not find donation records are Bureau of Economics, Bureau of Political-Military Affairs, Employment Standards Administration, European and Eurasian Affairs, National Foundation on the Arts and the Humanities, National Security Education Board, Office of Healthy Homes and Lead Hazard Control, Office of Vocational and Adult Education, Office of Energy Policy and New Uses, Border and Transportation Security Directorate, Office of Electricity Delivery and Energy Reliability, Domestic Nuclear Detection Office, Office of the Federal Coordinator for Alaska Natural Gas Transportation Projects, Board of Directors of the Hope for Homeowners Program, Independent Payment Advisory Board.

to bureaucrats who donate more, since they are likely to earn higher salaries and hence hold more influence and prestigious positions within the agency. The estimated agency ideology x_{jt} is the weighted average of the CF-score of bureaucrats employed by agency j:

$$x_{jt} = \frac{\sum_{i=1}^{n} x_{ijt}\phi_{ijt}}{\sum_{i=1}^{n} \phi_{ijt}}$$

The final dataset includes 2,230 observations covering 302 agencies over 9 presidential administrations. Some agencies do not have 9 observations – one per presidential term – because they were established or terminated during the period covered by the data. The median number of donors with a valid CF-score by agency-presidency is 29. This low number of donors is due to the fact that the list of agencies includes several small organizations. To compute the proportion of bureaucrats who are donors, I obtain data on the number of employees from the Office of Personnel Management (OPM) for the period 2005-2022. I calculate the average number of employees per presidential term, and I compute the share of employees who are donors for each agency-presidency pair. The average share of employees I detect in the DIME database is 9.4% of the total number of employees, and the median is 3.3%. Table E.5 below shows the average number of employees of the Environmental Protection Agency and the number of matched unique contributors.

Table E.5: EPA Employees and Donors.

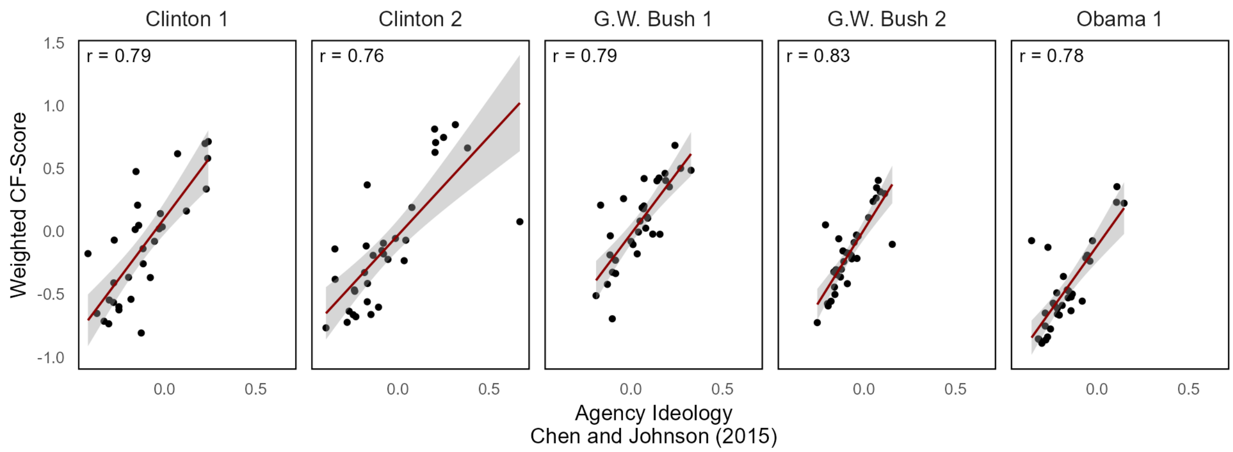
Presidency	**N. Employees**	**N. Donors**	**Share of Donors**
G.W. Bush 2	18,246	997	0.055
Obama 1	18,507	1,242	0.067
Obama 2	15,970	1,235	0.077
Trump	14,680	3,192	0.217
Biden	15,323	1,411	0.092

Notes: Number of employees and number of donors in the Environmental Protection Agency. Data on employees is from OPM FedScope, and data on donations is from DIME.

In the empirical analysis, I focus on the 64 agencies with an average of at least 200 donors per presidency and show that the results are robust to any alternative thresholds from 50 to 500 (see Section F3).

To validate these measures, Figure E.3 shows that the estimates positively correlate with those produced by Chen & Johnson (2015). On the horizontal axis, I plot the estimates from Chen & Johnson (2015), and on the vertical axis, I report the one I computed from the sample of 64 agencies with an average of at least 200 donors by presidency. There is a very strong correlation for each of the five presidential terms covered by both datasets. The correlation remains at similar levels if I include all agencies regardless of the number of donors by presidential term.

Figure E.3: Correlation between agency ideology estimates and estimates from Chen & Johnson (2015).



Notes: The figure displays the relationship between the estimates produced by Chen & Johnson (2015) and the estimates I computed directly from donation records across presidential administrations. Red line represents OLS linear fit of bivariate regression. Correlation coefficients reported on the top-left corner of each panel.

To verify the predictive validity of the measures, I exploit the panel-data structure of the agency ideology and examine whether agency ideology shifts to the right when a Republican president is in power. To do that, I estimate the following model:

$$x_{jt} = \gamma_j + \beta \text{Republican President}_t + \epsilon_{jt}$$

where x_{jt} is the ideology of the agency during presidential term t, γ_j are agency fixed effects, and β estimates the effect of transitioning to a Republican administration on agency ideology. Table E.6 shows the result for all agencies pooled, in Column (1), and for different types of agencies in Columns (2) to (4). As expected, Republican presidents move the ideology of agency to the right. Compared to the average ideology in the data, moving from a Democratic to a Republican administration shits the ideology of agency by more than 60% compared to the mean in the data. This effect is precisely estimated for different types of agencies. The effect is stronger for executive departments under tighter control of the President, but it is sizable for independent agencies too. On the other hand, in executive sub-agencies, mostly populated by careerists and where the reach of presidential appointments is lower, the effect of presidential transitions is noisier but still distinguishable from zero at 90% level.

Table E.6: Effect of Presidential Transitions on Agency Ideology.

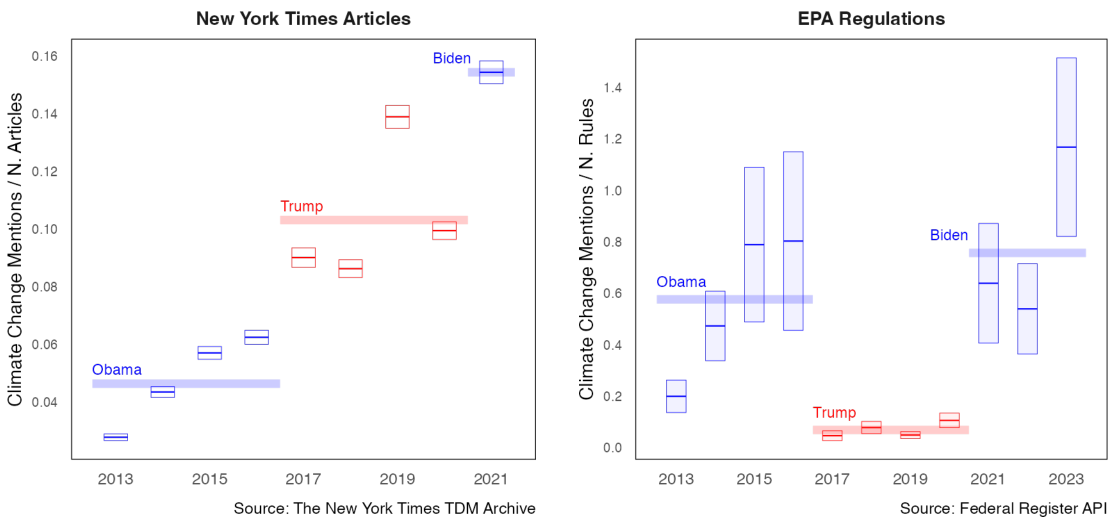
	Agency Ideology (Weighted CF-score)			
Sample:	All Agencies	Independent Agencies	Executive Departments	Executive Sub-agencies
	(1)	(2)	(3)	(4)
Republican Pres.	0.23***	0.28**	0.43***	0.09†
	(0.04)	(0.09)	(0.06)	(0.05)
Mean DV	-0.296	-0.247	-0.199	-0.353
R^2	0.43	0.40	0.46	0.44
Observations	555	131	150	247
Agency FE	✓	✓	✓	✓

Notes: OLS estimates. SE clustered by agency in parenthesis. The dependent variable is agency ideology (weighted CF-score). Unit of observation is at the agency-presidential term level. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

E1 Agency Priorities Across Presidential Administrations: The Case of the EPA

To show how presidential transitions affect agency policy priorities, I estimate how the Trump presidency, by appointing administrators committed to undoing environmental regulations, affected the priorities and actions of the EPA. To do that, I collect the text of all final rules and proposed regulations issued by the EPA between 2013 and 2022, and I count the number of times they contain climate change-related words: climate change, climate protection, climate mitigation, climate adaptation, and global warming. Second, I access the archived corpus of New York Times articles and repeat the same exercise. The figure shows that, while the media paid increased attention to climate change, the regulations of the EPA under Trump contained almost no mention of climate change language. Conversely, the average number of climate change mentions is much higher during the second term of Obama and the first years of the Biden administration. This evidence showcases the importance of studying the dynamic interaction between presidents, Congress, and the bureaucracy.

Figure E.4: Mentions of "Climate Change" in EPA regulations and New York Times articles.



Notes: The panel on the right shows the average number of mentions of climate change-related words in NYT articles over time and across three presidential administrations. Each bar shows the mean and the standard error of the mean.

F Robustness Tests: Agency Ideology

F1 Alternative Measures for the Outcome Variable

Table F.7: Ideological Distance and MCs' Use of Bureaucratic Information: Alternative Outcome Variables.

	# Sentences Citing Policy Information		
	Continuous (1)	Log-transformed (2)	Dichotomous (3)
Ideological Distance	-0.005**	-0.002**	-0.001*
	(0.002)	(0.001)	(0.001)
MC Covariates	✓	✓	✓
Floor/Committee	✓	✓	✓
Mean DV	0.03	0.03	0.018
R^2	0.297	0.279	0.231
Observations	870,419	870,419	870,419
MC-Agency FE	✓	✓	✓
Agency-Congress FE	✓	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Column (1) count, Column (2) dichotomous, Column (3) log-transformed count. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F2 Classifying Policy citations: Alternative Cutoffs

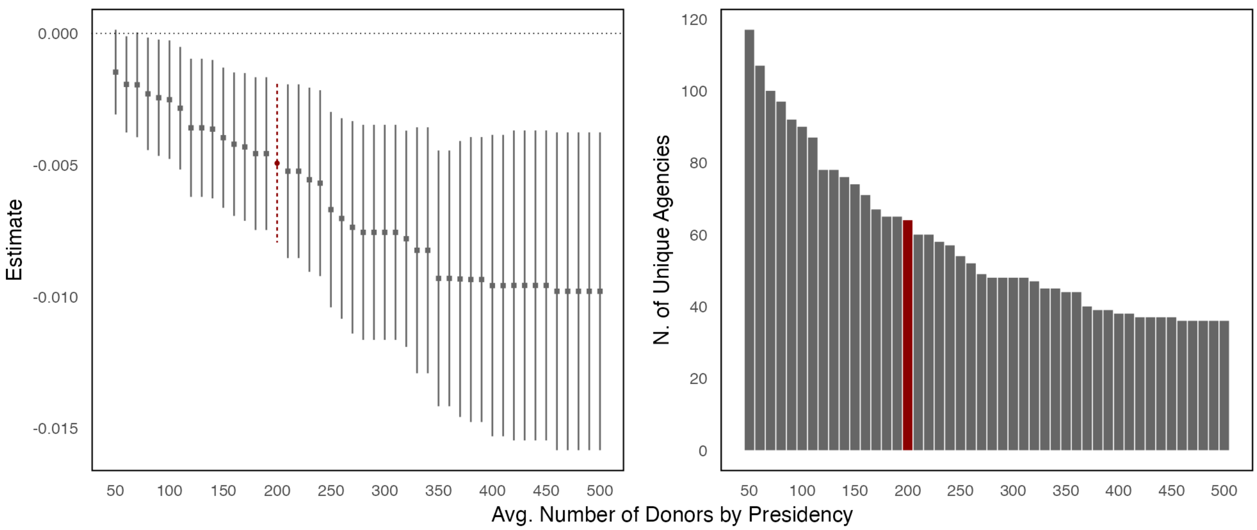
Table F.8: Ideological Distance and Members' Use of Policy Information: Alternative Cutoffs

	# Sentences Citing Policy Information					
	.30	.40	.50	.60	.70	.80
	(1)	(2)	(3)	(4)	(5)	(6)
Ideological Distance	-0.0047**	-0.0051**	-0.0049**	-0.0045**	-0.0045***	-0.0045***
	(0.0018)	(0.0017)	(0.0015)	(0.0014)	(0.0013)	(0.0012)
Floor/Committee	✓	✓	✓	✓	✓	✓
MC Covariates	✓	✓	✓	✓	✓	✓
Mean DV	0.037	0.033	0.030	0.028	0.025	0.021
R^2	0.308	0.300	0.297	0.290	0.284	0.269
Observations	870,419	870,419	870,419	870,419	870,419	870,419
MC-Agency FE	✓	✓	✓	✓	✓	✓
Agency-Congress FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Each column uses a different cutoff to determine whether the citation is classified as a policy citation based on the predicted probability returned by the LDA classifier. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Column (1) replicates Column (5) of Table 4, and Column (2) estimates the heterogeneous effect of ideological distance by venue. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F3 Sensitivity to Sample of Agencies

Figure F.5: Sensitivity of Main Results to the Number of Agencies Included in the Analysis Dataset.



Notes: The figure on the left reports the effect of ideological distance (with 95% confidence intervals) on members' number of citations of agency policy information when including agencies with different average numbers of donors per presidential term, which is reported on the horizontal axis. Coefficient in red is the one used for the main analysis, which includes the 55 agencies whose measure of ideology aggregated on average at least 200 bureaucrats per presidential term. Coefficients are OLS estimates with 95% confidence intervals. SE clustered by MC-agency dyad. Ideological distance is the absolute value of the difference between the ideology of members and agencies. The figure on the right reports the number of unique agencies entering the analysis dataset as a function of the different threshold imposed on the minimum number o average donors per presidential term.

F4 Floor and Committee Heterogeneity

Table F.9: Ideological Distance and Members' Use of Policy Information in Floor and Committees.

	# Sentences Citing Policy Information	
	(1)	(2)
Ideological Distance	-0.0049**	-0.0033*
	(0.0015)	(0.0016)
Floor	0.0348***	0.0375***
	(0.0023)	(0.0026)
Ideological Distance × Floor		-0.0029*
		(0.0012)
MC Covariates	✓	✓
Mean DV	0.03	0.03
R^2	0.297	0.297
Observations	870,419	870,419
MC-Agency FE	✓	✓
Agency-Congress FE	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Column (1) replicates Column (5) of Table 4, and Column (2) estimates the heterogeneous effect of ideological distance by venue. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F5 House and Senate Heterogeneity

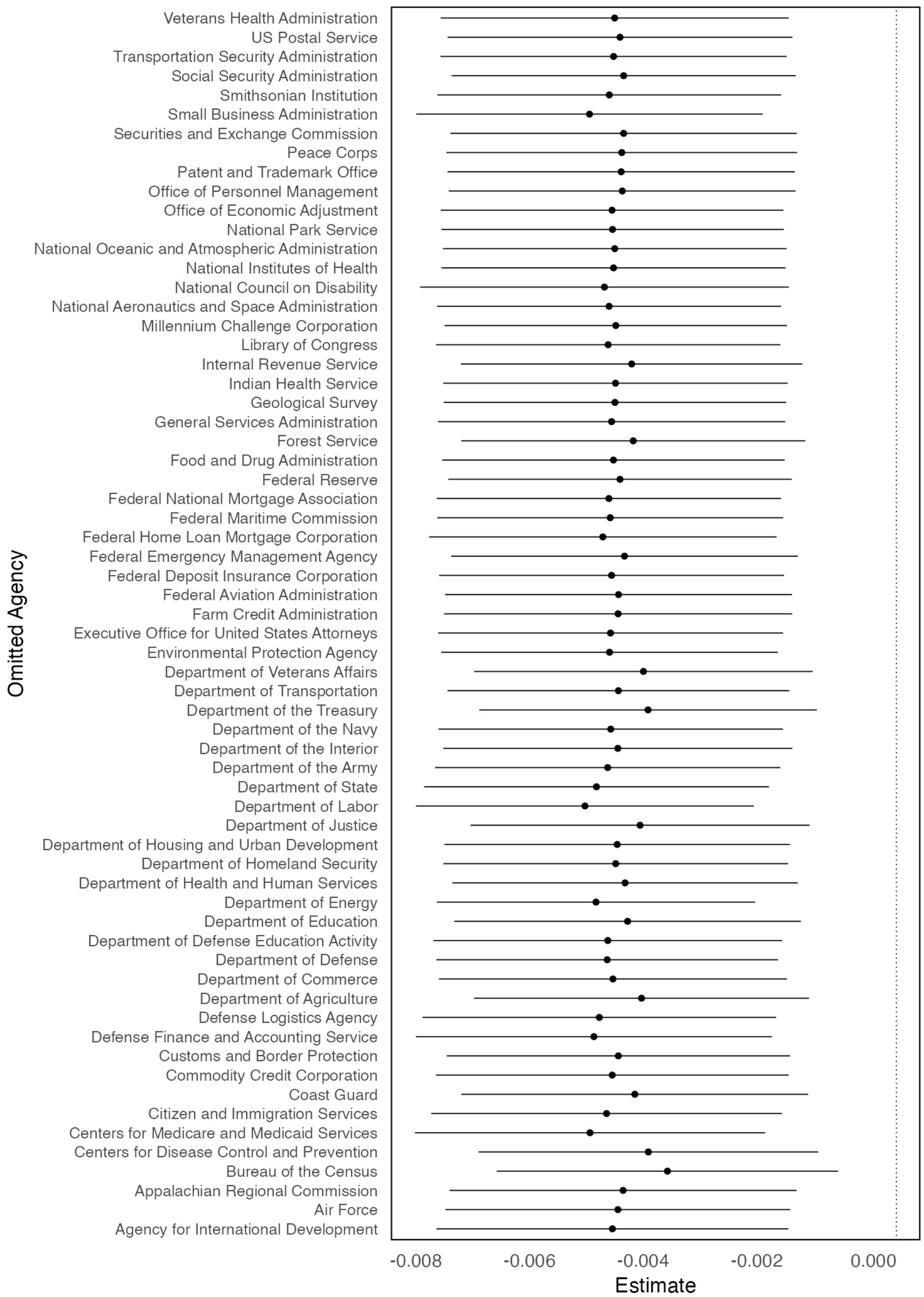
Table F.10: Ideological Distance and Members' Use of Policy Information in the House and Senate.

	# Sentences Citing Policy Information	
Chamber:	House	Senate
	(1)	(2)
Ideological Distance	-0.0023†	-0.0108†
	(0.0012)	(0.0056)
MC Covariates	✓	✓
Floor/Committee	✓	✓
Mean DV	0.0190	0.0760
R^2	0.262	0.356
Observations	696,511	173,908
MC-Agency FE	✓	✓
Agency-Congress FE	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Column (1) shows results for representatives, and Column (2) for senators. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

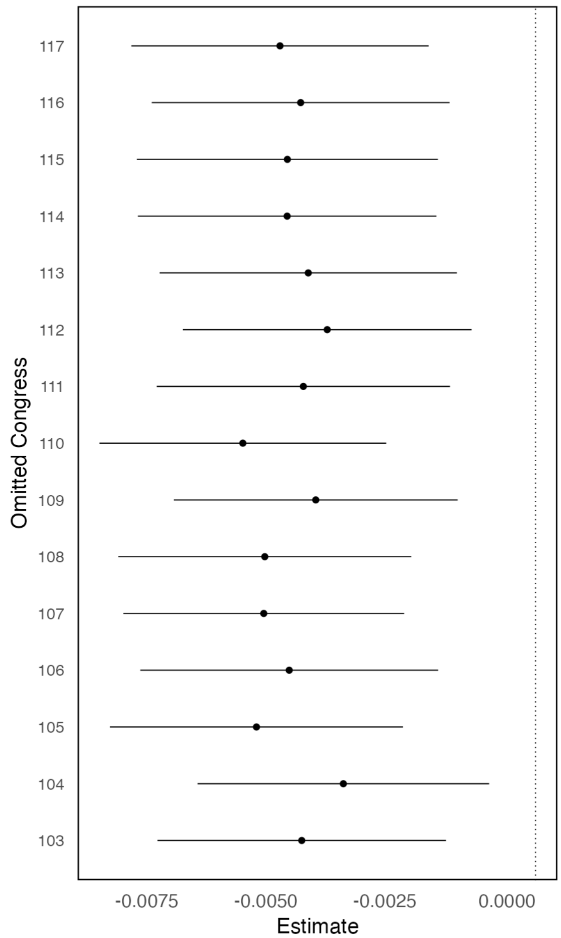
F6 Sensitivity of Results to Individual Agencies and Congresses

Figure F.6: Ideological Distance and MCs' Use of Bureaucratic Information: Dropping one agency at a time.



Notes: OLS estimates and 95% confidence intervals. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200), except for the one sequentially dropped and reported on the vertical axis. Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Data from speeches given on the floor and in committees. Agency names on the vertical axis indicate the agency removed from the analysis dataset.

Figure F.7: Ideological Distance and MCs' Use of Bureaucratic Information: Dropping one Congress at a time.



Notes: OLS estimates and 95% confidence intervals. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Data from speeches given on the floor and in committees. Congress on the vertical axis indicates that Congress was removed from the analysis dataset.

F7 Alternative Measures of Agency Ideology

In Table F.11 below, I show that the main results on the effects of ideological distance on members' reliance on bureaucratic policy information are robust to alternative ways of aggregating bureaucrats' ideology and donations.

- Column (1) uses the same measure used in the main analysis and reported in Table 4. Agency ideology is the weighted average of bureaucrats' CF-score, with weights equal to the total amount donated by bureaucrat i during presidency t.
- Column (2) uses the dynamic measure of member CF-score, namely period-specific estimates of recipient scores re-estimated in each election cycle while holding contributor scores constant.
- Column (3) uses the same measure used in the main analysis but relies on a different weighting strategy. Instead of weighting bureaucrats' ideology by the total dollars donated in each presidential term, it weights by the average amount donated.
- Column (4) weights by the number of donations.
- Column (5) shifts from CF-scores to DW-NOMINATE scores. First, bureaucrats' ideology is computed as the weighted average of the DW-NOMINATE of the recipient, with weights equal to the amount donated to each recipient. Then, agency ideology is the simple average of each bureaucrat's weighted DW-NOMINATE score.
- Column (6) replicates the steps in Column (5) but aggregates bureaucrats' weighted DW-NOMINATE score re-weighting each bureaucrat's score by the total amount donated, mirroring the assumption invoked when building the main measure of agency ideology, namely that better-paid bureaucrats – possibly holding more prestigious positions within the agency – earn and donate more.

One key advantage of using CF-scores compared to DW-NOMINATE scores, is that most recipients do not have an available DW-NOMINATE score and hence those donations cannot be used when computing bureaucrats' (and agencies') estimates.

When using the DW-NOMIANTE version of agency ideology, the measure of ideological distance is the absolute difference between the first dimension of members' DW-NOMINATE and the agencies' estimated ideology as the (weighted) average of bureaucrats' DW-NOMINATE.

Except for Column (5), the results are robust to alternative aggregation and weighting strategies an similar in size to those presented in Table 4. This is an important robustness test, for it shows that two alternative measures of ideology (from donations and roll-call voters) yield similar results.

Table F.11: Alternative Measures of Agency Ideology.

	# Sentences Citing Policy Information					
	(1)	(2)	(3)	(4)	(5)	(6)
Ideological Distance	-0.005**	-0.005**	-0.005**	-0.008***	-0.002	-0.007*
	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
MC Covariates	✓	✓	✓	✓	✓	✓
Floor/Committee	✓	✓	✓	✓	✓	✓
Mean DV	0.030	0.030	0.030	0.030	0.030	0.030
R^2	0.297	0.299	0.297	0.297	0.297	0.297
Observations	870,419	862,982	870,419	870,419	858,573	858,573
MC-Agency FE	✓	✓	✓	✓	✓	✓
Agency-Congress FE	✓	✓	✓	✓	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Each column reports results using a different measure of agency ideology. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F8 Alternative Measures of Policy Information: Quantitative Evidence

One concern recently highlighted in the literature (see, e.g. Barrie et al. 2024) is that GPT-generated output does not allow replicability. The authors find that GPT output varies if prompted multiple times over time and argue that the fact that GPT and other proprietary language models do not allow full transparency and replicability undermines open science.

To address these legitimate concerns, I perform two robustness tests.

First, I show that the task at hand (i.e., classifying policy information in members' statements citing bureaucracies) yields low over-time variability. I extracted a random sample of 300 statements used in the training data and re-classified them with GPT almost 2 months after the first classification. I use the same model and set the same parameters used in the training data stage and find that GPT's output is identical to the one obtained approximately 2 months before for 266 statements (88.7%).

Second, I show that the results are robust to using a fully replicable measurement strategy that detects members' use of quantitative evidence and statistical facts in members' statements citing bureaucratic agencies.

I apply a simple dictionary-based approach to the sentences quoting agencies, whereby every citation is assigned a score capturing the frequency of words belonging to a pre-defined dictionary of statistical facts and quantitative evidence. I use the licensed off-the-shelf LIWC dictionary (Pennebaker et al. 2015), which contains a comprehensive list of words related to quantifiers and numbers, as well as all digits and numbers used to express quantities, which I integrate with verbs capturing quantitative change. In Table F.12, I report the full list of words used to build the dictionary.

Table F.13: Dictionary of Statistical Facts and Quantitative Evidence.

Source	Words
LIWC Dictionary	billion*, doubl*, dozen*, eight*, eleven, fift*, first, five, four*, half, hundred*, infinit*, million*, nine*, onc, one, quarter*, second, seven*, singl, six*, ten, tenth, third, thirt*, thousand*, three, trillion*, twel*, twent*, twice, two, zero, zillion*, add, ad, all, allot, alot, amount, anoth, ani, approximat*, averag, bit, both, bunch, chapter, coupl, each, either, entire*, equal*, everi, extra, few, fewer, fewest, group*, inequal*, least, less, lot, lotof, lotsa, lotta, major, mani, mo, more, most, much, mucho, multipl, nada, none, part, percent*, piec, plenti, remain, sampl*, scarc, scarcer, scarcest, section, segment*, seri, several*, some, somewhat, ton, total, triple*, tripl, varieti, various, whole, (All digits that are not dates)
Verbs (Quantitative Description)	increase, decrease, reduce, boost, lower, decline, skyrocket, eliminate, enhance, rise, limit, accelerate, significantly, plummet, spike, overall, large, face, hurt, harm, end, nurture, criticize, cause, induce, suffer, exacerbate, result, inflict, prevent, worsen, consequence, impact, affect, effect, combat, minimize, maximize, ensure, allow, curb, avoid, curtail, save, mitigate, promote, cultivate, facilitate, create, adopt, sustain, develop, bolster, improve, expand, maintain, restore, intensify, decay, crumble, erode, collapse, evolve, neglect, stop, budget, lose, fund, regulate, provide, discourage, encourage, go, plan

The frequency measures I use are absolute frequency (i.e., the number of words in the citation that belong to the dictionary) and weighted frequency. Specifically, I compute the term-frequency inverse-document-frequency (*tf-idf*), which down-weights (up-weights) words that are in the dictionary but that appear in many (few) citations, for they are less (more) useful at differentiating between citations. I then estimate the same models presented in Table 4 with these alternative measures of the outcome variable. I report the results in Table F.14.

Table F.14: Ideological Distance and Members' Use of Quantitative Evidence.

	# Sentences Citing Quantitative Evidence	
	Absolute Frequency (1)	Weighted Frequency (*tfidf*) (2)
Ideological Distance	-0.014*	-0.018*
	(0.006)	(0.008)
MC Covariates	✓	✓
Floor/Committee	✓	✓
Mean DV	0.117	0.162
R^2	0.292	0.289
Observations	870,419	870,419
MC-Agency FE	✓	✓
Agency-Congress FE	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the sum of the absolute (Col 1), weighted (Col 2), and log-transformed absolute frequency of statistical facts and quantitative evidence in members' citations for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Columns (1) to (3) use data from speeches given on the floor and in committees. Columns (4) and (5) use data from floor and committee speeches separately. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F9 Stance Detection

To ensure that statements citing policy information produced by bureaucratic agencies do not do so while disapproving of the agency, I extract from members' statements citing policy information the stance of the statements. To do so, I classify each statement through a Natural Language Inference (NLI) approach.

NLI leverages pre-trained language models that are designed for entailment classification tasks. Simply put, a document *entails* a statement with some probability. For example, the statement "I fully support the EPA initiative." has a high probability of entailing "The author of this comment *supports* the proposed rule." NLI operates by taking a premise (i.e., the statement) and a set of hypotheses (i.e., statements that are either true or false with some probability given the premise) and identifies the hypothesis with the highest probability of being true based on the comment's text (Burnham et al. 2024). This approach is superior to a simple sentiment analysis, for stance is frequently uncorrelated with sentiment (Bestvater & Monroe 2023). To classify the statements, I used the publicly available and fully transparent DEBATE model (Burnham et al. 2024).

For example, the model returns a probability equal to .99 that the author of the following statement opposes the Department of Energy:

> *"The Department of Energy has proposed changes that cannot, under any stretch of the imagination, be called reforms."*

This exercise allows me to establish two main results. First, only 8% of statements quoting agencies are assigned an opposing stance. Second, the effect of ideological distance is robust to removing these opposing statements from the analysis and counting as statements quoting policy information those whose stance towards the agency is classified as neutral or supportive. In Table F.15 below, I report the results of the effect of ideological distance on the frequency of citations with policy information that are classified as either neutral or supportive.

Table F.15: Ideological Distance and Members' Use of Policy Information: Excluding Citations with Negative Stance.

	# Sentences Citing Policy Information	
	All stances	Neutral and Supporitve Stance
	(1)	(2)
Ideological Distance	-0.005**	-0.005***
	(0.002)	(0.001)
MC Covariates	✓	✓
Floor/Committee	✓	✓
Mean DV	0.030	0.029
R^2	0.297	0.295
Observations	870,419	870,419
MC-Agency FE	✓	✓
Agency-Congress FE	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination (Col 1) and the number of sentences citing agency policy information that are not classified as having an opposing stance. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Columns (1) to (3) use data from speeches given in the floor and in committees. Columns (4) and (5) use data from floor and committee speeches separately. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

F10 Placebo Outcomes: Mentions and (Non-policy) Citations of Agencies

Table F.16: Ideological Distance and Member's Mentions of Agencies.

	Mentions	Citations	Policy Citations
	(1)	(2)	(3)
Ideological Distance	-0.0424	-0.0032	-0.0049**
	(0.0440)	(0.0028)	(0.0015)
MC Covariates	✓	✓	✓
Floor/Committee	✓	✓	✓
Mean DV	1.266	0.066	0.03
R^2	0.395	0.354	0.297
Observations	870,419	870,419	870,419
MC-Agency FE	✓	✓	✓
Agency-Congress FE	✓	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Same sample of agencies used in main analysis (64 agencies with an average number of donors by presidency greater than 200). Outcome variable is the number of mentions (Column 1), citations (Column 2) and policy citations (Column 3) for each member-agency-congress-venue (floor/committee) combination. Ideological distance is the absolute value of the difference between the ideology of members and agencies. Data from speeches given in the floor and in committees. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

G Robustness Tests: Agency Independence

G1 The Role of Ideology for Independent and Non-independent Agencies

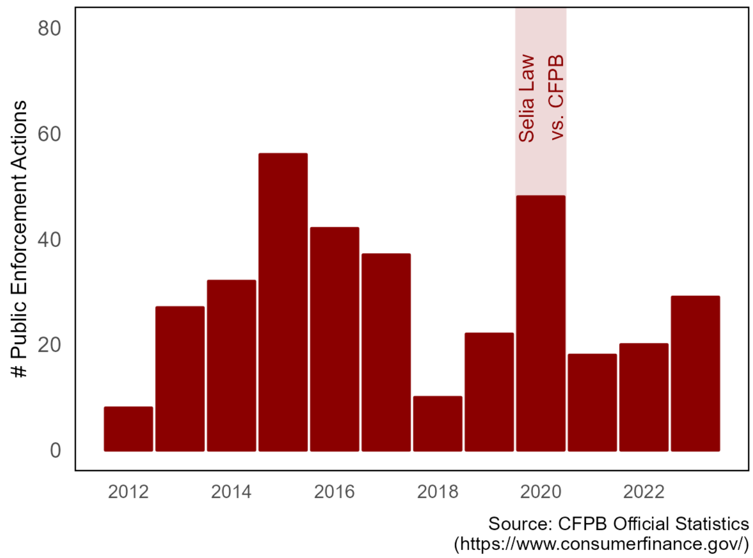
Table G.17: Ideological Distance and Members' Citations of Agency Information: Independent vs Non-independent Agencies

	# Sentences Citing Policy Information		
Agencies:	All	Independent	Non Independent
	(1)	(2)	(3)
Ideological Distance	-0.0049**	-0.0027	-0.0062**
	(0.0015)	(0.0027)	(0.0020)
Floor/Committee	✓	✓	✓
MC Covariates	✓	✓	✓
Mean DV	0.030	0.027	0.034
R^2	0.297	0.264	0.310
Observations	870,419	205,825	612,823
MC-Agency FE	✓	✓	✓
Agency-Congress FE	✓	✓	✓

Notes: OLS estimates. SE clustered by MC-agency dyad. Outcome variable is the number of sentences citing agency policy information for each member-agency-congress-venue (floor/committee) combination. Col (1) includes all agencies, Col (2) includes independent agencies, and Col (3) includes non-independent agencies. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

G2 CFPB Enforcement Actions

Figure G.8: CFPB Enforcement Activities Over Time.



Notes: Number of enforcement cases conduced by the Consumer Financial Protection Bureau over time.

G3 Alternative Measures for the Outcome Variable

Table G.18: Effect of *Seila v CFPB* Case: Alternative Outcome Variables.

	# Sentences Citing Policy Information		
	Continuous (1)	Log-transformed (2)	Dichotomous (3)
CFPB × Post (Trump)	-0.017*	-0.013*	-0.020**
	(0.008)	(0.005)	(0.007)
CFPB × Post (Biden)	0.027	0.012	0.011
	(0.017)	(0.008)	(0.010)
Mean DV	0.006	0.004	0.005
R^2	0.034	0.042	0.045
Observations	186,669	186,669	186,669
MC FE	✓	✓	✓
Time FE	✓	✓	✓
Agency FE	✓	✓	✓

Notes: OLS estimates. SE clustered by member-agency dyad. Outcome variables are different transformations of the number of sentences citing agency policy information for each member-agency-time period combination. Column (1) count, Column (2) dichotomous, Column (3) log-transformed count. CFPB is equal to 1 for member-CFPB pairs. Post (Trump) equals 1 for observations during the Trump presidency and after the ruling. Post (Biden) equals 1 for observations during the Biden presidency after the ruling. Reference time period is Pre (Trump), for observations during the Trump presidency before the ruling. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.

G4 Parallel Trends

To indirectly test that members' reliance on the CFPB's information was on parallel trends before the Court's ruling, estimate the effect of the ruling in three different post-ruling indicators: the real one and two placebo indicators in the pre- and post-treatment, respectively. First, I estimate the effect of the ruling using the actual post-treatment date of June 29, 2020. I subset the data to observations during the Trump presidency alone. Second, I subset the data to observations during the Trump presidency and in the period before the ruling. I then estimate the effect of the ruling using a placebo post-treatment indicator equal to June 29, 2019, one year before the actual date of the ruling. Third, I subset the data to observations during the Biden presidency and estimate the effect of the ruling using a placebo post-treatment indicator equal to June 29, 2021, one year after the actual date of the ruling.

The specification is identical to the one reported in Equation 4 with two notable exceptions. First, I estimate only the difference in members' use of policy information produced by agencies before and after the treatment (2 time periods). Second, I sequentially replace the actual post-treatment indicator with the two placebo indicators described above. Column (1) reports the results from the pre-ruling period. Column (3) reports the results from the post-ruling period. Column (2) reports the results when using the actual post-treatment indicator. I find a negative and statistically significant effect only when using the actual post-treatment indicator. These results suggest that the effects estimated in Table G.18 are due to the ruling and not to CFPB-specific trends.

Table G.19: Effect of *Seila v CFPB* Case: Placebo Post-ruling Indicators.

	# Sentences Citing Policy Information		
Presidency:	Trump		Biden
Post-ruling Date:	2019-06-29 (1)	2020-06-29 (2)	2021-06-29 (3)
CFPB × Post	-0.008 (0.007)	-0.020* (0.008)	0.024† (0.014)
Mean DV	0.004	0.005	0.004
R^2	0.030	0.033	0.034
Observations	592,446	592,446	588,930
MC FE	✓	✓	✓
Post FE	✓	✓	✓
Agency FE	✓	✓	✓

Notes: OLS estimates. SE clustered by member-agency dyad. Outcome variable is the number of sentences citing agency policy information for each member-agency-time period combination. Column (1) uses data from pre-ruling period during the Trump presidency. Column (3) uses data from post-ruling period during the Biden presidency. Column (2) used data from pre- and post-ruling during the Trump presidency. Post (Trump) equals 1 for observations during the Trump presidency and after the ruling. Post equals 1 for observations after the ruling. Signif. codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1.