Is Partisanship Bad for Bureaucratic Accountability?*

Luca Bellodi†

Abstract

Bureaucratic accountability rests on the ability of legislators to obtain information about bureaucracies. Yet how legislators process this information is an understudied topic. I argue that legislators follow partisan motives when updating their beliefs about bureaucracy. I introduce new estimates of partisan beliefs about 336 agencies over 40 years in the UK and the US using natural language processing techniques. I find that beliefs about bureaucracy are on average more positive by 3 percentage points for the governing party, irrespective of partisan congruence and ideological distance between the party and the agency. I show how bias follows from beliefs to behaviour. When co-partisan with the government, legislators' speeches about bureaucratic agencies are less grounded in quantitative evidence and legislators are less likely to oversee ideologically distant agencies. The bias that stems from legislators' partisan alignment with the government may hinder their ability to hold bureaucracy to account.

Word Count: 10,159

*For helpful comments, I am thankful to Tom O'Grady, Colin Provost, David Coen, Ruxandra Serban and seminar participants at the UCL PhD Workshop.

†Bocconi University (DONDENA Center) and UCL Political Science.

1 Introduction

A normative tenet of democratic government is that bureaucracies are accountable to elected legislators. To that end, principals design institutions and procedures to limit agency loss and ensure bureaucracies are held in check (McCubbins, Noll, and Weingast 1987; McCubbins and Schwartz 1984). Yet for accountability to be sustained, legislators need to observe the performance of agencies, update their beliefs, and respond accordingly. The literature is nonetheless silent on how legislators form their beliefs about bureaucracy and there is no evidence that they are truly Bayesian actors who see, update, and act. Work in political psychology has demonstrated how individuals – both voters and legislators – are motivated reasoners, and how their opinions follow directional motives aimed at confirming pre-existing ideological or partisan views (Iyengar et al. 2019; Druckman, Peterson, and Slothuus 2013; Baekgaard et al. 2017). In this paper, I show that partisanship also biases legislators' beliefs about bureaucracy and their ability to hold agencies to account.

Models of bureaucratic accountability and political oversight assume that political principals receive information about agencies and update their beliefs accordingly. Within the traditional principal-agent framework, this can happen in three different ways, either relying on legislators' direct observation, third-party reporting, or bureaucracies' reporting (Lupia and McCubbins 1994). The main issue at stake is therefore how to build institutions whereby political principals can access bureaucracies' private information in an optimal way. Even alternative accounts which move away from the canonical view of accountability aimed at reducing information asymmetries still focus on the information flow between account-giving agencies and account-holding principals (Busuioc and Lodge 2017; Schillemans and Busuioc 2015). What has not been addressed is how this information is processed by legislators.

"Tell me your party, and I'll tell you what you think" is a good summary of empirical and theoretical work on partisan motivated reasoning. Voters and legislators process information using directional motives consistent with pre-existing ideological or partisan preferences (Taber and Lodge 2006; Bartels 2002). Individuals strongly identify with their party-group and therefore make judgements about reality through a partisan lens (Green, Palmquist, and Schickler 2002). In what follows, I present a theory that builds on insights from motivated reasoning and identity theory scholarship to show co-partisanship with the government biases beliefs and behaviour of legislators *vis-à-vis* bureaucratic agencies.

With partisanship operating as an identity-shaping force, majority-party legislators perceive themselves to be members of the same team as the government. Because the executive is responsible for public policies, co-partisans with the government too feel responsible for what the bureaucracy does. This partisan co-membership triggers motivated reasoning, which in turn biases legislator' beliefs about bureaucracy. Partisanship defines political groups, and the process of belief formation is biased by group membership. This form of in-group bias triggers motivated reasoning which in turn biases beliefs about bureaucracy. As a result, other attributes of the bureaucracy such as ideological and partisan distance from legislators play a negligible role in explaining what legislators think about agencies, for they do not trigger in-group bias.

However, legislators might strategically hold biased beliefs to protect their party's image, while in fact holding true, unbiased private beliefs about bureaucracy. In rebuttal of this alternative explanation, I expect partisanship to hinder legislators' account-holding *behaviour* too, both in public and private forums. When co-partisan with the government, legislators change their rhetorical style when they argue about agencies in legislative debates (*public* accountability forum), resorting to a less factual and more politicised style. Furthermore, partisan bias limits legislators' informal oversight activities (*private* accountability forum), for it weakens the traditional relationship between agency-legislator ideological disagreement and oversight activities. All else equal, being co-partisan with the government makes accountability weaker.

Empirically, I introduce new data on partisan beliefs about 336 bureaucratic agencies in the US and the UK. I recover the semantic meaning legislators attach to the agencies they mention in legislative speeches by employing word-embedding techniques. I estimate word vector representations for every agency from local corpora of all the speeches given by Republican and Democratic congresspersons and Labour and Conservative MPs separately and in any given year, producing estimates for each party, agency, and year. I find that beliefs are approximately 3 percentage points more positive when there is partisan alignment between the party and the executive, while I find agency-party ideological distance or partisan congruence to play a negligible role at explaining beliefs. I present a micro-test of the motivated reasoning mechanism comparing how US congresspersons reacted to three major scandals involving three bureaucratic bodies, namely the Federal Management Emergency Agency, the Department of Veterans Affairs, and the Internal Revenues Service. I find that, faced with the same exogenous shocks about the reputation of agencies, co-partisan with the government are between 13-19 percentage points more likely to give a positive statement about the agencies involved in the scandal.

To rule out the alternative explanation that beliefs are strategically biased and privately true, I test whether co-partisanship with the government affects legislators' accountability behaviour. I find that the frequency of statistical facts and evidence-based words in speeches about bureaucracy decreases by about 5% compared to the average value for the US and by 9% compared to the average for the UK. Finally, co-partisans with the government are less likely to request information from ideologically distant agencies compared to when they are at the opposition, with the difference in probability of oversight between co-partisan and opposition member legislators resulting from a one-unit increase in ideological distance accounting to

-0.19 percentage points. These findings represent evidence against strategically-biased beliefs.

While other studies document accountability deficits (Schillemans 2011; Schillemans and Busuioc 2015), this is the first attempt at unveiling a "partisan" obstacle to bureaucratic accountability which is not the product of strategic choices or lack of skills of legislators. I present evidence on the key parts of the chain that leads partisanship to biased beliefs and weakened accountability. Importantly the fact that I find similar effects of partisanship on beliefs and accountability in two countries which embody very different administrative traditions (politicised in the US and neutral civil service in the UK) is convincing evidence of the importance of partisanship for belief formation and accountability practices.

Bringing partisan bias inside the study of bureaucracy has important implications for bureaucratic incentives and for theories of bureaucratic reputation. While we know how oversight institutions shape agencies' incentives to exert effort (Turner 2017) or to make informed policy decisions (Patty and Turner 2020), we do not know how partisan bias distorts such incentive structures. Similarly, while scholars mostly look at how agencies build their unique reputation and its effects on bureaucratic autonomy (Carpenter 2001), or at how agencies respond to reputational threats (Maor and Sulitzeanu-Kenan 2016), more attention should be paid to how reputation (namely beliefs aggregated in some ways) suffers from biases that do not pertain to the political leaning or administrative characteristics of the agency, and how legislators and agencies interact in light of such biases. Finally, understanding how legislators form their beliefs about bureaucracies is central to the effectiveness of accountability mechanisms designed to make sure agencies are kept in check, for partisan bias has the potential to undermine the effective control of bureaucratic policy-making.

2 In-Group Bias and Beliefs about Bureaucracy

In democratic government, the executive is responsible for administering public policies through bureaucratic bodies. Despite varying level of autonomy, bureaucratic agencies respond to the political will of the executive, and a bureaucracy doing a poor job has noticeable implications for the consensus of the incumbent party (James 2011; Nielsen and Baekgaard 2015; James and John 2007). This is consistent with the evidence from the retrospective voting literature (e.g., Erikson 1989), both at the national and local level (De Benedictis-Kessner and Warshaw 2020). Because legislators have a strong partisan identity, they follow partisan reasons when forming beliefs about what the government does or is responsible for, including bureaucratic bodies. Beliefs are more positive when the party is in power, and more negative when the party is at the opposition. This bias, I argue, limits legislators ability to hold agencies to account. There is a vast literature in political science that characterises partisanship as a political identity which is able to affect opinion and behaviour (Bartels 2002; Mason 2015). While Republicans might view positively what is done by their co-partisan president, they would evaluate the same situation under a Democratic president more negatively just for the fact they are not from the same team (Iyengar and Westwood 2015). In social identity theory, this *a priori* defence of somebody's own team is called in-group bias, for group membership is sufficient for automatically biasing individuals' judgements of the merits of the different groups (Tajfel 1987; Mason 2015). Legislators feel emotionally and socially attached to the party, and tend to view more positively what the party is doing or is responsible for. The process of belief formation about bureaucracy follows this dynamic too.

Hypothesis 1: Beliefs about bureaucracy are more positive for the governing party.

In theories of bureaucratic politics, interactions between legislators and agencies are often thought to be driven by the divergence of policy preferences between the two actors (Epstein and O'Halloran 1999). It would be reasonable to expect legislators' beliefs about bureaucracy to be a function of ideological agreement. Yet majority-party legislators, blinded by their ties to the government, will be unresponsive to changes in the partisan or ideological leaning of agencies. As motivated reasoning is triggered by in-group bias, ideological or partisan differences between legislators and the agency are of secondary importance, for they do not alter the co-membership of legislators with the government. Even though ideological or partisan incongruence between agencies and legislators is likely to occur, it is partisan alignment with the government that activates motivated reasoning, and the resulting bias is independent of the ideological or partisan position of each single agency. Partisan identity shapes the way legislators think about everything for which the government is responsible. As a result, the ideological distance and the partisan congruence between parties and agencies play a negligible role at shaping legislators' beliefs about bureaucratic agencies.

Hypothesis 2: Ideological or partisan differences between parties and agencies does not affect beliefs about bureaucratic agencies.

However, distinguishing between legislators' public and private beliefs complicates the picture. It might be possible that, when in power, it is in the interest of legislators to express satisfaction with their own party's achievements, including the performance of the agencies for which the governing party is responsible. Beliefs would therefore be strategically biased, while in fact legislators could privately have an objective view of the agency. Seemingly biased *public* beliefs would have no implications for the effectiveness of bureaucratic accountability, which would be informed by true and objective *private* beliefs. In observational terms, if legislators were not biased, their account-holding behaviour would not change when their party is no longer in power.

In fact, the reason why beliefs are important is because they affect behaviour. For instance, it has been shown theoretically that the desensitisation that stems from partisan bias and motivated reasoning – that is, motivated reasoners have a weaker reaction to information about performance – is detrimental to democratic accountability, for voters' beliefs are less likely to change legislators' incentives for effort (Little, Schnakenberg, and Turner 2021). If we translate this desensitisation mechanism to the relationship between legislators and bureaucracies, we might expect partisanship to hinder legislators' ability to objectively hold agencies to account: all else equal, being co-partisan with the government makes accountability weaker. Although private and public beliefs are of difficult empirical tractability, I can present expectations about observable behavioural implications that would result from beliefs being truly biased. In particular, in support of truly biased private beliefs, I argue that partisan bias weakens both public (spoken account-holding) and private (informal oversight) accountability practices.

Legislatures are a primary accountability forum where elected officials can fulfil their account-holding tasks through questions, interrogations, and criticism about the performance of bureaucracy (Miller 2005). When there is partisan alignment between a politician and the government, the former strongly identifies with the latter and evaluates the work of the bureaucracy – their group co-members – in a biased and more positive fashion. Since such more positive beliefs are not the product of anything factual, when majority party legislators debate about bureaucracy they are less likely to use statistical facts and objective quantitative evidence. Rather than accountability forums, legislative debates become an extension of partisan conflict. Partisanship weakening accountability in legislative speeches is in line with blame-avoidance accounts of legislators' use of information (Nielsen and Baekgaard 2015) and with accountability deficits identified in the literature (Schillemans 2011; Schillemans and Busuioc 2015). One key difference, however, is that the weakened accountability that results from partisan bias is not the result of a strategic choice or the inability of legislators, but it is rather an unintentional consequence of strong partisan identity. In observational terms, partisan motives affect legislators' argumentative style, which follows political rather than objective standards, making less frequent use of statistical facts and evidence. This partisanship-induced politicisation of what legislators say about agencies is detrimental to the ability of legislatures to effectively function as accountability forums.

Hypothesis 3: Government-legislators partisan alignment makes political speeches about bureaucracy less grounded in statistical facts and evidence.

Another accountability forum consists of political principals directly overseeing agencies. According to classical theories of bureaucratic accountability (McCubbins and Schwartz 1984), political oversight is more likely to happen when legislators and agencies are far apart on the ideological space. This proposition is the hallmark of bureaucratic accountability, for it entails that, faced with limited resources, elected legislators pay more attention to the activity of agencies that are ideologically far from their policy preferences, and hence deserve tighter control. Yet this relationship is not immune to partisan bias. Being co-partisan with the executive makes legislators less attentive to what agencies do and therefore inhibit their ability to hold bureaucracies to account by informally requesting information from agencies. Although recent work has found that, against common wisdom, legislators' individual oversight is mostly driven by policy valence rather than ideological disagreement (Lowande 2018), whatever role ideology plays in explaining oversight, it will be weaker for co-partisans with the government.

Hypothesis 4: Co-partisans with the government are less likely to oversee ideologically distant agencies.

Moving from beliefs to public and private accountability forums, this theory links partisan motivated reasoning with theories of bureaucratic accountability. When there is partisan alignment between the government and legislators, beliefs about agencies are more positive. This has implications for bureaucratic accountability: argumentative style is less grounded in statistical facts and evidence, and informal oversight is weaker.

3 Data

3.1 Beliefs about Bureaucracy

One obvious way to measure legislators' beliefs about bureaucracy would be to ask them directly what they think about agencies. Yet this approach would result in a snapshot of what *current* legislators think about agencies *today*, and it would not be possible to observe how beliefs change based on co-partisanship with the government. To capture party, agency, and time variation, I measure partisan beliefs about bureaucratic agencies from what legislators say in parliament and produce party-agency-year estimates of beliefs about 336 government departments and bureaucratic agencies (197 for the US and 139 for the UK) over a time frame of approximately 40 years. Legislative speeches can in fact aptly capture what legislators think about bureaucracy while allowing for a large time-coverage.[1]

I measure beliefs at party level from all speeches given by each party in any given year. I create estimates on a unidimensional scale, with larger values signifying more positive beliefs. The measurement strategy builds on the strategy presented by Bellodi (n.d.), who uses word-embedding models to produce reputation estimates for bureaucratic bodies in the

[1] Full lists of agencies as well as agency-type are reported in Section A.1 in the Appendix (see Tables A.1 and A.2).

US and the UK. Because reputation is defined as a "set of symbolic beliefs" (Carpenter 2001, 3–4), this measurement strategy is particularly suitable for estimating partisan beliefs about bureaucracy. Here I use a similar approach, but estimate embeddings separately from local corpora of speeches given by the main parties in each year and country. In particular, I estimate party p's beliefs about agency a by pooling all speeches given by legislators of party p in year t. The output of this strategy is a set of estimates of party p's beliefs about all agencies that are mentioned in the speeches given by each party in any given year. While this measurement strategy does not allow me to estimate the beliefs of each single legislator, it leaves me with sufficiently large corpora to estimate word embeddings for each year and each party separately, capturing variation along party and time.[2]

The intuition behind word-embedding models is that we can learn about the relationship between words and discriminate between words related to one word but not another by looking at the ratio of co-occurrence probabilities. Suppose we want to learn the relationship between the words FED and EPA in a year when the EPA is highly criticised and the FED is praised for its independent policy-making. To do so, we compare the probabilities of these two words co-occurring within a pre-defined segment of text with various context words k. We might expect word $k = \textit{independence}$ to co-occur with the word FED more often than with the word EPA, and word $k = \textit{critic}$ to co-occur more often with the word EPA than with the word FED. Similarly, we expect the word $k = \textit{horse}$ to be very unlikely to co-occur with any of the words, and the word $k = \textit{decision}$ with both. Table 1 represents these expectations in terms of hypothetical probabilities.

The ratio of co-occurrence probabilities for words related to FED is large, whereas for words related to EPA is low. Words related to both or neither have a ratio that approx-

[2]Estimating word embeddings for each legislator in each year would result in very small corpora and hence highly unstable embeddings. Different estimations would yield very different results. For instance, the median number of speeches per year given by US legislators is 60, and the median length is just 60 words.

Probability	$\boldsymbol{k}$ **= independence**	$\boldsymbol{k}$ **= critic**	$\boldsymbol{k}$ **= horse**	$\boldsymbol{k}$ **= decision**
$Pr(k\|FED)$	.1	.01	.001	.15
$Pr(k\|EPA)$	.01	.1	.001	.15
$Pr(k\|FED)/Pr(k\|EPA)$	$.1/.01 = 10$	$.01/.1 = 0.1$	$.001/.001 = 1$	$.15/.15 = 1$

Table 1: Example of co-occurrence probabilities for target words FED and EPA with related and unrelated context words. Only in the ratio does noise from non-discriminative words like horse and decision cancels out, so that large values correlate well with words associated with FED, and small values correlate well with words associated with EPA.

imates 1, because they do not help discriminate between which word is related to which. Co-occurrence ratios are therefore very powerful at encoding relevant semantic relations and at estimating the semantic similarity between terms. These probabilities can be calculated for each unique word that appears in a corpus of texts. Every word can therefore be represented as a real-value vector that encodes the semantic characteristic of the word. Slightly more formally, word embeddings are the coefficients of statistical models (i.e., neural networks) that capture the relationship between the ratio of co-occurrence probabilities. Like the genetic information encoded in a strand of DNA, the elements of such vectors carry semantic information about the word. Words that co-occur with similar words will therefore have a similar DNA and hence more similar vector representations. Distances between these vectors are informative about the semantic similarity of the words as used in the corpus from which they have been estimated Pennington, Socher, and Manning (2014). Therefore, by comparing the word embeddings of EPA and FED with a vector that combines several clearly positive terms such as "good," "excellent," "great," and so on, we can learn about the similarity between the EPA vector and a positivity vector. Similarly, if we do this separately for Democrats and Republicans, we can see how different these agency embeddings are from such positivity vectors.

Once I have word embeddings for every word in V, I can exploit the arithmetic properties of vector representations of words and build a vector that combines some unambiguously positive and negative words that will act as benchmark to measure the partisan belief about the agency. By deducting clearly negative embeddings from the sum of clearly positive embeddings, I obtain a word vector that captures positivity. The specific word vectors I used are:[3]

$$\begin{aligned} \vec{positivity} = \vec{great} + \vec{excellent} + \vec{successful} + \vec{effective} \\ - \vec{bad} - \vec{poor} - \vec{negative} - \vec{terrible} \end{aligned} \tag{1}$$

where the arrows signify the words are vectors. I finally measure the cosine similarity between the word embeddings of each agency and the $\vec{positivity}$ vector. The beliefs score will thus be the angular distance between the two embeddings. For instance, if the embedding $\vec{FED}$ is semantically very similar to $\vec{p}$, beliefs about the FED will be more positive, whereas if the $\vec{EPA}$ embedding is semantically distant, beliefs will be more negative. The resulting metric is normalised to take up values between 0 and 1, where greater values signify more positive beliefs.

I apply this method to a corpus of all floor speeches given by the two main parties between 1981-2016 for the US and 1980-2018 for the UK. Detailed information about speeches and the sample of agencies is reported in the Appendix (see Sections A.1 and A.2). First, I merge all the speeches at party level. Then I split them by year and obtain 86 local party-year corpora for the US and 78 for the UK. I then estimate word embeddings from each single local corpus. By doing so, I allow the semantic meaning of words to vary across country, party, and over time, for every estimation is performed on a different country-party-year corpus. The key advantage of this flexible estimation is that, for instance, the positivity vectors are allowed to vary based on the language used by each party in any given year. The final dataset

[3]The words have been chosen arbitrarily among clearly positive and negative words whose meaning is the same in both countries and did not change over time. This is similar to the seed words chosen by Rice and Zorn (2019) to set the benchmark for positivity and negativity dictionaries.

consists of 9,496 party-agency-year observations, 6,874 for the US and 2,622 for the UK.[4]

3.2 Weaker Accountability: Argumentative Style and Informal Oversight

I am interested in two types of accountability behaviour: legislators arguing about bureaucracies (public accountability forum), and legislators overseeing bureaucracies (private accountability forum).

I measure legislators' use of statistical facts and evidence when arguing about bureaucracy through a targeted dictionary-based analysis of legislative speeches, focusing on sections of text near the name of the agencies. Legislative speeches are assigned a score capturing the frequency with which words contained in a pre-defined list appear in the text. I use the LIWC dictionary (Pennebaker et al. 2015), which contains a comprehensive list of words related to quantifiers and numbers, such as "amount," "average," "equal," "less," "percentage," "twice," "total," as well as all digits and numbers used to express quantities.[5]

Dictionary-based approaches are deemed to be highly context-dependent and therefore need careful validation (Grimmer and Stewart 2013). To this end, the "fact-dictionary" derived from the LIWC list of words has been extensively and successfully validated by Hargrave and Blumenau (2020) in an almost identical setting as the one I study here: legislative speeches in the UK House of Commons. Furthermore, contrary to sentiment analysis tasks – where the meaning of words is likely to change across domains and over time (Rice and Zorn 2019) – words pertaining to statistical facts and quantitative evidence should be more representative of objective attributes and hence less dependent of the context in which are used.

[4]Some agencies will clearly be mentioned less often than others. To make sure results are not driven by agencies mentioned very few times, in Table A.8 in the Appendix I replicate the analysis on a restricted sample of observations where the number of mentions an agency receives in speeches from both parties is above the median.

[5]The full list of words is reported in Table A.4 in the Appendix.

To support this claim, in Section A.5 of the Appendix I report the results of an additional validation test which shows that the LIWC dictionary performs well at matching a manually labelled corpus of texts from different contexts. I find a positive and significant correlation between manual labels and the estimates of the dictionary method, thus strengthening our confidence in the low context-dependence of the dictionary (see Table A.3).

I estimate the use of facts and evidence for more than 500,000 speeches (196,689 for the UK and 326,357 for the US) given by a total of 3,842 different legislators. Text pre-processing steps are reported in Section A.4 of the Appendix. To ensure the analysis is performed over segments of text which are about the agencies, I limit the analysis to various symmetric windows of words centred around the names of the agencies, namely to segments of text that are 20 and 50 words before and after any name of agencies. Speeches can be long and about several topics. By looking at word usage within small segments of text around agency names I increase the likelihood that what legislators are saying is in fact about bureaucracy.[6]

I then compute for every term in each speech the term frequency-inverse document frequency (*tf-idf*), in order to reduce the importance of words that appear very often and in many speeches (Welbers, Van Atteveldt, and Benoit 2017). More precisely, tokens are assigned a weight which is equal to the logarithm of the inverse fraction of the speeches that contain the word. For instance, let us consider the words "approximately" and "less" which belong to the dictionary. If "less" appears in more speeches than "approximately," then "less" will receive a lower weight, for it is less helpful in discriminating between which word is more strongly representing the use of facts and evidence. For each speech, the final score is the sum of the *tf-idf* frequencies of tokens that appear in the dictionary.[7]

More formally, consider the full corpus a set of speeches, and each speech a set of words,

[6]As a falsification test, I also look at the speeches as a whole and find no effects of legislator-government partisan alignment on argumentative style (see Table 5).

[7]Results are robust to using the absolute frequency of facts-words appearing in the window of words. See Table A.13 in the Appendix.

whose cardinality represents the number of unique words in the speech. For each speech mentioning a bureaucratic agency, the use of facts and evidence is given by the following formula:

$$Fact_s = \sum_{t \in Dict} tf - idf_{t,s} \quad \text{with} \quad tf - idf_{t,s} = \frac{f_{t,s}}{|s|} \times log \frac{|S|}{|\{s \in S : t \in Dict\}|} \qquad (2)$$

where t is each token within the pre-defined windows of words for speech s, $Dict$ the list of words capturing the use of facts and statistical evidence, and $tf - idf$ is the term frequency-inverse document frequency of token t in speech s. The $Fact$ score is ultimately a function of the absolute frequency of the token t ($f_{t,s}$), the number of words in speech s ($|s|$), the number of speeches of the total corpus S, and the number of documents in the corpus that contain the token t ($|\{s \in S : t \in Dict\}|$).

As far as bureaucratic oversight is concerned, I use data on informal oversight collected by Lowande (2018), who obtained records of correspondence between 16 bureaucratic agencies and members of Congress during the 110th and 111th Congress. The measure of oversight is dichotomous, and equals one if legislator i during Congress t sent a request or an inquiry to the agency. Legislators' requests are classified into two different categories, depending on the subject of the request. *Casework* requests are those sent on behalf of constituents with particular grievances. *Policy* request are concerns about the substance of policy or programme implementation. Informal oversight is particularly suitable for the purpose of the test I present below, for it is the most private and behind-the-scene form of oversight legislators could make, and should therefore be insulated from the strategic logic of the public sphere.

3.3 Agency Partisanship & Ideology

To test expectations about the negligible role of agency attributes for beliefs, I need a measure of agency partisanship and ideology. While many political commentators, particularly in the US, easily infer the partisanship and the ideology of bureaucratic agencies, researchers face many challenges when conceptualising and operationalising measures of latent agency attributes. There are three challenges to such empirical enterprise (for a critical perspective, see Carpenter (2020)). First, data is hard to collect. While individual-level data on partisan affiliation or political preferences abound, these cannot always be mapped back at organisational level. Second, if data exists, its validity to infer partisan identification and ideological leaning should be carefully assessed. Third, if individual-level data is available and valid, it should be meaningfully aggregated at organisation level.

The US context and the availability of campaign contributions data provides a good solutions to these three issues. First, Campaign Finance Data from the US Federal Election Commission repository (see fec.gov) allows to track donations made from agencies' employees. Second, donation-based measures are regularly used in empirical political science and have been proved to be valid measures of political preferences (Bonica 2019). Third, existing measures of agency partisanship and ideology are commonly used in applied work, averaging individual-level data while taking into account the different rank of political donors (as proxied by the amount of the donation). Unfortunately, it is not possible to build such measures of agency partisanship and ideology for UK bureaucratic bodies, for no data on the employer of political donors is available in the UK.

I download raw bulk Campaign Finance Data from the Federal Election Commission repository for the bienniums from 1999/2000 to 2019/2020. The FEC data and metadata allows to compile a dataset at individual-donation level with information on the amount, beneficiary, and employer of contributors. I subset donations made by one of the 285 agencies in my initial sample of US bureaucracies to republican or democratic beneficiaries, for a total of approximately 11 million donations made from employees of 112 bureaucratic bodies. I build a measure of agency partisanship as the weighted share of republican donations, with weights equal to the amount of the donations. The underlying assumption of this weighting strategy, as used in other measures of agency attributes (see e.g., Chen and Johnson 2014), is that higher-ranked bureaucrats have a larger weight on the decisions of the agency and, because they have larger salaries, they make larger donations.

More formally, for agency i and biennium t, I estimate πREP_{it}, namely the percentage of republican donations, weighted by the amount of the donation ($\pi REP_{it} \in [0,1]$). To build a measure of partisan alignment between the party and the agency, I then use the following assignment function

$$\text{Party-Agency Partisan Alignment}_{it} = \begin{cases} \pi REP_{it} & \text{for the Republican Party} \\ 1 - \pi REP_{it} & \text{for the Democratic Party} \end{cases}$$

so that partisan alignment between the Republican party and the agency is equal to the weighted average of the donations to republican beneficiaries, whereas partisan alignment between the Democratic party and the agency is equal to the complementary percentage.[8]

The overall average number of donations per biennium across all the agencies is 3,591. The Central Intelligence Agency is the agency with the largest number of donations per biennium, with an average of 58,978 donations worth \$35.6 million per biennium. The agency with the largest average amount of donations is nonetheless the Office of Management and Budget, with an average amount of donations equal to \$91.5 million. Among the largest bureaucracies, the Air Force is the most republican department, whereas the most democratic

[8]By coding donations to Republican beneficiaries 1 and to Democratic beneficiaries 0, the mean of the party of the beneficiaries is equal to the share of donations to Republican parties. I weight this average with the amount of each donations.

are the Veterans Health Administration and the Office of Management and Budget (average weighted share of republican donations equal to .64, .35, and .28, respectively).

The measure of agency ideology is built on the same data but it makes a step further. I use Chen and Johnson (2014) donation-based estimates of agency ideology, which match the donations to the ideal point of the beneficiary, measured with the DW-NOMINATE scores of congresspersons receiving the donations. The dataset produced by Chen and Johnson (2014) covers 79 federal agencies across five presidencies, from the first Clinton Presidency to the first Obama Presidency. 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, Figueiredo, and Lewis 2019), and rule-making (Ellig and Conover 2014; Potter 2019). Once I have data on agency ideal point, I build a measure of ideological distance between the agency and the party or legislator by taking the absolute value of the difference between the two actors' ideal points. DW-NOMINATE scores for legislators are obtained from (J. Lewis et al. 2020). For the ideal point of the Democratic and Republican parties, I take the median ideal point of Democratic and Republican legislators.

4 Methods

To test the empirical expectations I present three tests. First, I exploit within-party variation and use a two-way fixed effects estimator to estimate the effect of partisan alignment on beliefs in both the US and the UK. Second, I show how variation in the party-agency partisan congruence or ideological distance is not associated with variation in beliefs. Third, I compare legislators' use of statistical facts and evidence when arguing about bureaucracy between co-partisans with the government and opposition members. Finally, I replicate the research design in Lowande (2018) and estimate the moderating effect of partisan alignment on the relationship between ideological disagreement and oversight.

First, I estimate the following model for both the US and the UK:

$$\text{Beliefs}_{p[a,t]} = \delta_p + \phi_a + \alpha_t + \beta\text{Party-Govt. Alignment}_{p[t]} + u_p \tag{3}$$

where $Beliefs_{p[a,t]}$ represents beliefs among party p, about agency a, and in year t. *Party-Govt. Alignment*$_{p[t]}$ is a dummy variable indicating whether there is party-government alignment in year t, ϕ_a are agency fixed effects to account for all time-invariant agency characteristics, δ_p are party fixed effects, and α_t are year fixed effects to account for common shocks. I then progressively add agency $\times$ year and party $\times$ agency to account differences in party attention to agencies and in time-changing agency characteristics (e.g., agency salience). Because treatment assignment is at the party-election level, I cluster standard errors at party-general election level for the UK and at party-congress level for the US.[9]

These specifications have a series of advantages. First and most importantly, agency dummies account for all time-invariant agency characteristics such as the history, culture, mission, statutory features, and policy sector of the agency, which might well confound the relationship between partisan alignment and beliefs. Some agencies might in fact perform functions or be in charge of policies more likely to be associated with more positive beliefs. Second, because different party might have systematically different beliefs, and these can change across agencies, party and party $\times$ agency dummies account for different party characteristics, organisation and views about the state and the various bureaucratic agencies. Third, by estimating agency $\times$ year fixed effects, I can sweep out the confounding effect of agency saliency. In fact, partisan biases can be exacerbated when agencies are under criticism or praises.

[9] In Tables A.5 and A.6 in the Appendix I show results are robust when clustering SE at party-prime minister (for the UK) and party-presidency level (for the US).

In the US, because of high levels of turnover in agency staff as a result of a new presidency, the effect of partisan alignment on beliefs might be confounded by a change in agency ideological and partisan leaning. Republican presidents might appoint conservative bureaucrats and Democrats' beliefs about the agency might decrease for reasons unrelated to their opposition status. If this were the case, my main hypothesis for which legislators follow partisan motives when forming their beliefs about bureaucratic agencies would be rejected. To test this, I add to Model (3) above the measure of the distance between the agency and the median ideal point of Republican and Democratic legislators and the measure of party-agency partisan congruence.

The second test moves from party-level to legislator-level analysis, and looks at the within-legislator change in argumentative style when arguing about bureaucracy as a result of being co-partisan with the government. I model the measure of use of facts and evidence as a function of partisan alignment using a two-way fixed effects estimator, in order to look at change in the use of facts within legislators and legislative debate (i.e., date) and therefore holding constant all unobserved sources of heterogeneity at the legislator and debate level. Since language can be correlated with specific agencies, I also include agency dummies to account for time-invariant agency-specific characteristics. Formally, I estimate the following model:

$$\text{Facts}_{i[l,a,d]} = \eta_l + \phi_a + \alpha_d + \beta\text{Leg-Govt. Alignment}_{l[d]} + \mathbf{X}'_{i[l,d]}\lambda + u_i \tag{4}$$

where $Facts_{i[l,a,d]}$ is the frequency of tokens considered facts and evidence in speech i given by legislator l about agency a in day d. η_l and ϕ_a are legislator and agency dummies, α_d date fixed effects, $\mathbf{X}'_{i[l,d]}$ a vector of covariates at legislator and speech level, and β the effect of being a co-partisan with the executive.

Finally, I estimate the effect of partisan alignment on bureaucratic oversight activities with a generalised difference-in-differences model with a multiplicative interaction term between ideological distance and partisan alignment. I follow Lowande (2018) and exploit the transition from the second Bush presidency to the first Obama presidency as source of variation in the ideology of the agency and estimate the causal effect of ideological distance with data consisting of only the 110th and 111th Congress. The appealing feature of this design is that it allows to isolate the variation in ideological distance resulting from a change in the presidency while committees' control and the internal organisation of the House and Senate remained fixed because Democrats maintained the majority in both chambers. In fact, the changing roles of legislators as a result of a new Congress might affect legislators' incentives to oversee agencies. I then build the following linear probability model:

$$\begin{aligned}\text{Oversight}_{l[a,c]} = \eta_l + \phi_a + \alpha_c + \beta_1 \text{Leg.-Agency Id. Dist.}_{l[a,c]} + \beta_2 \text{Leg-Govt. Alignment}_{l[c]} \\ + \beta_3 \text{Leg.-Agency Id. Dist.} \times \text{Leg-Govt. Alignment} + \mathbf{X}'\lambda + u_l \qquad (5)\end{aligned}$$

where $Oversight_{l[a,c]}$ is equal to 1 if legislator l in Congress c sent a request to agency a. The model includes agency, legislator, and congress dummies. I follow the original specification of Lowande (2018) and include both legislator- and agency-level time-varying covariates ($\mathbf{X}'$): the seniority of the legislator, whether the legislator is a member, chair, or ranking minority member of an oversight committee, agency budget, politicisation and logged number of employees. The continuous treatment *Leg.-Agency Id. Dist.*, which represents the ideological distance between legislator l and agency a, is the classical predictor of bureaucratic oversight. Here I am interested in estimating whether partisan alignment biases the oversight activities of legislators, which should be a function of ideological distance. The causal quantity of interest is therefore β_3, which accounts for the moderating effect of legislator-government

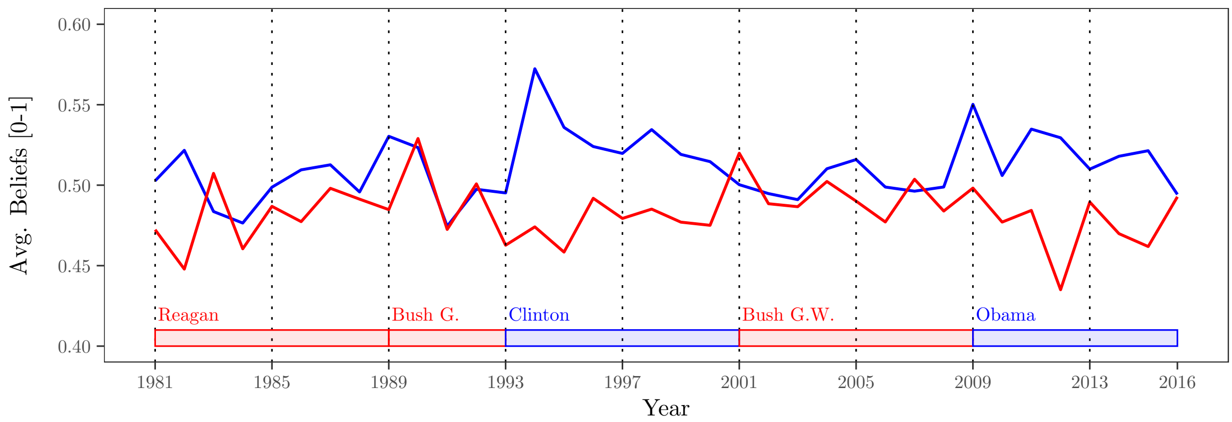

Figure 1: Parties' beliefs averaged across all agencies over time and across presidencies. Red for the Republican and blue for the Democratic party.

partisan alignment on the relationship between ideological distance and oversight.

5 Results

By just looking at Figures 1 and 2, it is clear how partisanship matters for beliefs. The figures plot the political parties' beliefs averaged across all agencies for the Democratic and Republican parties for the US, and for the Conservative and Labour parties for the UK, together with the party in government (horizontal bar at the bottom of the plots).

On average, when there is a Democratic President, the beliefs of the Democratic party are more positive compared to when there is a Republican President, and *vice versa* for the Republican party's beliefs. The UK shows a similar trend. When the government changes colour, partisan beliefs change too. The Conservative party's beliefs about bureaucracy are more positive during the Cameron and May governments than during the Blair and Brown premiership. The Labour party's beliefs too, despite being on avera more positive than the Conservatives', follow government cycles, more positive under Labour governments, more negative under Conservative governments.

Even though I am interested in within-party change in beliefs as a result of being co-partisan with the government, the figures also show high levels of partisan polarisation with

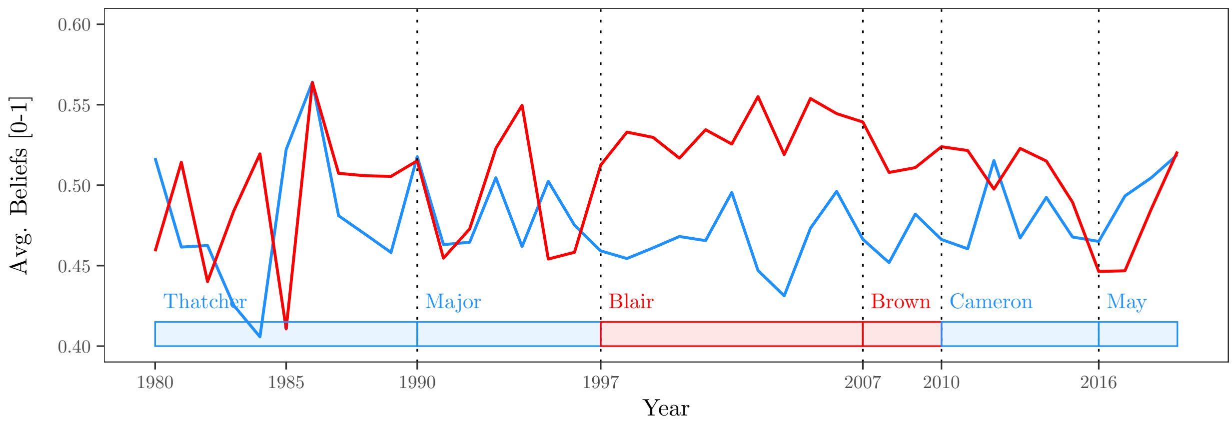

Figure 2: Parties' beliefs averaged across all agencies over time and across governments. Red for the Labour and light blue for the Conservative party.

respect to the bureaucracy (McCarty, Poole, and Rosenthal 2006). Partisan differences in beliefs are more pronounced under Democratic presidents in the US, where the gap between Democrats' and Republicans' beliefs widens. In the UK, except for the golden era of the New Public Management reforms of the mid-80s, when bureaucracy was considered a neutral body of experts (Hood 1991), partisan beliefs remain highly polarised, particularly during the Civil Service Reform Programme of Blair's New Labour government.

Who is in charge of government matters for the process of belief formation. For both the US and the UK, the party-government partisan alignment has a significant effect on beliefs about bureaucratic bodies. In Table 2 I report the results of the first test for the UK. Sample size is large and estimates are highly precise even when estimating party-agency and year-agency fixed effects, and therefore accounting for differences in party attention to agencies and in time-changing agency characteristics. Because beliefs are measured from 0 to 1, the effects of partisan alignment can be interpreted as changes in percentage points. All else being equal, partisan alignment is associated with beliefs on average 3 percentage points more positive, as large as 0.21 times the standard deviation of beliefs in the sample. Similarly, the effect is as large as the average difference in beliefs between the Conservative and Labour parties (i.e.,0.03).

The US data allows for an even more powerful test, for I can adjudicate not only between

DV:	Beliefs [0,1]		
Country:	UK		
Model:	(1)	(2)	(3)
Variables			
Party-Govt. Partisan Align.	0.027***	0.027***	0.031**
	(0.004)	(0.006)	(0.009)
Fixed-effects			
Party	✓	✓	
Year	✓		
Agency	✓		
Agency-Year		✓	
Party-Agency			✓
Year-Agency			✓
Fit statistics			
Observations	2,622	2,622	2,622
R^2	0.257	0.594	0.652
Within R^2	0.009	0.017	0.022

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

Table 2: Partisanship and Beliefs, UK Data. OLS estimates. Units are party-agency-year observations.

party-government and party-agency characteristics. In fact, because party-agency partisan or ideological alignment do not trigger in-group bias, I do not expect them to affect beliefs.[10]

In Table 3 I report the results. The effect of partisan alignment between the party and the government are positive, precisely estimated, and in the expected direction. Being aligned with the president is associated with an increase in beliefs about bureaucracy by 2-3 percentage points, depending on the specification. Importantly, the effect holds when conditioning on the ideological distance and the partisan congruence between the agency and the party (Models (3) and (4)). This lends support to *Hypothesis 1* and *Hypothesis 2*. Partisan alignment with the government changes beliefs, and ideological differences have a

[10]The sample of agencies for which there is available data for beliefs, party-agency ideological distance and partisan alignment are 197 out of the 197 agencies for which I produce beliefs estimates. In Table A.7 in the Appendix I show how the results hold when limiting the analysis to the sample of agencies for which all the three variables are available.

DV:	Beliefs [0,1]					
Country:	US					
	Party-Government				Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Party-Govt. Partisan Align.	0.018***	0.019***	0.030***	0.011*		
	(0.003)	(0.004)	(0.007)	(0.005)		
Party-Agency Partisan Align.				0.017	0.019	
				(0.015)	(0.016)	
Party-Agency Id. Dist.			0.010			-0.027
			(0.029)			(0.033)
Fixed-effects						
Party	✓					
Year	✓					
Agency	✓					
Party-Agency		✓	✓	✓	✓	✓
Year-Agency		✓	✓	✓	✓	✓
Fit statistics						
Observations	6,874	6,874	1,674	1,340	1,340	1,674
R^2	0.273	0.682	0.715	0.684	0.683	0.706
Within R^2	0.006	0.014	0.035	0.008	0.002	0.003

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

Table 3: Partisanhsip, Ideology, and Beliefs, US Data. OLS estimates. Units are party-agency-year observations.

negligible effect.[11] Models (5) and (6) replicate the estimation from Model (3), but this time I use the measure of party-agency ideological distance and partisan congruence. Absent bias in legislators' beliefs about bureaucracy, we would expect party-agency ideological or partisan differences to be better predictor of beliefs about agencies. However, in-group bias is not triggered by party-agency partisan or ideological congruence and none of the coefficient is distinguishable from zero at 95% level.[12]

The size of the effect of party-government alignment is 0.22 times the standard deviation

[11] In the Appendix (Table A.9) I show how the results are robust to adding additional covariates: agency budget, number of employees, and an indicator of politicisation (data from D. E. Lewis (2008)).

[12] It should be acknowledged that including both measures of party-agency and party-government alignment is likely to lead to post-treatment bias, hence the results of Model (4) should be interpreted with caution.

of beliefs in the sample, and as big as the average difference in beliefs between Democrats and Republicans (i.e., 0.03). The effect is of similar magnitude to the one estimated on UK data. The fact that, despite very different administrative traditions (i.e., politicised versus neutral civil service), the estimated effect in the US and UK are similar suggests that partisanship is able to affect beliefs in very different administrative systems. This is additional evidence in support of *Hypothesis 1*.

5.1 Testing the Mechanism: Scandals in the US Federal Bureaucracy

Beliefs are biased because co-partisan legislators follow partisan motives when forming their beliefs about bureaucracy. The results in Tables 2 and 3 show a negative effect of party-government partisan alignment on beliefs, but they offer limited support for the mechanism (e.g., motivated reasoning triggered by in-group bias). Furthermore, causal identification rests on the strong assumption of no omitted-variable bias. Because β is identified by comparing over-time changes in partisan beliefs when the government switches colour, bias would arise if changes in government co-occurred with changes in the characteristics of the legislators and how they interact with bureaucratic agencies. In this section I provide a stronger micro-test of the mechanism, while at the same time strengthening causal identification by looking at how legislators react to scandals in the US federal bureaucracy.

Empirical tests of partisan motivated reasoning generally relies on experimental manipulation in order to estimate partisan differences in reaction to identical information. To match this scenario as much as possible, I rely on exogenous shocks to the reputation of bureaucratic agencies resulted from scandals and compare the reaction of co-partisan and opposition-party legislators. I focus on major scandals involving three large federal bureaucracies in the United States: the response of the Federal Emergency Management Agency to Hurricane Katrina in

August 2005, the falsified-appointment case of the Department of Veterans Affairs in April 2014, and the Internal Revenue Service's undue scrutiny on conservative groups seeking tax-exempt status in May 2014. These scandals cover two presidencies of two different parties (the second G.W. Bush and Obama administrations) and are therefore not limited to one specific direction (either Democratic or Republican) of co-partisanship. Qualitative information on the scandals is reported in Section A.3 of the Appendix.

I estimate the effect of legislator-government co-partisanship on the sentiment of statements about the three agencies just before and after the date of the scandal. From the total sample of US floor speeches, I subset speeches given from 4 to one month before and after the scandal. I split the speeches into sentences and keep only sentences which mention the name of the agency involved in the scandal. I then apply a simple sentiment analysis to each sentence using the commonly-used Lexicoder Sentiment Dictionary provided within the *quanteda* library (Benoit et al. 2018). The dictionary contains lists of positive and negative words. I then count the number of words in each sentence belonging to the positive and negative lists and build a dichotomous measure of positivity equal to 1 if the sum of positive words is greater or equal than the sum of negative words, and 0 otherwise.

To identify the effect of partisan alignment with the government I leverage within-legislator pre- and post-scandal variation in the positivity of statements about the agency in a difference-in-differences design. In particular, I estimate the ATT of legislator-government partisan alignment with the following equation

$$\begin{aligned}\text{Positive Statement}_{i[l,a,t]} &= \eta_l + \phi_a + \alpha_d + \eta \text{Leg.-Govt. Alignment}_{l[t]} \\ &\quad + \gamma \text{Post-Scandal} + \beta \text{Leg.-Govt. Alignment} \times \text{Post-Scandal} + u_i\end{aligned} \tag{6}$$

β is the difference-in-differences estimator and identifies the effect of legislator-government partisan alignment on the probability of giving a positive statement about the agency affected by the scandal for co-partisans with the executive. η_l, ϕ_a, and δ_t are dummies to account for legislator, agency, and month-year differences. Despite the limited external validity, the appealing feature of this design is that it allows me to isolate how co-partisanship shapes legislators' *subjective* reaction to *objective* information (i.e., a clear national-level scandal).[13]

Table 4 shows the results across four different samples, namely all sentences given by legislators in floor speeches 4, 3, 2, 1 month(s) before and after the scandal. When we look at speeches given 2-1 month(s) before and after, and therefore increase the internal validity of the design, co-partisans are between 13-19 percentage points more likely to give positive statements about the agency involved in the scandal compared to the most likely counterfactaul. These findings suggest that, even when facing the same unambiguous information about bureaucracy, legislators react in a partisan way: more positive if aligned with the government, more negative if at the opposition.

This is a conservative test of the mechanism, for it exposes legislators to clearly negative information about the bureaucracy. In fact, motivated reasoning might be more pronounced when the valence of the information is ambiguous. The fact that I find large effects in co-partisans' reaction to scandal strengthens our confidence in the positive bias induced by co-partisanship with the government.

5.2 From Beliefs to Behaviour

What I have shown so far is that beliefs are biased, and that bias stems from partisan alignment between the party and the government. To show that biased beliefs is not simply

[13] In Table A.10 in the Appendix, I report falsification tests with placebo post-treatment indicators (i.e., placebo scandal dates 2 and 4 months before the true scandal date) in support of the parallel trend assumption in the pre-treatment period.

DV:	Pr(Positive Statement = 1)			
Months before/after Scandal	4 months	3 months	2 months	1 month
Model:	(1)	(2)	(3)	(4)
Variables				
Leg.-Govt. Partisan Alig.	0.052	0.050	-0.005	-0.057
	(0.031)	(0.034)	(0.048)	(0.092)
Post-Scandal	-0.000	-0.010	-0.071	-0.093
	(0.049)	(0.050)	(0.054)	(0.070)
Leg.-Govt. Partisan Alig. × Post-Scandal	0.026	0.027	0.133**	0.191*
	(0.032)	(0.032)	(0.044)	(0.081)
Fixed-effects				
Legislator	✓	✓	✓	✓
Month-Year	✓	✓	✓	✓
Agency	✓	✓	✓	✓
Fit statistics				
Observations	6,233	5,677	3,766	2,418
R^2	0.129	0.133	0.163	0.182
Within R^2	0.001	0.001	0.004	0.004

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

Table 4: ATT of legislator-government partisan alignment on the probability of giving a positive statement about bureaucracy estimated from four different samples of statements given 4, 3, 2, 1 month(s) before and after the date of the scandal.

a mask of strategic partisan rhetoric, I now show that partisan bias affects accountability practices too.

For this second test, the analysis is performed at legislator level. In support of *Hypothesis 3*, the two-way fixed effects results reported in Table 5 show that, compared to when they are at the opposition, when legislators argue about bureaucracy and are aligned with the government they are less likely to use analytical language grounded in statistical facts and evidence. The effect of partisan alignment is statistically significant across the different widows of words used and for both the UK and the US. In particular, focusing of segments of speeches 20 words before and after the name of the agency, the frequency of "fact-words" when there is partisan alignment decreases by -0.02 points for the US (average frequency among all speeches is 0.44) and by -0.03 points for the UK (average frequency among all

DV:	Facts (*tf-idf*)					
Country:	US			UK		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Leg.-Govt. Partisan Alig.	-0.021***	-0.014***	0.006	-0.031***	-0.034***	0.003
	(0.004)	(0.003)	(0.006)	(0.008)	(0.009)	(0.012)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	280,337	280,337	280,337	171,155	171,155	171,155
R^2	0.205	0.236	0.252	0.166	0.232	0.667
Within R^2	0.130	0.164	0.161	0.076	0.138	0.603

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

Table 5: Partisanhsip, Ideology, and Beliefs, US Data. OLS estimates. Dependent variable if tf-idf of facts-words in speeches. Controls include legislator's age and speech length (log number of words) and, for UK data only, legislator's seniority (i.e., log number of days in house) and whether the legislator holds government positions. In Table A.11 in the Appendix I show US results are robust when adding legislator-agency idelogical distance as a covariate.

speeches is 0.36). If we consider the average use of facts as a baseline, being co-partisan reduces the use of facts by approximately 5% for the US and 9% for the UK. Importantly, Models (3) and (6) show there is no effect of legislator-government partisan alignment on argumentative style when we focus on the entire speech, suggesting that partisan bias does not affect the speech as a whole, but rather the portion of speech about bureaucracy captured by the windows of words. In the Appendix I also show that the results for the US are robust to adding legislator-agency ideological distance as a covariate (see Table A.11) and that there is no statistically significant association between legislator-agency partisan congruence or ideological distance and argumentative style (see Table A.12).

Finally, the classical prediction for which ideologically divergent agencies are more likely to be subject to legislators' scrutiny is confirmed but only for legislators who are at the

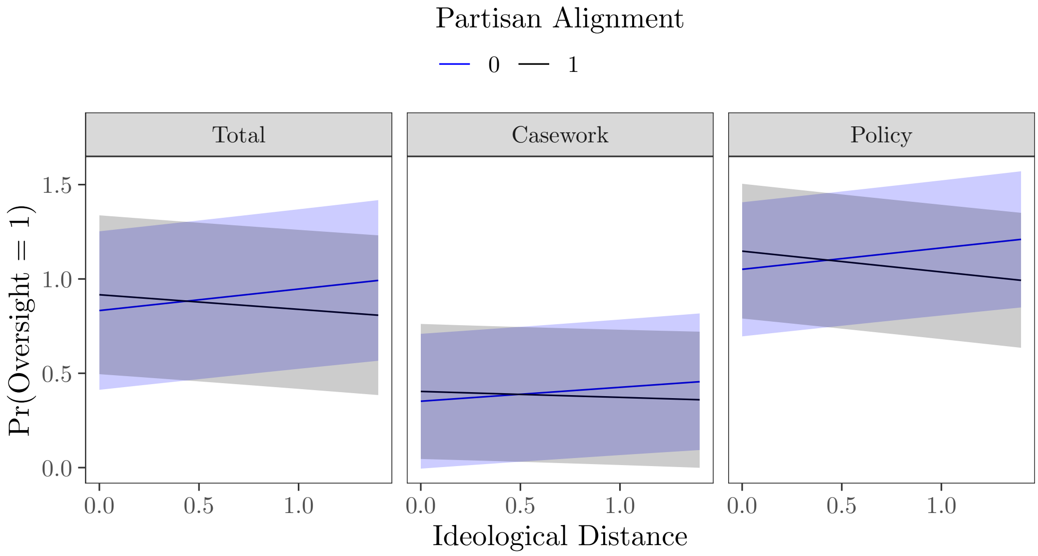

Figure 3: Predicted probability with 95% confidence intervals of legislator oversight based on ideological distance and partisan alignment with the executive (agency- and legislator-level covariates, agency, and legislator FE included in the estimation. SE clustered by legislator). Labels indicate the type of legislators' request to agencies.

opposition. Legislators are less likely to oversee ideologically distant agencies when they are aligned with the government.

While ideological distance has a negligible effect on the probability of oversight when excluding partisan alignment, its effect is much clearer once we allow it to vary based on co-partisanship with the government. Most importantly, there is a large slope change between co-partisan legislators and members of the opposition. The results are robust across each category of oversight but the moderating effect of co-partisanship is particularly strong for policy-relevant requests. A one-unit increase in ideological distance increases the probability of oversight by 11 percentage points for members of the opposition, whereas it decreases by 19 percentage points for co-partisans. Figure 3 plots the predicted probability against ideological distance for both co-partisans and opposition members. The slope change is large and highly significant.[14]

The effect of partisan alignment with the government on the relationship between ideo-

[14] Regression results as well as diagnostics tests for common support of the moderator (Hainmueller, Mummolo, and Xu 2018) are reported in Section A.7 of the Appendix (see Table A.14 and Table A.15). Figure A.1 shows density function of ideological distance across the two values of the moderator.

logical distance and oversight is very different based on the majority or opposition status of legislators. In fact, for co-partisans with the government the probability of oversight *decreases* with ideological distance. These findings are surprising and deserve further research.

6 Discussion

Holding bureaucracies to account is a central topic in political science. While scholars have shown how legislators design institutions to hold agencies to account, no attention has been paid to the possible biases that systematically affect the way legislators form their beliefs about bureaucracy, and how such biases affect the practice of accountability. In this paper, I have shown that legislators follow partisan motives when forming their beliefs about bureaucracy. Partisan alignment with the government, via in-group bias, triggers motivated reasoning and biases beliefs about bureaucracy: more positive when the party is in government, more negative when the party is at the opposition. Importantly, in support of the in-group bias mechanism, I find that bias stems only from partisan alignment with government, rather than as a result of ideological or partisan differences with the bureaucracy. Bias quickly moves from beliefs to argumentative style and oversight activities. Since the reasons which make beliefs more positive belong to partisan identity and co-partisanship with the government, when legislators talk about bureaucracies and they are of the same team as the government they use a more politicised style and resort to statistical facts and evidence less frequently. Finally, the commonly held proposition whereby ideological distance triggers oversight is much weaker when legislators are co-partisan with the government.

As far as the generalisability of the results is concerned, there are reasons to believe partisanship biases beliefs about bureaucracy in other countries with strong and rooted political parties, for I find similar effects of partisanship on beliefs in two countries with a very different administrative system. However, limits to the generalisability of the results arise from other features of the US and UK political systems. Both countries are advanced democracies with high levels of polarisation and majoritarian institutions. The relationship between elected legislators, partisanship, and bureaucracy in more proportional types of democracies or in countries with weaker administrative capacity might display different patterns from those observed in the cases studied here. Future research could look at such different contexts and at how bureaucracies – knowing legislators are motivated reasoners – make policies and interact with legislators.

This paper represents the first attempt at studying the effect of partisan motivated reasoning on belief formation and accountability behaviour. Doubts about the efficacy of majority-party legislators' ability to objectively hold agencies to account might be warranted.

References

Baekgaard, Martin, Julian Christensen, Casper Mondrup Dahlmann, Asbjørn Mathiasen, and Niels Bjørn Grund Petersen. 2017. "The Role of Evidence in Politics: Motivated Reasoning and Persuasion among Politicians." *British Journal of Political Science* 49 (3): 1117–40. https://doi.org/10.1017/S0007123417000084.

Bartels, M. 2002. "Beyond the Running Tally: Partisan Bias in Political Perceptions." *Political Behavior* 24 (2): 117–50. https://doi.org/10.1038/ajg.2011.482.

Bellodi, Luca. n.d. "A Dynamic Measure of Bureaucratic Reputation: New Data for New Theory." *American Journal of Political Science*, n.d. http://lucabellodi.com/Reputation_AJPS.pdf.

Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "Quanteda: An r Package for the Quantitative Analysis of Textual Data." *Journal of Open Source Software* 3 (30): 774. https://doi.org/10.21105/joss.00774.

Bertelli, Anthony M, Dyana P Mason, Jennifer M Connolly, and David A Gastwirth. 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: 513–44. https://doi.org/10.1093/jopart/mut040.

Blumenau, Jack. 2021. "House of Commons Parliamentary Debates, 1979-2019." Colchester, Essex: UK Data Service.

Bolton, Alexander, John M. de Figueiredo, and David E. Lewis. 2019. "Elections, Ideology, and Turnover in the U.S. Federal Government." https://doi.org/10.1017/CBO9781107415324.004.

Bonica, Adam. 2019. "Are donation-based measures of ideology valid predictors of individual-level policy preferences?" *Journal of Politics* 81 (1): 327–33. https://doi.org/10.1086/700722.

Busuioc, E. Madalina, and Martin Lodge. 2017. "Reputation and Accountability Relationships: Managing Accountability Expectations through Reputation." *Public Administration Review* 77 (1): 91–100. https://doi.org/10.1111/puar.12612.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy.* Princeton: Princeton University Press.

———. 2020. "On Categories and the Countability of Things Bureaucratic: Turning From Wilson (Back) to Interpretation." *Perspectives on Public Management and Governance* 3 (2): 83–93. https://doi.org/10.1093/ppmgov/gvz025.

Carrillo-de-Albornoz, Jorge, Javier Rodríguez Vidal, and Laura Plaza. 2018. "Feature engineering for sentiment analysis in e-health forums." *PLoS ONE* 13 (11): 1–25. https://doi.org/10.1371/journal.pone.0207996.

Chen, Jowei, and Tim Johnson. 2014. "Federal employee unionization and presidential control of the bureaucracy: Estimating and explaining ideological change in executive agencies." *Journal of Theoretical Politics* 27 (1): 151–74. https://doi.org/10.1177/0951

629813518126.

De Benedictis-Kessner, Justin, and Christopher Warshaw. 2020. “Accountability for the Local Economy at All Levels of Government in United States Elections.” *American Political Science Review*, 1–17. https://doi.org/10.1017/S0003055420000027.

Druckman, James N., Erik Peterson, and Rune Slothuus. 2013. “How elite partisan polarization affects public opinion formation.” *American Political Science Review* 107 (1): 57–79. https://doi.org/10.1017/S0003055412000500.

Ellig, Jerry, and Christopher J. Conover. 2014. “Presidential priorities, congressional control, and the quality of regulatory analysis: an application to healthcare and homeland security.” *Public Choice* 161 (3-4): 305–20. https://doi.org/10.1007/s11127-014-0201-3.

Epstein, David, and Sharyn O’Halloran. 1999. *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers.* Cambridge, UK: Cambridge University Press.

Erikson, Robert S. 1989. “Economic Conditions and The Presidential Vote.” *American Political Science Review* 83 (2): 567–73.

Gentzkow, Matthew, Jesse M. Shapiro, and Matt Taddy. 2018. “Congressional Record for the 43rd-114th Congresses: Parsed Speeches and Phrase Counts.” Palo Alto, CA: Stanford Libraries [distributor]. https://data.stanford.edu/congress_text.

Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts andMinds.* New Haven, CT: Yale University Press.

Grimmer, Justin, and Brandon M. Stewart. 2013. “Text as data: The promise and pitfalls of automatic content analysis methods for political texts.” *Political Analysis* 21 (3): 267–97. https://doi.org/10.1093/pan/mps028.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2018. “How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice.” *Political Analysis* 27 (2): 163–92. https://doi.org/10.2139/ssrn.2739221.

Hargrave, Lotte, and Jack Blumenau. 2020. “The Declining Importance of Gender Stereotypes for Politicians’ Style in the UK.”

Hood, Christopher. 1991. “A Public Management For All Seasons?” *Public Administration* 69: 3–19. https://doi.org/10.2307/25305228.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. 2019. “The origins and consequences of affective polarization in the United States.” *Annual Review of Political Science* 22: 129–46. https://doi.org/10.1146/annurev-polisci-051117-073034.

Iyengar, Shanto, and Sean J. Westwood. 2015. “Fear and Loathing across Party Lines: New Evidence on Group Polarization.” *American Journal of Political Science* 59 (3): 690–707. https://doi.org/10.1111/ajps.12152.

James, Oliver. 2011. “Performance measures and democracy: Information effects on citizens in field and laboratory experiments.” *Journal of Public Administration Research and Theory* 21 (3): 399–418. https://doi.org/10.1093/jopart/muq057.

James, Oliver, and Peter John. 2007. "Public management at the ballot box: Performance information and electoral support for incumbent english local governments." *Journal of Public Administration Research and Theory* 17 (4): 567–80. https://doi.org/10.1093/jopart/mul020.

Lewis, David E. 2008. *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*. Princeton University Press.

Lewis, Jeffrey, Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. 2020. "Voteview: Congressional Roll-Call Votes Database." voteview.com.

Little, Andrew, Keith Schnakenberg, and Ian Turner. 2021. "Motivated reasoning and democratic accountability." *American Political Science Review*, no. July. https://doi.org/10.31235/osf.io/esfy6.

Lowande, Kenneth. 2018. "Who Polices the Administrative State?" *American Political Science Review* 112 (4): 874–90. https://doi.org/10.1017/S0003055418000497.

Lupia, Arthur, and Mathew D. McCubbins. 1994. "Designing bureaucratic accountability." *Law and Contemporary Problems* 57 (1): 91–126. https://doi.org/10.2307/1191988.

Maor, Moshe, and Raanan Sulitzeanu-Kenan. 2016. "Responsive Change: Agency Output Response to Reputational Threats." *Journal of Public Administration Research and Theory* 26 (1): 31–44. https://doi.org/10.1093/jopart/muv001.

Mason, Lilliana. 2015. ""I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization." *American Journal of Political Science* 59 (1): 128–45. https://doi.org/10.1111/ajps.12089.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: The MIT Press.

McCubbins, M, Roger G Noll, and Barry R Weingast. 1987. "Administrative Procedures as Instruments of Political Control." *Journal of Law, Economics, & Organization* 3 (2): 243–77.

McCubbins, M, and T Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28 (1): 165–79.

Miller, Gary J. 2005. "the Political Evolution of Principal-Agent Models." *Annual Review of Political Science* 8 (1): 203–25. https://doi.org/10.1146/annurev.polisci.8.082103.104840.

Moore, Emily H. 2018. "Polarization, Excepted Appointments, and the Administrative Presidency." *Presidential Studies Quarterly* 48 (1): 72–92. https://doi.org/10.1111/psq.12417.

Nielsen, Poul A., and Martin Baekgaard. 2015. "Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment." *Journal of Public Administration Research and Theory* 25 (2): 545–69. https://doi.org/10.1093/jopart/mut051.

Patty, John W., and Ian R. Turner. 2020. "Ex Post Review and Expert Policymaking: When Does Oversight Reduce Accountability?" *Journal of Politics*. https://doi.org/https://doi.org/10.1086/708913.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The Development and Psychometric Properties of LIWC2015.* Austin: University of Texas at Austin. https://doi.org/10.1068/d010163.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "GloVe: Global Vectors for Word Representation." *Empirical Methods in Natural Language Processing (EMNLP)*, 1532–43.

Potter, Rachel A. 2019. *Bending the Rules. Pocedural Politicking in the Bureaucracy.* Chicago: University of Chicago Press.

Rice, Douglas R., and Christopher Zorn. 2019. "Corpus-based dictionaries for sentiment analysis of specialized vocabularies." *Political Science Research and Methods*, 1–17. https://doi.org/10.1017/psrm.2019.10.

Roberts, Patrick S. 2006. "FEMA and the prospects for reputation-based autonomy." *Studies in American Political Development* 20 (1): 57–87. https://doi.org/10.1017/S0898588X06000010.

Schillemans, Thomas. 2011. "Does horizontal accountability work? Evaluating potential remedies for the accountability deficit of agencies." *Administration and Society* 43 (4): 387–416. https://doi.org/10.1177/0095399711412931.

Schillemans, Thomas, and E. Madalina Busuioc. 2015. "Predicting public sector accountability: From agency drift to forum drift." *Journal of Public Administration Research and Theory* 25 (1): 191–215. https://doi.org/10.1093/jopart/muu024.

Selin, Jennifer L. 2015. "What Makes an Agency Independent?" *American Journal of Political Science* 59 (4): 971–87. https://doi.org/10.1111/ajps.12161.

Spirling, Arthur, and Pedro L Rodriguez. 2019. "Word Embeddings What works, what doesn't, and how to tell the difference for applied research."

Taber, Charles S, and Milton Lodge. 2006. "Motivated Skepticism in the Evaluationof Political Beliefs." *American Journal of Political Science* 50 (3): 755–69.

Tajfel, Henri. 1987. "The achievement of inter-group differentiation." In *Differentiation Between Social Groups*, edited by Henri Tajfel, 77–100. London: Academic Press.

TIGTA. 2013. "Inappropriate Criteria Were Used to Identify Tax-Exempt Applications for Review." https://www.treasury.gov/tigta/auditreports/2013reports/201310053fr.pdf.

Timeline. 2017. "Snapshots from FEMA's disastrous failure in the wake of Hurricane Andrew." timeline.com/fema-hurricane-andrew-snapshots-7a764f017614.

Turner, Ian R. 2017. "Working smart and hard? Agency effort, judicial review, and policy precision." *Journal of Theoretical Politics* 29 (1): 69–96. https://doi.org/10.1177/0951629816630430.

Vox. 2015. "The VA scandal of 2014, explained." vox.com/2014/9/26/18080592/va-scandal-explained.

Welbers, Kasper, Wouter Van Atteveldt, and Kenneth Benoit. 2017. "Text Analysis in R." *Communication Methods and Measures* 11 (4): 245–65. https://doi.org/10.1080/193124

58.2017.1387238.

Appendix

Contents

A.1 Sample of Agencies A2

A.1.1 Sample of UK Agencies A2

A.1.2 Sample of US Agencies A7

A.2 Measuring Partisan Beliefs About Bureaucracy A14

A.3 Scandals: Qualitative Description A16

A.4 Text Pre-Processing A17

A.5 Dictionary Approach: Validation A18

A.6 Robustness Checks A20

A.7 Test 4: Diagnostics A29

A.1 Sample of Agencies

In the following tables, I report all the agencies for which I produce beliefs estimates, the number of observations for each agency (i.e., the number of years for which I produce an estimate), the average number of mentions per year, and the average beliefs among both parties.

A.1.1 Sample of UK Agencies

Table A.1: Sample of UK agencies and descriptive statistics.

			Conservative		Labour	
Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Advisory Committee on Business Appointments	Non-dept. Agency	2	28	0.31	5	0.51
Advisory Council on the Misuse of Drugs	Non-dept. Agency	7	16	0.44	14	0.43
Advisory, Conciliation and Arbitration Service	Non-dept. Agency	18	45	0.39	48	0.46
Animal and Plant Health Agency	Non-dept. Agency	1	20	0.59	5	0.34
Armed Forces' Pay Review Body	Non-dept. Agency	5	10	0.40	11	0.46
Arts Council England	Non-dept. Agency	4	11	0.41	13	0.52
Arts Council of Wales	Non-dept. Agency	2	12	0.31	16	0.45
Atomic Energy Authority	Non-dept. Agency	10	16	0.44	12	0.46
Bank of England	Non-dept. Agency	40	137	0.47	129	0.50
Boundary Commission for England	Non-dept. Agency	1	10	0.24	16	0.45
Boundary Commission for Wales	Non-dept. Agency	1	7	0.26	9	0.55
British Business Bank	Non-dept. Agency	1	8	0.45	6	0.59
British Council	Non-dept. Agency	30	36	0.54	33	0.56
Broads Authority	Non-dept. Agency	4	18	0.49	30	0.36
Cabinet Office	Government Dept.	36	70	0.51	68	0.52
Care Quality Commission	Non-dept. Agency	16	118	0.44	89	0.43
Central Arbitration Committee	Non-dept. Agency	2	10	0.32	6	0.47
Children and Family Court Advisory and Support Service	Non-dept. Agency	1	5	0.35	12	0.39
Civil Justice Council	Non-dept. Agency	1	6	0.48	7	0.27
Coal Authority	Non-dept. Agency	5	32	0.35	61	0.39
College of Policing	Non-dept. Agency	4	43	0.41	17	0.57
Commissioner for Public Appointments	Non-dept. Agency	3	11	0.35	10	0.42
Committee on Climate Change	Non-dept. Agency	12	22	0.34	44	0.49
Committee on Standards in Public Life	Non-dept. Agency	14	18	0.41	21	0.45

Table A.1: Sample of UK agencies and descriptive statistics. *(continued)*

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Beliefs	Avg. Mentions	Avg. Beliefs
Committee on Toxicity of Chemicals in Food, Consumer Products and the Environment	Non-dept. Agency	1	6	0.46	6	0.15
Companies House	Non-dept. Agency	9	15	0.35	17	0.54
Competition Appeal Tribunal	Non-dept. Agency	1	12	0.39	11	0.51
Construction Industry Training Board	Non-dept. Agency	10	22	0.42	20	0.49
Consumer Council for Water	Non-dept. Agency	1	5	0.49	6	0.60
Courts and Tribunals Service	Non-dept. Agency	6	21	0.49	10	0.47
Criminal Cases Review Commission	Non-dept. Agency	7	34	0.43	15	0.48
Criminal Injuries Compensation Authority	Non-dept. Agency	1	9	0.55	8	0.49
Department for Business, Energy and Industrial Strategy	Government Dept.	4	25	0.46	19	0.35
Department for Digital, Culture, Media and Sport	Government Dept.	18	20	0.42	19	0.49
Department for Education	Government Dept.	38	95	0.52	90	0.53
Department for Environment Food and Rural Affairs	Government Dept.	19	72	0.41	63	0.48
Department for Exiting the European Union	Government Dept.	4	52	0.48	25	0.32
Department for International Development	Government Dept.	22	74	0.57	72	0.56
Department for International Trade	Government Dept.	4	96	0.68	18	0.43
Department for Transport	Government Dept.	19	151	0.53	144	0.63
Department for Work and Pensions	Government Dept.	19	108	0.48	152	0.49
Department of Health and Social Care	Government Dept.	40	175	0.54	173	0.54
Department of Social Security	Government Dept.	16	62	0.44	73	0.37
Disclosure and Barring Service	Non-dept. Agency	2	6	0.39	10	0.40
Driver and Vehicle Licensing Agency	Non-dept. Agency	21	26	0.44	28	0.49
Driver and Vehicle Standards Agency	Non-dept. Agency	2	14	0.49	7	0.50
Economic and Social Research Council	Non-dept. Agency	1	8	0.47	5	0.45
Education and Skills Funding Agency	Non-dept. Agency	1	15	0.39	8	0.12
Environment Agency	Non-dept. Agency	25	135	0.51	117	0.53
Equality and Human Rights Commission	Non-dept. Agency	11	16	0.43	30	0.43
Export Guarantees Advisory Council	Non-dept. Agency	1	7	0.26	10	0.47
Financial Conduct Authority	Non-dept. Agency	9	203	0.59	109	0.58
Financial Reporting Council	Non-dept. Agency	5	12	0.44	12	0.57
Financial Services Authority	Non-dept. Agency	21	135	0.45	164	0.58
Fire Service College	Non-dept. Agency	2	14	0.39	8	0.46

Table A.1: Sample of UK agencies and descriptive statistics. *(continued)*

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Foreign and Commonwealth Office	Government Dept.	40	226	0.57	202	0.52
Gambling Commission	Non-dept. Agency	4	55	0.54	22	0.49
Gangmasters and Labour Abuse Authority	Non-dept. Agency	1	11	0.52	5	0.62
Government Equalities Office	Non-dept. Agency	3	16	0.44	10	0.56
Groceries Code Adjudicator	Non-dept. Agency	5	22	0.53	9	0.59
Health and Safety Executive	Non-dept. Agency	40	60	0.45	84	0.54
Health Education England	Non-dept. Agency	6	32	0.53	15	0.47
Highways England	Non-dept. Agency	5	147	0.49	30	0.49
Home Office	Government Dept.	40	390	0.68	434	0.66
Horserace Betting Levy Board	Non-dept. Agency	1	5	0.41	11	0.53
Human Fertilisation and Embryology Authority	Non-dept. Agency	11	17	0.37	24	0.45
Human Tissue Authority	Non-dept. Agency	1	14	0.27	38	0.43
Independent Anti-slavery Commissioner	Non-dept. Agency	1	6	0.47	7	0.58
Independent Commission for Aid Impact	Non-dept. Agency	4	20	0.44	10	0.44
Independent Office for Police Conduct	Non-dept. Agency	2	24	0.55	12	0.85
Independent Parliamentary Standards Authority	Non-dept. Agency	11	121	0.48	86	0.51
Independent Reconfiguration Panel	Non-dept. Agency	1	15	0.45	7	0.32
Industrial Injuries Advisory Council	Non-dept. Agency	1	7	0.37	10	0.60
Information Commissioner's Office	Non-dept. Agency	4	24	0.39	13	0.38
Insolvency Service	Non-dept. Agency	3	11	0.42	10	0.38
Institute for Apprenticeships and Technical Education	Non-dept. Agency	2	5	0.24	6	0.42
Intellectual Property Office	Non-dept. Agency	2	9	0.34	6	0.62
Judicial Appointments Commission	Non-dept. Agency	1	7	0.35	24	0.35
Law Commission	Non-dept. Agency	40	57	0.45	45	0.51
Legal Aid Agency	Non-dept. Agency	3	18	0.49	13	0.43
Legal Services Board	Non-dept. Agency	1	6	0.32	5	0.21
Low Pay Commission	Non-dept. Agency	9	30	0.39	39	0.51
Marine Accident Investigation Branch	Non-dept. Agency	4	16	0.41	10	0.41
Marine Management Organisation	Non-dept. Agency	2	14	0.36	10	0.38
Maritime and Coastguard Agency	Non-dept. Agency	5	11	0.47	11	0.52
Medical Research Council	Non-dept. Agency	24	16	0.44	13	0.51
Medicines and Healthcare products Regulatory Agency	Non-dept. Agency	11	18	0.46	22	0.41
Migration Advisory Committee	Non-dept. Agency	5	33	0.34	8	0.39

Table A.1: Sample of UK agencies and descriptive statistics. *(continued)*

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Ministry of Defence	Government Dept.	40	236	0.66	201	0.58
Ministry of Housing, Communities and Local Government	Government Dept.	2	150	0.59	47	0.37
Ministry of Justice	Government Dept.	13	163	0.61	128	0.56
National Data Guardian	Non-dept. Agency	1	42	0.34	6	0.54
National Infrastructure Commission	Non-dept. Agency	2	28	0.46	14	0.65
National Institute for Health and Care Excellence	Non-dept. Agency	7	35	0.48	20	0.40
Natural Environment Research Council	Non-dept. Agency	2	10	0.31	9	0.56
Network Rail	Non-dept. Agency	28	204	0.52	224	0.56
Northern Ireland Office	Government Dept.	37	26	0.44	20	0.46
Nuclear Decommissioning Authority	Non-dept. Agency	2	6	0.37	9	0.35
Office for Budget Responsibility	Non-dept. Agency	10	74	0.43	96	0.39
Office for National Statistics	Non-dept. Agency	23	50	0.39	52	0.42
Office for Students	Non-dept. Agency	4	31	0.44	20	0.53
Office of Communications	Non-dept. Agency	20	111	0.49	112	0.55
Office of Gas and Electricity Markets	Non-dept. Agency	18	50	0.48	77	0.45
Office of Gas Supply	Non-dept. Agency	5	13	0.41	9	0.39
Office of Rail and Road	Non-dept. Agency	15	21	0.42	28	0.43
Office of Tax Simplification	Non-dept. Agency	4	16	0.47	7	0.41
Office of the Children's Commissioner	Non-dept. Agency	2	7	0.49	5	0.25
Office of the Public Guardian	Non-dept. Agency	1	14	0.34	22	0.21
Oil and Gas Authority	Non-dept. Agency	1	50	0.52	11	0.46
Parole Board	Non-dept. Agency	14	28	0.41	20	0.39
Payment Systems Regulator	Non-dept. Agency	2	8	0.41	5	0.52
Pensions Regulator	Non-dept. Agency	5	26	0.41	8	0.56
Planning Inspectorate	Non-dept. Agency	9	42	0.34	14	0.50
Police Service of Northern Ireland	Non-dept. Agency	15	19	0.49	20	0.54
Prison Service	Non-dept. Agency	25	50	0.47	67	0.51
Privy Council Office	Non-dept. Agency	4	24	0.44	13	0.38
Prudential Regulatory Authority	Non-dept. Agency	2	40	0.44	18	0.55
Public Health England	Non-dept. Agency	8	100	0.60	26	0.53
Pubs Code Adjudicator	Non-dept. Agency	2	12	0.43	8	0.38
Rail Safety and Standards Board	Non-dept. Agency	2	8	0.38	14	0.51
Rural Payments Agency	Non-dept. Agency	11	23	0.37	16	0.32

Table A.1: Sample of UK agencies and descriptive statistics. *(continued)*

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Science and Technology Facilities Council	Non-dept. Agency	2	6	0.30	15	0.43
Sea Fish Industry Authority	Non-dept. Agency	4	20	0.42	13	0.58
Secret Intelligence Service	Non-dept. Agency	16	13	0.37	15	0.52
Security Industry Authority	Non-dept. Agency	5	28	0.46	70	0.48
Senior Salaries Review Body	Non-dept. Agency	5	9	0.41	10	0.33
Sentencing Council	Non-dept. Agency	6	31	0.42	14	0.45
Single Financial Guidance Body	Non-dept. Agency	2	10	0.31	11	0.64
Social Mobility Commission	Non-dept. Agency	3	16	0.40	28	0.43
Social Security Advisory Committee	Non-dept. Agency	15	18	0.34	15	0.50
Space Agency	Non-dept. Agency	2	10	0.45	6	0.46
Sports Council for Wales	Non-dept. Agency	1	10	0.10	14	0.53
Sports Grounds Safety Authority	Non-dept. Agency	1	11	0.56	5	0.45
Treasury	Government Dept.	40	818	0.64	759	0.60
UK Export Finance	Government Dept.	1	27	0.54	15	0.48
Valuation Office Agency	Non-dept. Agency	6	18	0.39	14	0.45
Valuation Tribunal Service	Non-dept. Agency	1	5	0.54	9	0.66
Victims' Commissioner	Non-dept. Agency	3	13	0.35	12	0.55
Youth Justice Board	Non-dept. Agency	9	15	0.37	25	0.49

A.1.2 Sample of US Agencies

Table A.2: Sample of US agencies and descriptive statistics.

			Republican		Democratic	
Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Administrative Conference of the United States	Independent Agency	2	8	0.18	10	0.43
Agency for Healthcare Research and Quality	Ex. Sub-agency	2	6	0.39	5	0.31
Agency for International Development	Independent Agency	36	45	0.48	62	0.50
Agency for Toxic Substances and Disease Registry	Ex. Sub-agency	3	19	0.51	9	0.40
Agricultural Research Service	Ex. Sub-agency	19	14	0.44	17	0.47
Air Force	Ex. Sub-agency	36	526	0.69	480	0.66
Alcohol and Tobacco Tax and Trade Bureau	Ex. Sub-agency	1	9	0.09	8	0.65
American Battle Monuments Commission	Ex. Sub-agency	4	7	0.25	13	0.49
AMTRAK	Independent Agency	1	5	0.33	10	0.32
Animal and Plant Health Inspection Service	Ex. Sub-agency	19	16	0.41	16	0.54
Appalachian Regional Commission	Ex. Sub-agency	20	18	0.43	26	0.52
Benefits Review Board	Ex. Sub-agency	1	6	0.36	16	0.40
Board of Veterans Appeals	Ex. Sub-agency	4	28	0.37	38	0.50
Bonneville Power Administration	Ex. Sub-agency	11	22	0.42	14	0.42
Broadcasting Board of Governors	Independent Agency	5	7	0.43	11	0.51
Bureau of Alcohol, Tobacco, Firearms, and Explosives	Ex. Sub-agency	26	27	0.40	43	0.51
Bureau of Economic Analysis	Ex. Sub-agency	1	10	0.48	13	0.53
Bureau of Indian Affairs	Ex. Sub-agency	33	30	0.44	27	0.45
Bureau of Indian Education	Ex. Sub-agency	1	5	0.51	6	0.51
Bureau of Labor Statistics	Ex. Sub-agency	35	17	0.39	29	0.45
Bureau of Land Management	Ex. Sub-agency	36	51	0.43	50	0.48
Bureau of Ocean Energy Management	Ex. Sub-agency	2	8	0.48	7	0.43
Bureau of Prisons	Ex. Sub-agency	27	20	0.41	19	0.47
Bureau of Reclamation	Ex. Sub-agency	36	51	0.50	50	0.56
Bureau of the Census	Ex. Sub-agency	36	41	0.40	67	0.42
Centers for Disease Control and Prevention	Ex. Sub-agency	22	19	0.46	38	0.50
Centers for Medicare and Medicaid Services	Ex. Sub-agency	16	113	0.44	75	0.48
Central Intelligence Agency	Independent Agency	36	201	0.52	321	0.56
Citizen and Immigration Services	Ex. Sub-agency	7	21	0.49	14	0.56

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Civil Rights Division	Ex. Sub-agency	12	16	0.48	20	0.53
Commission on Civil Rights	Independent Agency	9	16	0.37	18	0.41
Commodities Futures Trading Commission	Independent Agency	21	74	0.52	89	0.54
Commodity Credit Corporation	Gvt.-owned Corp.	22	41	0.44	49	0.41
Consumer Financial Protection Bureau	Independent Agency	7	123	0.46	170	0.55
Consumer Product Safety Commission	Independent Agency	27	33	0.47	67	0.49
Corporation for National and Community Service	Independent Agency	7	11	0.49	12	0.44
Corporation for Public Broadcasting	Non-for-profit Organiz.	16	61	0.48	66	0.47
Council of Economic Advisers	Ex. Office of Presid.	25	13	0.41	21	0.43
Council on Environmental Quality	Ex. Sub-agency	10	9	0.35	11	0.47
Court Services and Offender Supervision Agency	Ex. Sub-agency	1	11	0.58	16	0.37
Customs and Border Protection	Ex. Sub-agency	14	30	0.53	32	0.53
Defense Advanced Research Projects Agency	Ex. Sub-agency	13	11	0.43	23	0.55
Defense Contract Audit Agency	Ex. Sub-agency	3	16	0.49	36	0.44
Defense Finance and Accounting Service	Ex. Sub-agency	1	48	0.44	7	0.40
Defense Intelligence Agency	Ex. Sub-agency	28	24	0.44	19	0.46
Defense Logistics Agency	Ex. Sub-agency	8	9	0.38	14	0.43
Department of Agriculture	Ex. Department	36	155	0.55	189	0.55
Department of Commerce	Ex. Department	36	139	0.55	160	0.56
Department of Defense	Ex. Department	36	652	0.66	819	0.66
Department of Education	Ex. Department	36	147	0.49	139	0.50
Department of Energy	Ex. Department	36	260	0.58	294	0.58
Department of Health and Human Services	Ex. Department	36	178	0.52	174	0.53
Department of Homeland Security	Ex. Department	16	637	0.79	887	0.73
Department of Housing and Urban Development	Ex. Department	36	33	0.48	40	0.49
Department of Justice	Ex. Department	36	486	0.60	546	0.60
Department of Labor	Ex. Department	36	132	0.51	158	0.52
Department of State	Ex. Department	36	72	0.52	77	0.50
Department of the Army	Ex. Department	24	14	0.39	15	0.47
Department of the Interior	Ex. Department	36	133	0.53	150	0.52
Department of the Navy	Ex. Department	36	487	0.68	473	0.67
Department of the Treasury	Ex. Department	36	712	0.45	740	0.48
Department of Transportation	Ex. Department	36	130	0.57	144	0.53

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Beliefs	Avg. Mentions	Avg. Beliefs
Department of Veterans Affairs	Ex. Department	36	225	0.64	298	0.68
Domestic Nuclear Detection Office	Ex. Sub-agency	3	13	0.47	11	0.48
Drug Enforcement Administration	Ex. Sub-agency	36	83	0.50	72	0.57
Economic Development Administration	Ex. Sub-agency	28	71	0.43	130	0.56
Economic Research Service	Ex. Sub-agency	1	10	0.40	10	0.55
Election Assistance Commission	Ex. Sub-agency	5	28	0.49	32	0.48
Environmental Protection Agency	Independent Agency	36	720	0.57	753	0.56
Equal Employment Opportunity Commission	Independent Agency	32	37	0.41	40	0.48
Export-Import Bank of the United States	Independent Agency	1	12	0.54	5	0.53
Farm Credit Administration	Independent Agency	9	25	0.45	28	0.43
Farm Service Agency	Ex. Sub-agency	10	12	0.42	15	0.49
Federal Agricultural Mortgage Corporation	Gvt.-owned Corp.	4	39	0.58	30	0.47
Federal Aviation Administration	Ex. Sub-agency	35	34	0.48	50	0.51
Federal Bureau of Investigation	Ex. Sub-agency	36	327	0.56	350	0.59
Federal Communications Commission	Independent Agency	36	174	0.50	194	0.51
Federal Deposit Insurance Corporation	Independent Agency	36	70	0.44	90	0.45
Federal Election Commission	Independent Agency	36	61	0.41	55	0.50
Federal Emergency Management Agency	Ex. Sub-agency	36	128	0.49	227	0.49
Federal Energy Regulatory Commission	Ex. Sub-agency	36	81	0.47	126	0.48
Federal Highway Administration	Ex. Sub-agency	34	24	0.43	26	0.43
Federal Housing Administration	Ex. Sub-agency	36	90	0.42	102	0.48
Federal Housing Finance Agency	Independent Agency	4	15	0.43	16	0.46
Federal Labor Relations Authority	Independent Agency	3	11	0.47	7	0.39
Federal Law Enforcement Training Center	Ex. Sub-agency	9	17	0.48	10	0.40
Federal Maritime Commission	Independent Agency	17	23	0.46	35	0.49
Federal Mediation and Conciliation Service	Independent Agency	1	8	0.09	11	0.50
Federal Motor Carrier Safety Administration	Ex. Sub-agency	7	9	0.35	12	0.47
Federal Prison Industries	Gvt.-owned Corp.	10	36	0.42	30	0.45
Federal Railroad Administration	Ex. Sub-agency	22	22	0.48	32	0.45
Federal Reserve	Independent Agency	36	192	0.40	306	0.45
Federal Student Aid	Ex. Sub-agency	1	8	0.44	6	0.41
Federal Trade Commission	Independent Agency	36	103	0.48	129	0.53
Federal Transit Administration	Ex. Sub-agency	7	8	0.38	12	0.53
Financial Crimes Enforcement Network	Ex. Sub-agency	6	9	0.49	9	0.46

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Beliefs	Avg. Mentions	Avg. Beliefs
Financial Management Service	Ex. Sub-agency	6	19	0.34	13	0.38
Financial Stability Oversight Council	Ex. Sub-agency	5	25	0.56	22	0.48
Fish and Wildlife Service	Ex. Sub-agency	36	73	0.52	65	0.52
Food and Drug Administration	Ex. Sub-agency	36	318	0.57	370	0.61
Food and Nutrition Service	Ex. Sub-agency	4	7	0.36	12	0.40
Food Safety and Inspection Service	Ex. Sub-agency	2	12	0.54	8	0.42
Foreign Claims Settlement Commission	Ex. Sub-agency	1	8	0.38	6	0.36
Forest Service	Ex. Sub-agency	36	235	0.53	216	0.57
General Services Administration	Independent Agency	36	76	0.50	91	0.53
Geological Survey	Ex. Sub-agency	36	25	0.44	29	0.50
Government National Mortgage Association	Gvt.-owned Corp.	5	21	0.42	22	0.42
Grain Inspection, Packers, and Stockyards Administration	Ex. Sub-agency	2	36	0.41	17	0.52
Health Resources and Services Administration	Ex. Sub-agency	14	10	0.47	12	0.49
Housing Finance Agency	Independent Agency	1	7	0.62	5	0.45
Immigration and Customs Enforcement	Ex. Sub-agency	17	106	0.46	57	0.53
Independent Payment Advisory Board	Independent Agency	3	166	0.41	65	0.47
Indian Health Service	Ex. Sub-agency	30	29	0.44	32	0.44
Institute of Peace	Independent Agency	7	8	0.46	16	0.49
Inter-American Foundation	Independent Agency	2	10	0.50	12	0.52
Internal Revenue Service	Ex. Sub-agency	36	602	0.44	431	0.48
International Boundary and Water Commission	Ex. Sub-agency	2	6	0.38	8	0.61
International Trade Administration	Ex. Sub-agency	8	10	0.49	10	0.39
International Trade Commission	Independent Agency	35	62	0.43	58	0.45
Maritime Administration	Ex. Sub-agency	21	18	0.42	16	0.45
Marshals Service	Ex. Sub-agency	22	15	0.48	17	0.51
Merit Systems Protection Board	Independent Agency	19	13	0.35	19	0.41
Millennium Challenge Corporation	Independent Agency	4	16	0.39	8	0.46
Minority Business Development Agency	Ex. Sub-agency	3	15	0.48	25	0.44
Missile Defense Agency	Ex. Sub-agency	7	19	0.43	17	0.46
National Aeronautics and Space Administration	Independent Agency	36	220	0.62	315	0.67
National Archives and Records Administration	Independent Agency	1	12	0.47	8	0.21
National Capital Planning Commission	Independent Agency	4	15	0.47	16	0.44
National Consumer Cooperative Bank	NA	1	17	0.46	20	0.41
National Credit Union Administration	Independent Agency	10	15	0.45	18	0.40

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Beliefs	Avg. Mentions	Avg. Beliefs
National Geospatial-Intelligence Agency	Ex. Sub-agency	4	26	0.45	14	0.49
National Highway Traffic Safety Administration	Ex. Sub-agency	31	29	0.45	31	0.49
National Institute of Standards and Technology	Ex. Sub-agency	28	41	0.51	49	0.57
National Institutes of Health	Ex. Sub-agency	36	75	0.55	113	0.56
National Labor Relations Board	Independent Agency	28	68	0.44	69	0.45
National Mediation Board	Independent Agency	4	32	0.33	40	0.38
National Nuclear Security Administration	Ex. Sub-agency	13	23	0.46	25	0.46
National Oceanic and Atmospheric Administration	Ex. Sub-agency	34	17	0.48	27	0.51
National Park Service	Ex. Sub-agency	36	111	0.56	133	0.61
National Reconnaissance Office	Ex. Sub-agency	7	37	0.54	35	0.47
National Science Foundation	Independent Agency	36	107	0.56	130	0.60
National Security Agency	Ex. Sub-agency	35	36	0.41	48	0.51
National Technical Information Service	Ex. Sub-agency	2	16	0.45	32	0.52
National Telecommunications and Information Administration	Ex. Sub-agency	16	11	0.43	24	0.53
National Transportation Safety Board	Independent Agency	32	28	0.46	40	0.48
Natural Resources Conservation Service	Ex. Sub-agency	10	11	0.42	13	0.41
Nuclear Regulatory Commission	Independent Agency	35	79	0.48	100	0.51
Occupational Safety and Health Administration	Ex. Sub-agency	34	123	0.49	107	0.51
Occupational Safety and Health Review Commission	Independent Agency	2	29	0.52	13	0.58
Office of Economic Adjustment	Ex. Sub-agency	2	9	0.42	14	0.40
Office of Energy Efficiency and Renewable Energy	Ex. Sub-agency	3	20	0.45	10	0.54
Office of Federal Procurement Policy	Ex. Office of Presid.	6	16	0.34	16	0.46
Office of Foreign Assets Control	Ex. Sub-agency	3	15	0.36	38	0.48
Office of Government Ethics	Independent Agency	11	27	0.47	33	0.45
Office of Justice Programs	Ex. Sub-agency	4	8	0.37	14	0.46
Office of Labor-Management Standards	Ex. Sub-agency	1	69	0.66	13	0.32
Office of Management and Budget	Ex. Office of Presid.	36	254	0.42	304	0.45
Office of National Drug Control Policy	Ex. Office of Presid.	22	30	0.50	20	0.49
Office of Personnel Management	Independent Agency	36	52	0.44	55	0.51
Office of Science and Technology	Ex. Office of Presid.	8	8	0.30	13	0.27
Office of Special Counsel	Independent Agency	11	14	0.41	17	0.47

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Beliefs	Avg. Mentions	Avg. Beliefs
Office of the Comptroller of the Currency	Ex. Sub-agency	2	10	0.47	8	0.29
Office of the Director of National Intelligence	Independent Agency	3	7	0.46	9	0.45
Office of Thrift Supervision	Ex. Sub-agency	11	26	0.41	44	0.44
Parole Commission	Ex. Sub-agency	5	10	0.50	12	0.41
Patent and Trademark Office	Ex. Sub-agency	27	39	0.46	43	0.47
Peace Corps	Independent Agency	35	71	0.53	122	0.57
Pension Benefit Guaranty Corporation	Independent Agency	19	39	0.36	40	0.42
Pentagon	Ex. Sub-agency	36	187	0.44	369	0.45
Pipeline and Hazardous Materials Safety Administration	Ex. Sub-agency	4	16	0.52	27	0.45
Postal Regulatory Commission	Independent Agency	1	16	0.39	9	0.25
Privacy and Civil Liberties Oversight Board	Independent Agency	3	9	0.42	9	0.57
Public Buildings Service	Independent Agency	1	8	0.35	7	0.52
Public Health Service	Ex. Sub-agency	33	37	0.49	42	0.48
Railroad Retirement Board	Independent Agency	4	17	0.32	12	0.41
Risk Management Agency	Ex. Sub-agency	2	16	0.44	20	0.62
Rural Housing Service	Ex. Sub-agency	1	7	0.50	6	0.34
Rural Utilities Service	Ex. Sub-agency	2	10	0.42	12	0.47
Securities and Exchange Commission	Independent Agency	36	155	0.59	240	0.62
Securities Investor Protection Corporation	Non-for-profit Organiz.	3	28	0.23	36	0.55
Selective Service System	Independent Agency	7	20	0.34	23	0.47
Small Business Administration	Independent Agency	36	181	0.57	233	0.58
Social Security Administration	Independent Agency	36	50	0.46	59	0.48
State Justice Institute	Non-for-profit Organiz.	3	14	0.51	16	0.36
Substance Abuse and Mental Health Services Administration	Ex. Sub-agency	12	14	0.49	19	0.46
Surface Transportation Board	Independent Agency	8	10	0.43	15	0.48
Tennessee Valley Authority	Independent Agency	22	18	0.44	21	0.50
Trade and Development Agency	Independent Agency	3	20	0.49	13	0.52
Transportation Security Administration	Ex. Sub-agency	16	125	0.57	134	0.59
Tricare Management Activity	Ex. Sub-agency	19	39	0.38	56	0.50
US Postal Service	Independent Agency	36	169	0.53	165	0.52
Veterans Benefits Administration	Ex. Sub-agency	8	10	0.45	10	0.50
Veterans Employment and Training Service	Ex. Sub-agency	3	8	0.42	8	0.47
Veterans Health Administration	Ex. Sub-agency	14	11	0.40	18	0.46

Table A.2: Sample of US agencies and descriptive statistics. *(continued)*

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Beliefs**	**Avg. Mentions**	**Avg. Beliefs**
Wage and Hour Division	Ex. Sub-agency	1	6	0.41	20	0.60
Western Area Power Administration	Ex. Sub-agency	3	13	0.46	24	0.40
Womens' Bureau	Ex. Sub-agency	1	5	0.55	12	0.62

A.2 Measuring Partisan Beliefs About Bureaucracy

The core assumption of the measurement strategy is that we can capture legislators beliefs by studying what legislators say in legislatures.

The corpus from which I estimate the word embeddings are all the legislative speeches from 1980 to the most recent available data in the two chambers of the US Congress and the UK House of Commons. US congressional speeches were downloaded from the Social Science Data Collection of Stanford University (Gentzkow, Shapiro, and Taddy 2018), while UK parliamentary speeches were downloaded from UK Data Service ReShare (Blumenau 2021), for a total of almost 4,9 million speeches (2.52 mln speeches for the US and 2.37 for the UK). I created a list of agencies as comprehensive as possible from both existing datasets and government official websites, for a total of 636 bureaucratic bodies, 285 for the US and 351 for the UK.[15]

I employ the unsupervised learning algorithm GloVe (Pennington, Socher, and Manning 2014), a count-based model that produces vector representations of words by doing dimensionality reduction on a co-occurrence matrix. The first step is to create a term co-occurrence matrix X of dimension $V \times V$, where V is a vocabulary consisting of all the unique tokens that appear in the corpus, where each corpus consists of all speeches of legislators belonging to each party in any given year. Each element X_{ij} is a number representing how many times word i co-occurs in the context of word j, with the context simply being a pre-defined window of words whose size depends on the particular task at hand. Let X_j be the sum of the co-occurrences of any word i with the context word j (i.e., the sum of the j^{th} column), and $P(i|j) = X_{ij}/X_j$ be the probability that word i appears in the context of word j.

Word vectors are then estimated with a neural network, namely a statistical model containing one layer of latent variables (the dimensions of the word vectors) between the textual input (term co-occurrence matrix) and the output data (the word vectors). To avoid the model from weighting all the co-occurrences equally, word vectors are estimated for every word in V by training a log-bilinear model with a weighted least-squares objective that tries to predict the context word j in which word i is used. Very summarily, the model minimises the following equation J,

$$J = \sum_{i=1}^{V} \sum_{j=1}^{V} f(X_{ij})(w_i^T w_j + b_i + b_j - log(X_{ij}))^2$$

where $V = \{v_1, v_2, ..., v_V\}$ is the vocabulary, w_i is the vector of the target word, w_j is the vector of the context word, and b_i and b_j are scalar bias terms. $f(X_{ij})$ is a function that determines the weight to each pair of words based on how often they co-occur; pairs of words that co-occur more often will have greater weight. The final output is a word embedding for every word in the vocabulary.

I train the GloVe algorithm on a local corpus of parliamentary speeches for every year, every country, and for the two main parties in the US and the UK (Republican, Democratic, Labour, Conservative). I follow standard practice in text-analysis and I lemmatise the tokens, remove punctuations, digits, capitalisation, and stop-words to increase the precision of the estimation. I also remove two-character words for they are deemed to carry no semantic information. Agencies referred to in more than one way (e.g., CIA and Central Intelligence Agency) were replaced in the text with standardised tokens. I then create a vocabulary with all the tokens appearing at least five times in all the corpus, because words appearing very few

[15]For the US, I used the samples in Bertelli et al. (2013) and Selin (2015). For the UK, I created a list of agencies from gov.uk/government/organisations.

times do not convey semantic information. I create a term co-occurrence matrix specifying a window size of 12 tokens and estimate 300-dimensional word vectors with a weighting function $X_{max} = 10$. This means that any pair of words for which the co-occurrence count is greater than 10 will receive a weight of 1, whilst the other weights $w_i \in [0, 1)$. I estimate the model through 100 iterations, with a convergence threshold of 0.001, and a learning rate appropriate to the size of the corpus, equal to 0.1. I use these parameters because they are deemed to be the most appropriate for semantic tasks (Spirling and Rodriguez 2019).

A.3 Scandals: Qualitative Description

I focus on three major scandals affecting US bureaucracy that uncontroversially undermined the reputation of agencies.

- FEMA "was criticized for poor preparation and a slow response to Hurricane Katrina" (Roberts 2006, 57) and its response the Hurricane Katrina on 23 August 2005 is still acknowledged as "another grand failure for FEMA" (Timeline 2017).

- A report published by the US Treasury Inspector General for Tax Administration on 14 May 2013 found that the Internal Revenue Service targeted conservative groups applying for tax-exempt status (TIGTA 2013).

- Finally, "the Department of Veterans Affairs in 2014 was embroiled in a scandal over massive wait times in its health-care system." In some hospitals, the staff falsified appointment records to appear to meet the 14-day target. Some patients died while they were on the waiting list (Vox 2015).

A.4 Text Pre-Processing

I implement the dictionary-based measurement through the following steps: the speeches are the same used to estimate beliefs (see Section A.2 for more details on the corpus of speeches). Since I do not need a minimum number of speeches to measure legislators' use of facts and evidence, I keep all the speeches given by every political party. To compare speeches about bureaucracy, I keep only the speeches which mention at least one agency. I removed punctuation and converted all the tokens to lower case. Agencies referred to in more than one way (e.g., CIA and Central Intelligence Agency) were replaced in the text with standardised token.

A.5 Dictionary Approach: Validation

Dictionary-based approaches to analyse text are deemed to be highly context-dependent and therefore need careful validation (Grimmer and Stewart 2013). Words' semantics can in fact change from one context to another. This issue is particularly concerning for sentiment analysis tasks, for the valence of words is likely to change over time and across domains.

The "fact-dictionary" derived from the LIWC lists of words I use to measure legislators' use of facts and evidence when arguing about bureaucracy has been extensively validated by (Hargrave and Blumenau 2020) in an almost identical setting as the one I study here: legislative speeches in the UK House of Commons. Moreover, context-dependence seems less problematic for facts and evidence-related dictionaries, whose words are more representative of quantities and objective attributes and less reflective of emotions. To back this claim with data, I compare the estimates of the dictionary approach with a manually labelled corpus of text from a very different context: posts and comments of medical online forums on breast cancer, crohn, and various allergies.

The corpus is assembled by Carrillo-de-Albornoz, Vidal, and Plaza (2018), who train a classification model to estimate patients' opinion about health services. Coders classified each sentence of each post as communicating "experience," "fact," or "opinion." The benchmark I use to assess the validity of the dictionary is thus the number of sentences classified as "fact" in each post. I then apply the dictionary-approach to the corpus of posts (N = 480) and I model the relationship between the dictionary and manual estimates. Table A.3 below reports regression estimates of OLS and various count models where the number of fact sentences is regressed on the *fact* estimates consisting of the sum of the *tf-idf* of each term in the fact dictionary that appears in the post, as per Equation *tfidf*. The coefficients suggest that the *Dictionary Measure* is a strong predictor of the number of fact sentences in forum posts. This strengthens our confidence of the validity of the dictionary for capturing the use of facts and evidence in texts and its weak dependence of context.

Estimators:	(1) OLS	(2) Poisson	(3) Neg. Bin.	(4) Logit
Variables				
Dictionary Measure	2.760*	0.817***	0.773***	0.540**
	(1.296)	(0.145)	(0.110)	(0.198)
Fit statistics				
Observations	480	480	480	480
Squared Correlation	0.180	0.338	0.331	0.014
Pseudo R^2	0.038	0.159	0.048	0.011
BIC	2,434.170	1,865.475	1,585.082	651.933
Over-dispersion			1.082	

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table A.3: Manually labelled fact scores regressed on fact scores produced with automated text analysis. Different estimators. Logistic regression with dichotomised outcome = 1 if number of fact sentences in post > 1, 0 otherwise.

In the following table I report the list of the LIWC facts dictionary.

LIWC Dictionary - Statistical Facts and Evidence
000, 000-day, 000-hour, 000-mile, 000-minute, 000-month, 000-odd, 000-page, 000-plus, 000-to-000, 000-week, 000-year, 000-year-old, 000-year-olds, 000,000, 000a, 000b, 000g, 000m, 000nd, 000p, 000rd, 000s, 000st, 000st-century, 000th, 000th-century, add, added, adding, adds, amount, amounts, another, approximately, average, billion, billion-worth, billions, bit, bits, bunch, chapter, couple, double, double-dip, doubled, doubling, doubly, dozen, dozens, eight, eighteen, eighth, either, eleven, entire, entirely, entirety, equal, equalisation, equalise, equalities, equality, equally, equals, every, extra, fewer, fifteen, fifth, fifthly, fifths, fifty, first, five, four, four-year, four-year-old, four-year-olds, fourth, fourthly, group, group's, grouped, grouping, groupings, groups, half, hundred, hundreds, inequalities, inequality, infinite, infinitely, least, less, lot, lots, majority, many, million, million-worth, millionaires, millions, much, multiple, nine, none, one, part, partly, percentage, percentages, piece, pieces, plenty, quarter, quarterly, quarters, remaining, sample, samples, scarce, second, section, series, seven, seven-day, seven-year, seven-year-olds, sevenoaks, seventh, several, single, six, six-month, six-week, six-year, sixth, sixth-form, sixthly, somewhat, ten, tenth, third, thirty, thousand, thousands, three, total, trillion, triple, tripled, twelve, twenty, twice, two, variety, various, whole, zero

Table A.4: LIWC list of statistical facts and evidence used to measure legislators' argumentative style when arguing about bureaucracy. 000 captures numbers.

A.6 Robustness Checks

In Tables A.5 and A.6 below I show that the results reported in Tables 2 and 3 are robust to clustering SE at party-prime minister and party-presidency level.

DV:	Beliefs [0,1]		
Country:	UK		
Model:	(1)	(2)	(3)
Variables			
Party-Govt. Partisan Align.	0.027***	0.027***	0.031**
	(0.004)	(0.006)	(0.010)
Fixed-effects			
Party	✓	✓	
Year	✓		
Agency	✓		
Agency-Year		✓	
Party-Agency			✓
Year-Agency			✓
Fit statistics			
Observations	2,622	2,622	2,622
R^2	0.257	0.594	0.652
Within R^2	0.009	0.017	0.022

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

Table A.5: Partisanship and Beliefs, UK Data. OLS estimates with SE clusterd at party-prime minister level. Units are party-agency-year observations.

DV:	Beliefs [0,1]					
Country:	US					
	Party-Government				Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Party-Govt. Partisan Align.	0.018***	0.019***	0.030***	0.011**		
	(0.002)	(0.003)	(0.005)	(0.003)		
Party-Agency Id. Dist.			0.010			-0.027
			(0.033)			(0.043)
Party-Agency Partisan Align.				0.017	0.019	
				(0.015)	(0.016)	
Fixed-effects						
Party	✓					
Year	✓					
Agency	✓					
Party-Agency		✓	✓	✓	✓	✓
Year-Agency		✓	✓	✓	✓	✓
Fit statistics						
Observations	6,874	6,874	1,674	1,340	1,340	1,674
R^2	0.273	0.682	0.715	0.684	0.683	0.706
Within R^2	0.006	0.014	0.035	0.008	0.002	0.003

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

Table A.6: Partisanship and Beliefs, UK and US Data. OLS estimates with SE clusterd at party-presidency level. Units are party-agency-year observations.

In Table A.7 below I show that the results hold when subsetting the data to the agencies for which data on ideology, party-agency, and party-government partisan alignment are available.

DV:	Beliefs [0,1]				
Country:	US				
	Party-Government			Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Party-Govt. Partisan Align.	0.018*	0.033**	0.023*		
	(0.007)	(0.011)	(0.008)		
Party-Agency Partisan Align.			0.005		0.021
			(0.027)		(0.020)
Party-Agency Id. Dist.		0.020		-0.006	
		(0.037)		(0.038)	
Fixed-effects					
Party-Agency	✓	✓	✓	✓	✓
Year-Agency	✓	✓	✓	✓	✓
Fit statistics					
Observations	1,164	606	552	606	552
R^2	0.710	0.705	0.682	0.693	0.675
Within R^2	0.013	0.040	0.023	0.000	0.001

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

Table A.7: Partisanhsip, Ideology, and Beliefs, US Data. Robustness tests with a subset of data including only the 21 agencies for which there is data on beliefs, ideology, and partisanship. OLS estimates. Units are party-agency-year observations.

In Table A.8 below I report regression estimates from a sub-sample of the dataset where the total number of mentions of agencies is above the median value (i.e., 90 for the UK and 105 for the US).

DV:	Beliefs [0,1]					
Country:	UK			US		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Party-Govt. Partisan Align.	0.075***	0.075***	0.077***	0.016***	0.016***	0.032**
	(0.007)	(0.010)	(0.011)	(0.003)	(0.004)	(0.010)
Party-Agency Id. Dist.						0.007
						(0.038)
Fixed-effects						
Party	✓	✓		✓		
Year	✓			✓		
Agency	✓			✓		
Agency-Year		✓				
Party-Agency			✓		✓	✓
Year-Agency			✓		✓	✓
Fit statistics						
Observations	1,304	1,304	1,304	3,426	3,426	1,176
R^2	0.318	0.626	0.667	0.302	0.700	0.729
Within R^2	0.081	0.138	0.144	0.005	0.011	0.043

Clustered (Party-Gen. Elections/Congress) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table A.8: Partisanhsip, Ideology, and Beliefs, UK and US Data. Robustness tests on limited sample where agencies' number of mentions is above the median. OLS estimates. SE clustered by party-general elections for the UK and by party-congress for the US. Units are party-agency-year observations.

In Table A.9 below I report the regression results with additional covariates from D. E. Lewis (2008). Agency politicisation is measured as the ratio of managers who are presidential appointees, whereas (authorised) budget and employees are measured in dollars and units. Data available only from 1988 to 2005.

DV:	Beliefs [0,1]		
Country:	US		
Model:	(1)	(2)	(3)
Variables			
Party-Govt. Partisan Align.	0.028***	0.036**	0.039**
	(0.004)	(0.009)	(0.011)
Party-Agency Id. Dist.		0.015	0.012
		(0.019)	(0.027)
Politicisation	-0.007	-0.062	-0.062
	(0.007)	(0.056)	(0.052)
Log N. Employees	0.030	0.009	0.009
	(0.020)	(0.066)	(0.075)
Log Budget	-0.003	-0.005	-0.005
	(0.003)	(0.003)	(0.002)
Fixed-effects			
Party	✓	✓	
Year	✓	✓	✓
Agency	✓	✓	✓
Party-Agency			✓
Fit statistics			
Observations	1,978	508	508
R^2	0.301	0.334	0.374
Within R^2	0.017	0.026	0.029

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

Table A.9: Partisanhsip, Ideology, and Beliefs, US Data. Robustness tests with additional covariates. OLS estimates. Units are party-agency-year observations. Total number of mentions of agencies always included.

In Table A.10 I report falsification tests of the difference-in-differences strategy of Section 5.1. Placebo post-treatment indicators have been set to 2 and 4 months before the true date of the scandal and the sample consists of statements given 2 months before and after the placebo scandal date.

DV:	Pr(Positive Statement = 1)	
Placebo Scandal Date (Months before true scandal date) :	-2 months	-4 months
Model:	(1)	(2)
Variables		
Leg.-Govt. Partisan Alig.	0.003	-0.062
	(0.063)	(0.052)
Placebo Post-Scandal	-0.032	-0.082
	(0.082)	(0.078)
Leg.-Govt. Partisan Alig. × Placebo Post-Scandal	-0.002	0.050
	(0.057)	(0.059)
Fixed-effects		
Legislator	✓	✓
Month-Year	✓	✓
Agency	✓	✓
Fit statistics		
Observations	1,831	1,950
R^2	0.182	0.187
Within R^2	0.000	0.001

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

Table A.10: ATT of legislator-government partisan alignment on the probability of giving a positive statement about bureaucracy with placebo post-treatment indicator. Sample consists of statements given 2 months before and after the placebo date of the scandal.

In Table A.11 below I replicate the estimation of Table 5 for the US conditioning on the ideological distance between the agency and the legislator giving the speech.

DV:	Facts (*tf-idf*)		
Country:	US		
Window Size:	20	50	Total
Model:	(1)	(2)	(3)
Variables			
Leg.-Govt. Partisan Alig.	-0.024***	-0.017**	0.006
	(0.007)	(0.005)	(0.009)
Leg.-Agency Id. Dist.	-0.009	0.003	-0.017
	(0.023)	(0.018)	(0.021)
Fixed-effects			
Legislator	✓	✓	✓
Agency	✓	✓	✓
Date	✓	✓	✓
Fit statistics			
Observations	99,771	99,771	99,771
R^2	0.196	0.234	0.298
Within R^2	0.121	0.159	0.198

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

Table A.11: Partisanhsip, Ideology, and Beliefs, US Data. OLS estimates. Dependent variable if tf-idf of facts-words in speeches. Controls include legislator's age and speech length (log number of words). The estimated effect of legislator-government alignment remain distinguishable from 0 and in the expected direction even when conditioning on legislator-agency ideological distance.

In Table A.12 below I replicate the estimation of Table 5 replacing legislator-government partisan alignment (treatment) with two alternative treatments: legislator-agency partisan alignment and legislator-agency ideological distance.

DV:	Facts (*tf-idf*)					
Country:	US					
	Ideology			Congruence		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Leg.-Agency Id. Dist.	0.013	0.019	-0.023			
	(0.023)	(0.018)	(0.020)			
Leg.-Agency Partisan Align.				0.011	0.002	-0.039
				(0.018)	(0.013)	(0.021)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	99,771	99,771	99,771	41,100	41,100	41,100
R^2	0.196	0.234	0.298	0.271	0.303	0.334
Within R^2	0.121	0.159	0.198	0.128	0.162	0.210

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

Table A.12: Legislator-Agency Partisan Alignment and Ideological Distance. US Data. OLS estimates. Dependent variable if tf-idf of facts-words in US speeches. Controls include legislator's age and speech length (log number of words).

In Table A.13 below I replicate the estimation of Table 5 using the absolute frequency of facts-words instead of the *tf-idf* of facts-words as dependent variable.

DV:	Facts (Abs. Frequency)					
Country:	US			UK		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Leg.-Govt. Partisan Alig.	-0.135***	-0.207***	0.291	-0.083***	-0.147***	0.053
	(0.028)	(0.049)	(0.322)	(0.021)	(0.039)	(0.192)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	280,337	280,337	280,337	171,155	171,155	171,155
R^2	0.205	0.236	0.252	0.166	0.232	0.667
Within R^2	0.130	0.164	0.161	0.076	0.138	0.603

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

Table A.13: Robustness Analysis: Argumentative Style, US and UK Data. OLS estimates. Dependent variable is the absolute frequency of facts-words in speeches. Controls include legislator's age and speech length (log number of words) and, for UK data only, legislator's seniority (i.e., log number of days in house) and whether the legislator holds government positions.

A.7 Test 4: Diagnostics

Table A.14 reports full regression results of Model (5).

	No Interaction			Multiplicative Interaction		
DV:	Total	Casework	Policy	Total	Casework	Policy
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Leg.-Agency Id. Dist.	0.015	0.014	-0.001	0.114***	0.074**	0.113***
	(0.020)	(0.019)	(0.018)	(0.027)	(0.025)	(0.025)
Leg.-Govt. Partisan Alig.				0.084***	0.052***	0.096***
				(0.014)	(0.014)	(0.015)
Leg.-Govt. Partisan Alig. × Leg.-Agency Id. Dist.				-0.191***	-0.105***	-0.224***
				(0.033)	(0.031)	(0.035)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Congress	✓	✓	✓	✓	✓	✓
Fit statistics						
Observations	17,552	16,455	16,455	17,552	16,455	16,455
R^2	0.503	0.500	0.430	0.504	0.500	0.431
Within R^2	0.005	0.002	0.006	0.007	0.002	0.009

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

Table A.14: Linear probability models. Binary variable, 1 for informal legislator oversight of agency (for each type of request: casework, policy, and both). Covariates include committee membership, committee chair, ranking, budget in millions of dollars, agency politicisation, log number of staff.

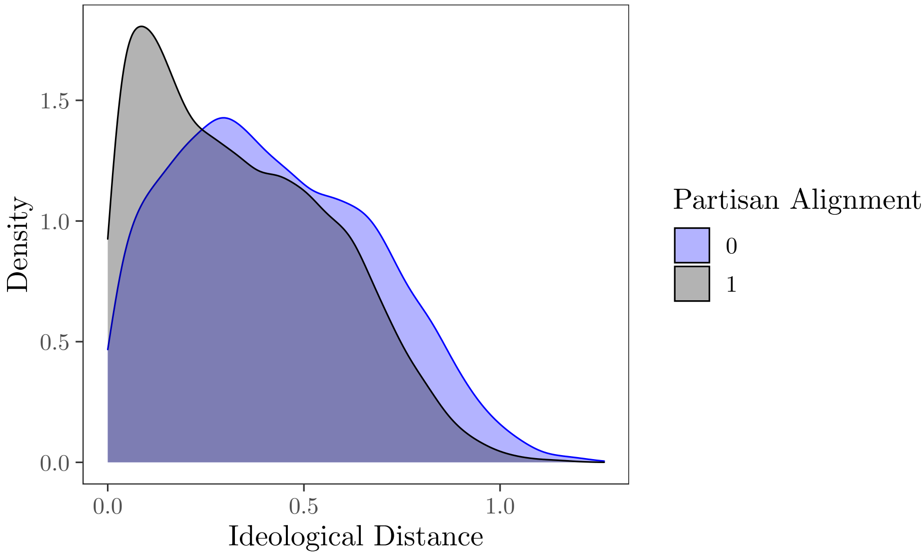

Figure A.1: Density function of ideological distance by partisan-alignment with the government.

	Leg.-Agency Partisan Alignment	
	No	**Yes**
N. Observations	8,320.00	9,232.00
Leg.-Agency Id. Dist.		
Mean	0.43	0.34
SD	0.25	0.24
Min	0.00	0.00
1st Quartile	0.22	0.14
Median	0.40	0.31
3rd Quartile	0.62	0.52
Max	1.27	1.20

Table A.15: Summary statistics of the treatment ideological distance variable (Leg.-Agency Id. Dist.) across two values of moderators. The absence of marked differences between the values of the two columns suggests there is common supoprt of the moderator.