Reputation, Partisanship, and Ideology in the Administrative State

Luca Bellodi
Student Number 16092574

Department of Political Science
University College London
United Kingdom

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

Contents

Declaration

I, Luca Bellodi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

In democratic government, elected politicians are accountable to voters for the policies they pass. Policies, however, are administered by bureaucracies with possibly large levels of discretion and different preferences from those of elected politicians. The relationship between elected politicians and unelected bureaucrats is therefore both empirically and normatively important: politicians ought to make sure that their electoral mandate translates into policies; and unelected bodies shall be held accountable by politicians, in order to ensure that the administration of policies remains aligned with the will of government. The interactions between politicians and bureaucracy is a classic topic in political science research. While early work on the study of politicians-bureaucracy interactions considered bureaucracy as a passive actor controlled by politicians, recent advancements in the scholarship have demonstrated how bureaucracies can become autonomous actors able to influence their political masters. In this dissertation I study how reputation, partisanship, and ideology affect two main types of interactions: politicians influencing bureaucracies and bureaucracies influencing politicians.

Reputation. The political role of bureaucracies consolidates with theories of bureaucratic reputation, which posit that autonomous policy-making occurs when agencies can build a reputation for uniqueness among multiple audiences. However, the literature lacks a valid measure of reputation that changes over time and across agencies, thus limiting the reach of the theory. In Chapter 1, I introduce a new measure of bureaucratic reputation which applies word-embedding techniques to legislative speeches and show how scholars can now test theories of reputation more rigorously and answer new questions in political science.

Partisanship. In the hierarchy of government, politicians are superior to bureaucracies, and they can oversee bureaucracies to ensure they align to politician' directives. However, politicians' control of the bureaucracy is a trade-off which is subject to political constraints. In Chapter 2, I show that partisanship biases legislators' statements about bureaucracy and hinders their ability to hold agencies to account.

Ideology. The legitimacy of the political and autonomous role of unelected bureaucrats

rest with their ability to produce information that can be used by politicians to reduce uncertainty over policy outcomes. However, information can also be a channel of influence for the bureaucracy. In Chapter 3, I demonstrate how bureaucratic influence in the legislative process – namely the extent to which legislators use the information produced by bureaucracy – decreases with ideological divergence between legislators and bureaucratic bodies, and how statutory independence can reduce the salience of the ideological divide between legislators and agencies.

These theoretical contributions are combined with methodological advancements that expand the use of computational methods to the study of bureaucracy and politicians. I collect large original data and introduce several innovative techniques to measure bureaucratic reputation, politicians' statements about bureaucracy, and legislators' use of bureaucratic information in the legislative process, showing how these measurement strategies can contribute to classic and new questions about political-administrative interactions.

Impact Statement

This dissertation is about bureaucracies and legislators. It addresses two normatively important questions.

The first question is about bureaucratic accountability. Elected politicians generally lack the time and expertise to pass detailed legislation, implement policies, or manage large programmes, and they therefore delegate these tasks to bureaucracies. Bureaucracies, however, do not have democratic authority and therefore need to be held in check in order to ensure that the link between citizens and representatives is not distorted by the action of unelected officials. For accountability to work, politicians need to monitor the performance of agencies, update their beliefs, and respond accordingly. However, in this dissertation I show that partisanship biases legislators' statements about bureaucracy and their ability to effectively hold agencies to account.

The second question is about evidence-based policy-making. In increasingly complex societies, expertise and evidence-based policy are key pillars of good governance and high-quality policies. Bureaucracies, because of their sectoral expertise, play an important role in the provision of information and evidence that can inform the decisions taken by legislators. A crucial question is therefore how to facilitate a constructive communication between legislators and bureaucratic bodies. In this dissertation I examine the role of ideology in inhibiting politicians' use of evidence produced by bureaucratic agencies. I find that, the more ideologically distant legislators and bureaucracies are, the less likely it is that legislators use the evidence produced by bureaucracies when discussing policy in legislatures. However, I find that when bureaucratic agencies are insulated from political influence, ideology plays a weaker role. Bureaucratic independence can therefore mend the ideological divide between legislators and bureaucrats and facilitate evidence-based policy-making.

The findings of this dissertation can help researchers and practitioners understand how partisanship and ideological conflict affect the smooth functioning of accountability mechanisms and evidence-based policy-making. A good understanding of these effects can inform the institutional design of bureaucratic agencies and oversight procedures, in the attempt at curbing the suboptimal incentives that legislators have as a result of partisanship and ideological leaning.

Future research on bureaucratic and legislative politics can benefit from the methodological advancements presented in this dissertation. Data-driven analysis of bureaucracy-legislators interactions is difficult, just as it is hard to look inside the black box of government. However, in this dissertation I show that, by leveraging new methodologies widely used in other fields of research such as computer science and computational linguistics, political scientists can exploit alternative sources of data such as parliamentary debates and committees' transcripts to measure important quantities of interest, from the reputation of agencies, to politicians' beliefs, and legislators' use of the information produced by bureaucracy. By introducing these methods in the study of bureaucracy, this dissertation opens new research paths that can contribute to a better understanding of the impact of bureaucracy on democratic governance.

Acknowledgments

This dissertation is the final product of a long journey which embraced my intellectual, academic, and personal growth. I hope it will contribute to our understanding of how elected politicians interact with unelected bureaucracies in modern democracies. I also hope it will convince my parents I am not employed by MI5.

I am indebted to so many who helped me navigate the PhD, I would not be able to thank them all here. However, one exception is in order. I wish to thank my supervisors, Colin Provost, David Coen, and Tom O'Grady for pushing me to do better and for guiding me through the challenges of becoming an independent researcher.

In the last years I have learned how hard it is to build credible counterfactuals. However, I can easily say that I would have not been able to write "The End" without the support of my family and of my colleagues and now friends at 31 Tavistock Square, London. The biggest risk of the PhD is to pause life until this big enterprise is completed. I'm sure I've done that many times, but they reminded me to put tough moments in perspective and to move on with joy.

My dad often teases me about my work, asking me "you search, search, and search, but where's the treasure?". The PhD made me cross paths with many people, and I like to think that when I met Alin, my wife, as a long-term unintended consequence of a research visit at Columbia, New York, I found that too. My treasure. She helped me through the last sprint of this PhD and into our new life together. No doubts she deserves the last, honorary line of my acknowledgements.

Introduction: From Congressional Dominance to Bureaucratic Autonomy

On December 7, 2016, President-elect Donald Trump announced his intention to nominate Edward Scott Pruitt as the fourteenth Administrator of the Environmental Protection Agency. Mr. Pruitt, as a self-declared "leading advocate against the EPA's activist agenda", rejected scientific consensus on climate change and his ideas were in direct contrast with those of the previous Obama administration.

In 1972, John Edgar Hoover died, after serving for over 48 years as director of the US Federal Bureau of Investigation. His legacy will always be characterised by two antithetical primacies: building the FBI into a modern organisation and an investigative powerhouse, while using it to threaten politicians and control presidents.

These controversial pages of American politics remark upon the two-way interaction between politics and administration. On the one hand, politicians influencing bureaucracy; on the other hand, bureaucrats influencing politicians. These two dynamics are the main topic of this dissertation.

Over the last decades, scholarly work on the interaction between politicians and bureaucracy underwent a long intellectual journey which, for simplicity, I shall present clustered around three moments. First, theory posited and empirics tried to support the idea that bureaucratic bodies were passive actors in the policy-making process. As enshrined in the iconic concept of "congressional dominance," bureaucracy was considered under tight control of politicians, who could use it consistently with their political and electoral gains. Second, with the consolidation of the principal-agent model to the study of political institutions, control started being characterised as the result of a trade-off that politicians made *vis-à-vis* information asymmetries and costly monitoring of agency behaviour. Bureaucracies were portrayed as rational actors with preferences and resources they could deploy in order to attain desired outcomes. Third, the downward direction to the study of bureaucracy, one that used to conceive politicians as first movers, lost its near-monopoly in the scholarship: bureaucratic expertise and authority were now seen as sources of political power able to influence elected politicians, the political agenda, and policy outcomes. The question was no longer exclusively "when do politicians control bureaucracy?" yet a new reversed question emerged in scholarly work: "when do bureaucracies control politicians?"

In this dissertation I build on this evolving literature which portrays bureaucracies from administrative subordinates of elected politicians to autonomous policy-makers. Once the relationship between politicians and bureaucracies transcends formal and statutory boundaries to fully politicise, new questions in the study of bureaucracy emerge. Here, I focus on one overarching question: How does partisanship and ideology affect politicians' and bureaucrats' competition for power and political influence? I will approach this question from three different angles. First, I will document the "making" of bureaucratic autonomy through bureaucratic reputation. Second, I will look at how partisanship alters politicians' incentives and ability to control bureaucracy. And finally, I will show that ideological agreement between politicians and bureaucracy can bolster bureaucracies' influence in the legislative process.

The (Myth) of Congressional Dominance

An influential literature rooted in social choice theory started taking shape in the early 80s which posited that, mostly with reference to the United States, bureaucracy was controlled by Congress (Fiorina 1981; McCubbins and Schwartz 1984; Mccubbins 1985; McCubbins, Noll, and Weingast 1987; Weingast and Moran 1983). The theory was centred on an empirical fact and a theoretical assumption. The assumption was that legislators seek re-election and congressional politics is organised in a way to equip legislators with the resources necessary to remain in power. The fact was about the importance of congressional committees for policy outcomes. Legislators self-assign to the most salient committees for their electoral district. Committees specialise in few specific policy issues and thanks to their agenda-setting prerogatives, enjoy large influence over policy. Congress hence wield significant rewards and sanctions in imposing their will on the bureaucracy: it controls agencies' budget – necessary for agency survival and stability –, it can engage in oversight activities, and can pass new legislation aimed at altering the organisation and tasks of agencies. Congress can reject and obstruct proposed projects, terminate top bureaucrats' careers, and control appointments. The resulting incentive mechanisms produce agency compliance with congressional directives (Weingast and Moran 1983).

Empirical support for this theory was mostly offered by single-agency analyses that were used more as real-world examples to display the theoretical assertions rather than as a research design to rigorously test empirical predictions. Such fragile empirical foundations of the theory of congressional dominance crumbled under its own logic when scholars of Congress and the bureaucracy started to realise that the theory was in fact leaving very little space to the bureaucracy. It was increasingly clear that the theory which was claiming to explain bureaucratic behaviour was in fact purely about Congress and committees' agenda-setting power. The empirical analyses demonstrated that the concept of control itself was vague and – most importantly – no attention was paid to the the bureaucratic side of the story, to the preferences of agencies, to the resources they can marshal to achieve their goals, and to their markedly political role. These theoretical and empirical weaknesses became evident when the scholarship aligned with the new economics of organisation: the principal-agent model (Moe 1984).

The main contribution of the principal-agent model to the study of the relationship between elected politicians and bureaucratic bodies was to remind scholars that rationality did not inform the choices of principals alone, but of agents too. Control could not be assumed as a result of acknowledging the influential role of committees in policy-making. Conversely, principals *try* to control the behaviour of their agents, who are in turn driven by their own interests and make decisions on the basis of information only imperfectly available to the principals (Hammond and Knott 1996; G. J. Miller 2005). Moreover, political principals find it costly to monitor and oversee agencies, and sanctioning mechanisms are themselves subject to political constraints, inter-committee coordination, and presidential approval (Moe 1984). Control is ultimately a trade-off and principals might be better off by leaving the bureaucracy with large levels of discretion. For instance, early work on delegation of authority to bureaucracy shows how Congress delegates larger discretion to bureaucracies ideologically similar, and when the policy sector in which the agency operates is highly complex (Epstein and O'Halloran 1999).[1]

Once the scholarship on the bureaucracy takes hold of the political and active role of agencies in their interaction with politicians, considering principals first movers is no longer a natural choice in the study of bureaucracy. In fact, further advancements in the literature started to question this top-down, principal-centred approach, proposing a

[1]For simplicity, I make no reference to the literature on multiple principals, chiefly Congress, Presidents, and Courts. For a review see Moe and Wilson (1994); Moe (2012).

new type of relationship that unfolds along a two-way street. If politicians can control bureaucrats, bureaucrats too can use their expertise and authority to influence the decisions and preferences of politicians and exert influence on the choices that should ultimately rest with democratically-elected officials. Krause (1996), for instance, depicted politicians' decisions to control the agency through budgetary allocation as the result of the interactions between both the agency and the principals. In his study of the US Securities and Exchange Commission (SEC), he shows that the budgetary preferences of the government with respect to the SEC are influenced by the SEC's regulatory performance. Similarly, Potter (2019) show how bureaucrats can influence the oversight by the President, Congress, and interest groups – one of the mechanisms whereby politicians were supposed to "dominate" the bureaucracy – by strategically manipulating the timing, quality, or space allowed for consultation of their rule-making activity (Bolton, Potter, and Thrower 2016; Potter 2017), and Moe (2006) shows that bureaucrats can take part in the electoral process to choose their own political principals. What this strand of the literature shows is that bureaucracy can ultimately act like autonomous political actors in order to strike a better deal for themselves.

Bureaucratic autonomy thus became a central topic in the literature on bureaucratic politics, and rapidly moved to scholarly traditions outside the United States. Often labelled as *The Regulatory State* (Majone 1994), a large scholarship emerged on the phenomenon of "agencification," whereby governments established statutory independent bureaucratic agencies responsible for economic or social regulation. While initially focused on the European Union and on EU member states, the study of independent regulatory agencies became a general standard for the study of the bureaucracy, and was applied to other continents (Pavón Mediano 2018) and on a global scale (Jordana, Levi-Faur, and Fernández-i-Marín 2011). These accounts aimed at explaining why elected politicians increasingly delegated large stocks of discretion to unelected officials politically insulated from politicians' control (also referred to as non-majoritarian bodies (Coen and Thatcher 2005; Thatcher and Sweet 2002)). A central argument put forward in the literature hinges on the credibility of government commitment to policies. In order to minimise the risk of policy reversal posed by changing governments, policy authority was delegated to bodies immune from political influence (Gilardi 2002, 2007). The most striking example of this phenomenon is the establishment of independent central banks, something that, although nowadays appears the natural set-up of monetary institutions and policy, it represented an epochal change in the second half of the twentieth century (Cukierman, Webb, and Neyapti 1992). In the empirical context of national and European bureaucratic bodies, this strand of the literature too interpreted the relationship between political principals and bureaucratic agents as unfolding along multiple directions. Scholarly work highlighted the role of bureaucratic bodies in the formulation of policies (Bach 2012; Maggetti 2009), in their flexible relationship with overseers (Busuioc 2009; Schillemans and Busuioc 2015a), and in the interaction with other bureaucracies (Levi-Faur 2011; Maggetti 2014).

One important theoretical account which explains how agencies become autonomous actors is centred on the concept of bureaucratic reputation. Bureaucratic reputation becomes a prominent theory of bureaucratic politics with Carpenter's *The Forging of Bureaucratic Autonomy* (2001a), where he shows how, during the US Progressive Era, bureaucratic agencies and departments built and cultivated strong reputations across multiple audiences and eventually managed to autonomously decide policies and programmes and exert influence on even the most powerful politicians. Since then, a growing body of scholarship, theoretical and empirical, adopted a reputation-based perspective to the study of agencies. Accountability became the practice of sustaining reputation across multiple forums (Busuioc and Lodge 2016, 2017), external communication was seen as a tool for bureaucratic bodies to handle threats to their reputation (Maor and Sulitzeanu-Kenan 2016), and bureaucratic outcomes were partly shaped by reputational concerns (Krause and Douglas 2005).

The concept of bureaucratic autonomy consolidates the new role of bureaucracies in the political arena. From passive actors controlled by Congress and politicians to political protagonists of the policy-making process. This dissertation builds on these recent advancements in the literature on politicians-bureaucracy relations. It takes stock of the political role of the bureaucracy and of the two-way street along which principals influence agents and agents influence principals. The ambition of this dissertation is to expand the scholarship on politicians-bureaucracy interactions both theoretically and empirically. This dissertation innovates theories of congressional control of bureaucracy by considering legislators' oversight activity as limited by partisan considerations. In support of the idea that control is a trade-off which is subject to political constraints, I show how partisanship hinders legislators' ability to hold agencies to account. Moreover, this dissertation provides a solid investigation of the influence exerted by bureaucracies on legislators via expertise and information. The literature summarised above highlights how bureaucracies can use their expertise as a political resource to attain policy goals. By showing how legislators use the information produced by bureaucracy, this dissertation strengthen the upward direction of influence which starts from bureaucracies and arrive to legislators making policies. The theoretical innovations are backed by methodological advancements in the empirical study of bureaucracy. Each chapter, which is summarised in the sections below, uses big unstructured data – e.g., legislative speeches, campaign contributions, congressional hearings reports, to name but a few – to produce estimates of agency characteristics that allow for comparative analysis and credible statistical inference. These theoretical and methodological innovations represent new powerful tools that can be deployed to advance of the scientific study of the interactions between politicians and bureaucracy and to address other questions in political science.

Scoping the Contribution

While the literature summarised above and the arguments I will put forward in this dissertation are general in scope, the empirical analysis will focus on two countries: the United States and the United Kingdom. There are two blocks of reasons why I chose these two countries for studying reputation, partisanship, and ideology in the administrative state, one pragmatic, one substantive.

The block of pragmatic reasons, which – it should be clearly acknowledged – has been equally important to the block of substantive reasons when selecting the case studies, consists of the large availability of data on bureaucracy and on legislators for these two countries, in particular for the United States. Making use of natural language processing techniques in countries with different languages poses challenges that were beyond the scope of this work, which is already empirically dense. However, because of a sort of scientific path-dependence, the quantitative analysis I present in this dissertation enormously benefited from the availability of other datasets on legislators' ideology, bureaucracies' attributes, and legislative data on the US that scholars have build over the years, and therefore is skewed towards the US case. While the main empirical enterprise is supported with data form the US and the UK, many additional tests I use to corroborate my theoretical expectations would have not be possible absent this data on US bureaucracy.

As for the substantive block of reasons, both the US and the UK are countries with strong and high-capacity bureaucracies organisationally distinct from elected government. In order to study the relationship between politicians and bureaucracy, the latter ought to be organisationally identifiable in the hierarchy of government. Organisational autonomy means that both countries have institutions to ensure that bureaucracies are held in check and that bureaucracies and elected politicians have different tasks. Second, bureaucracies in these countries enjoy significant levels of discretion and expertise, and are in charge of administering large programmes and policies. These two characteristics – organisational identity and large policy responsibilities – grant bureaucracies a sizeable endowment to compete in the political market. In fact, if bureaucracies were completely deprived of their administrative authority, they would have no resources to marshal in their interactions with politicians. Third, on the politicians side, both the US and the UK host highly organised and rooted political parties with clear ideological stances. The fact that political competition in these two countries unfolds along partisan and ideological lines represents a good premise for studying how ideology and partisanship define the relationship between politicians and bureaucracies.

These political and institutional features, while informing the selection of the case studies, also raise immediate concerns about the implications of this dissertation for the literature on bureaucracy in general. On the one hand, many of the characteristics of US and UK administrative politics are not unique. Several countries across the world have autonomously organised bureaucracies and government agencies in charge of administering policies and programmes. Similarly, many countries have strong political parties, institutions and procedures to hold bureaucracy to account, and politicians and bureaucracies in charge of different and non-overlapping tasks. However, it is hard to recall other countries with a bureaucracy as politicised as the one in the US, at least as administrative politicisation is interpreted by political commentators. While this primacy is likely to be driven by a selective attention issue – also because of the US political and economic stature – the mixture between politics and administration in the United States is arguably matchless in politically and economically similar countries. For this reason, I decided to match the analysis of the administrative state in the US with that of the UK.

In fact, these two countries nicely lend themselves to comparative analysis. While sharing many institutional and administrative features, the UK civil service is deemed a neutral body of experts to the service of the government of the day (or at least inspired to neutrality), and lacks the high levels of political contestation of US bureaucracy. The joint analysis of these two cases should therefore allow the theoretical insights of this dissertation to travel across administrative traditions with different levels of politicisation and responsiveness to elected politicians. Conversely, this work will be silent about the interactions between politicians and bureaucracies in countries where bureaucracy is not an autonomous apparatus in the hierarchy of government and where partisanship and ideology are not salient currencies of political competition. In such cases it is questionable whether it even makes sense to study the interactions between two actors so closely intertwined or where authoritarian politics jeopardises inter-party and ideological competition. Let us consider for instance countries where bureaucrats are puppets of the government, ideological clones of ministries, or obsequious servant of political elites. It seems reasonable to expect partisan and ideological conflict to be extremely thin in such contexts. Clearly, there is no conclusive list of countries where the findings of this dissertation can and cannot have the ambition to travel to, and the task of judging whether this work is useful at understanding politicians-bureaucracy relationship on a country-by-country basis is ultimately left with the reader.

Bureaucratic Reputation: Back to Measurement

Theories of bureaucratic reputation represent a golden standard for scholars interested in bureaucratic autonomous policy-making. However, the literature on bureaucratic reputation is constrained by the lack of a consistent and validated measure of reputation able to match the far-reaching scope of theoretical and conceptual work. In fact, research on bureaucratic reputation remains mostly limited to a one-agency, one-country, one-year approach.

In **Chapter 1**, I use computational linguistics techniques and propose a new method to measure bureaucratic reputation from millions of legislative speeches in the US and the UK. I first propose that what politicians say in parliament can be a good proxy of what key audiences think about agencies and present a measurement strategy that employs word-embedding techniques to estimate how agencies appear in the language space compared to very positive concepts. Most importantly, I introduce a measure of bureaucratic reputation that varies across agencies, over time, and between countries. I show the measure responds well to several validation exercises employing both qualitative and quantitative data. I provide two applications of this method showing how it can be used to systematically test existing theories as well as to open new research paths in the study of political-administrative interactions. First, I show how, consistent with Carpenter's account of bureaucratic autonomy, agencies with a better reputation report to operate – on average – with higher levels of autonomy. Second, I show how the flexibility of the proposed measure allows researchers to target audience-specific measures of reputation. I split the corpus of speeches by party and show how Republicans and Democrats, in the US, and Conservative and Labour MPs, in the UK, polarise about bureaucracy. I demonstrate that "bureaucratic polarisation" is larger for departmental and more politicised bodies, while it decreases for more independent agencies, opening new research agendas that study how partisanship and ideology contribute to politicians' evaluation of bureaucracy. These new questions are addressed in Chapter 2.

Partisanship and Bureaucratic Accountability

In **Chapter 2**, I build on the measurement strategy proposed in Chapter 1 and focus on the top-down relationship between politicians and bureaucracy, namely politicians controlling and monitoring agencies. The measure of reputation is derived from a large set of statements about bureaucracy, and in Chapter 1 I show that there are partisan differences in how political parties evaluate bureaucracies. In times of increasing partisan polarisation, namely larger inter-party and smaller within-party differences in beliefs, shall scholars study politicians and bureaucracy through a partisan lens? If partisanship is a new currency that defines what politicians should care about and which position they should endorse, politicians' ability to oversee agencies is likely to be a function of partisan considerations. In Chapter 2 I hence address the following questions: Can partisanship blind politicians and distort their ability to hold bureaucracy to account? Classical models of bureaucratic accountability assume that politicians observe what agencies do, update their beliefs, and respond accordingly. If politicians selectively hold bureaucracy to account along partisan lines, can they still be effective at overseeing agencies?

I answer these questions and present a theory of selective accountability that builds on legislators' incentives to protect the image of their party. Because politicians care about their party's image, and a bureaucracy doing a poor job has negative consequences for the reputation of the governing party, politicians will selectively evaluate bureaucracy: more positively when in government, more negative at the opposition. Because the government is responsible for the public policies administered by bureaucracy, government politicians have an incentive to assess the activities of agencies more favourably when their party is in power, and this co-partisanship makes statements about agencies more positive. Partisanship ultimately reduces politicians' responsiveness to information about bureaucracy. As a result, the bedrock of bureaucratic accountability, namely acquiring information about bureaucracy, weakens for government legislators. Therefore partisanship may hinder bureaucratic accountability.

Empirically, I apply the measurement strategy introduced in Chapter 1 and estimate the sentiment of partisan statements about hundreds of agencies and over 40 years in the UK and the US. I present two studies on two dimensions of selective accountability: selective evaluation and selective information acquisition. I find that for the governing party statements about bureaucracy are on average 3 percentage points more positive, irrespective of the party-agency ideological distance or partisan congruence. I also demonstrate how partisanship affects legislators' incentives to acquire information on bureaucracy in congressional hearings, and to use such information in legislative speeches. When there is partisan alignment between the President and the chair of a congressional committee, the probability of a bureaucracy being heard as a witness drops sharply. Consistently, I also find that government politicians are less likely to report quantitative information when discussing about bureaucracy in floor debates. By incorporating partisanship in the study of bureaucratic accountability, this chapter sheds light on a new political constraint to politicians ability and willingness to control bureaucracy.

Bureaucratic Information as a Channel of Influence

If Chapter 2 deals with the downward dimension of political-administrative relations, **Chapter 3** deals with the upward dimension: bureaucracy influencing politicians. I propose an information-based conceptualisation of influence and interpret bureaucratic influence as a function of politicians resorting to the information and evidence produced by bureaucracy when passing and debating policy. Building on cheap talk models of strategic communication (Crawford and Sobel 1982; Gailmard and Patty 2012), I argue that bureaucratic influence is more likely to occur when agencies and politicians have similar preferences over policy outcomes. Consider an expert bureaucracy sending a signal or information to a legislator who will then make a policy decision. Because the legislator cannot verify the quality or veracity of the information, truthful communication is only achieved when both the bureaucracy and the legislator have similar preferences over policy outcomes. Ideology is used by politicians as a heuristic to decide whether to to use the information produced by bureaucracy or not. The more ideologically apart, the less likely it is that legislators will use the information produced by agencies. I expand this argument by looking at the role of institutional independence on the politicisation of the information. Operating as a credibility-enhancing mechanism, agency independence reduces the salience of ideological differences and hence limit the negative effect of ideological divergence on legislators' use of the information sent by the agency.

To test this argument, I present a new measurement strategy to estimate politicians' use of bureaucratic information that uses syntactic dependency parsing. First, I build extraction rules that match several syntactic structures capturing every instance in which an actor is reporting what said by another actor. Second, I match every instance in which the "source" of the information reported in a speech is a bureaucratic agency. Third, I isolate the content of the information reported by politicians (i.e., the quote) and measure the frequency of words considered as statistical facts and evidence, and hence tap into the expertise and technical information of agencies. I apply this method to a corpus of 6.8 million floor and committee speeches given by US Congresspersons. I find support for the key prediction of the theory, although there is little evidence in support of the moderating role of agency independence. Ideological distance reduces bureaucratic influence and politicians are less likely to use the information produced by agencies ideologically apart. Bureaucracies can shape policies through information, and legislators do not oppose the influence of agencies when their policy preferences are aligned with that of politicians. Ideology remains an important driver of the political space that bureaucracies can create for themselves in the policy-making process.

Mission-Oriented Measurement: New Measurement for New Theory

The leitmotiv of this dissertation is that the scientific study of the bureaucracy can vastly benefit from the recent advancements in computational methods. With this dissertation, I make two key contributions to the literature on bureaucratic politics. First, I show how natural language processing methods can be used to advance our understanding of administrative processes and attributes. I combine large and automatic data collection with novel measurement strategies to study key concepts in the literature on bureaucracy. I introduce two new strategies to measure bureaucratic reputation and bureaucratic influence on legislative politics using respectively word embeddings and syntactic analysis. I also produce estimates of bureaucracies' partisanship from millions of individual bureaucrats' donations to partisan candidates. Second, I show how these methods can be applied to test new and old theories about political-administrative interactions. In Chapter 1 I test Carpenter's model of reputation-driven autonomy and in Chapter 3 I extend and test theories of strategic communication between politicians and bureaucracy. Moreover, in Chapter 2 I present and test a new theory on partisan selective accountability. The measurement and the theoretical novelty of this dissertation can ultimately contribute to our understanding of power relationship between elected and unelected bodies in democratic systems of government.

Chapter 1

A Dynamic Measure of Bureaucratic Reputation: New Data for New Theory

Abstract

Bureaucratic reputation is one of the most important concepts used to understand the behaviour of administrative agencies and their interactions with multiple audiences. Despite a rich theoretical literature discussing reputation, we do not have a comparable measure across agencies, between countries, and over time. I present a new strategy to measure bureaucratic reputation from legislative speeches with word-embedding techniques. I introduce an original dataset on the reputation of 465 bureaucratic bodies over a period of forty years, and across two countries, the US and the UK. I perform several validation tests and present two applications of this method to investigate (1) whether reputation leads to autonomy; and (2) whether partisanship and agency politicisation matter for reputation. I find that agencies with a better reputation report to operate under higher levels of autonomy. I also find that agencies enjoy a better reputation among the members of the party in government, with partisan differences less pronounced for independent bodies. I finally discuss how this measurement strategy can contribute to classical and new questions about political-administrative interactions.

1.1 Introduction

The political science literature has made important strides to enhance our understanding of the drivers of bureaucratic behaviour and the sources of bureaucratic power. Ever since the first models of political control of the bureaucracy, scholars have mostly focused on structural features of the bureaucracy, namely formal discretion, administrative procedures, and oversight mechanisms, considered to be the main tools to control bureaucratic policy-making (Epstein and O'Halloran 1999; Huber and Shipan 2002; McCubbins, Noll, and Weingast 1987; Moe 1990). Yet later advancements in the scholarship have gradually crumbled this structuralist approach to the study of the bureaucracy, and interpreted bureaucratic behaviour as unfolding through formal and informal channels or, in other words, implicit and explicit contracts (Carpenter and Krause 2015). One of the most prominent attempts to theorise what happens beyond formal contracts is Carpenter's reputation-based account of bureaucratic autonomy (Carpenter 2001a, 2010).

It is thanks to bureaucratic reputation – "a set of symbolic beliefs about an organisation embedded in a network of multiple audiences" – that agencies become autonomous actors and manage to secure the policies they favour despite the opposition of even the most powerful politicians (Carpenter 2001a, 3–4). During the US progressive era, well-esteemed bureaucracies such as the US Post Office Department and the Department of Agriculture were indeed consistently able to induce Congress and the President to consider and pass legislation that was quite different from their original preferences. This and subsequent works on bureaucratic reputation inaugurated a new tradition of scholarship that integrates formal (e.g., structure, capacity, and procedures) and informal accounts to better understand the interactions between the bureaucracy and other political actors. However, measuring reputation is a daunting task and scholarly work has been limited to a one-country, one-agency, one-year approach, failing to match the innovative scope of bureaucratic reputation as a new prominent account of bureaucratic politics.

For bureaucratic reputation to become a general theory of bureaucratic behaviour and to talk to other subfields in political science, we need empirical work that is able to identify the effects and causes of reputation more systematically. If we want reputation to talk to theories of delegation (Thomson and Torenvlied 2011), interest groups (Nelson and Yackee 2012), political oversight (Lowande 2018), and rule-making (Potter 2017), we need measures that allow theories of reputation to be jointly tested with and against alternative explanations of administrative outcomes.

In this paper I build on recent advancements in natural language processing and propose a new method to measure bureaucratic reputation from millions of speeches given by politicians in parliament. I employ word-embedding techniques to understand how politicians talk about bureaucratic agencies and derive word vector representations for each agency in every year and then measure the distance of these vectors from a vector that captures positivity. Agencies with better reputations will be "closer" to this positivity embedding than agencies with worse reputation. I introduce an original dataset on bureaucratic reputation for 465 agencies across two countries – the US and the UK – and over almost forty years. I use both quantitative and qualitative information to demonstrate the validity of these measures, showing that the estimates react meaningfully to important changes or scandals that involved the agency, and that they positively correlate with related measures such as public opinion about government agencies.

I present two applications of this method. First, I perform the first systematic test of Carpenter's reputation-based account of bureaucratic autonomy. I discuss a two-faceted concept of bureaucratic autonomy, one centred on political constraints and one on goal attainment. While I am not able to test whether agencies enjoying a better reputation are more likely to change policy, I find a positive relationship between reputation and self-reported level of autonomy bureaucrats have when performing their job. Second, I examine politicians' polarisation when they talk about bureaucracy. By splitting the initial corpus of speeches by political party, I estimate partisan measures of reputation and show how reputation differs by political party, party in government, and agency politicisation. I show that agencies enjoy a better reputation among the members of the party in government, but the difference between reputation among majority and opposition party members is smaller for independent agencies and non-departmental bodies. Finally, I discuss some limitations of the proposed measurement strategy and outline directions for future research.

This measurement strategy opens new paths to the study of political-administrative interactions, offering new data for both classical questions on control and delegation, as well as new questions that integrate theories of bureaucratic politics with insights from other subfields of political science such as partisan identity and political polarisation.

1.2 The Need for a New Measure of Reputation

Bureaucratic reputation has informed an innovative literature that bridges the principal-agent paradigm with informal political-administrative dynamics. Krause and Douglas (2005), for instance, show that presidential, congressional, and independent regulatory commissions are concerned with homogenous reputational considerations about the quality of their decisions that outweigh the different political pressures resulting from their degree of insulation from other political actors. Maor (2007) describes agency independence and its scientific gold standard as reputation protection mechanisms, able to legitimise bureaucratic decisions, and Krause and Corder (2007) find that bureaucracies under tight political control make less accurate (and more optimistic) economic forecasts because they discount future reputation costs associated with their mistakes at a steeper rate than independent organisations.

Works on bureaucratic accountability have employed reputation-based accounts too. Busuioc and Lodge (2016), for instance, conceive accountability as the practice of sustaining and cultivating reputation across multiple audiences, far beyond formal requirements aiming to reduce informational asymmetries and agency slack. Empirical support for this theoretical claim is offered by Gilad, Maor, and Bloom (2013), who showcase how reputation explains the communication strategies of agencies facing media attacks: silence in domains where the agency is well esteemed, and attention and responsiveness where reputation is weak. Reputation ultimately allows agencies to "generate public support, to achieve delegated autonomy and discretion from politicians, to protect the agency from political attack, and to recruit and retain valued employees" (Carpenter 2002, 491).

However, while bureaucratic reputation has been put forward as a new currency of bureaucratic politics, able to explain autonomous policy-making, delegation, and accountability, the discipline still lacks a comparable measure of reputation across agencies, between countries, and over time. While the number of empirical works on reputation has increased in the last few years, scholars still employ either qualitative data (Busuioc 2016; Gilad and Yogev 2012) or quantitative proxies such as the valence of media coverage (Maor and Sulitzeanu-Kenan 2013). When measured quantitatively, reputation has mostly been studied from the supply-side, namely by looking at how agencies respond to reputational threats (Maor, Gilad, and Bloom 2013) or try to manage their reputation through external communication. Lee and Whitford (2013), for instance, measure reputation as the number of freedom-of-information-act request denials and the time to respond to a request, Anastasopoulos and Whitford (2019) look at twitter profiles of agencies, and Busuioc and Rimkutė (2019) perform a quantitative content analysis of agency annual reports.

However, reputation is a set of beliefs among audiences and it therefore seems appropriate to measure bureaucratic reputation from the demand-side, trying to understand what these audiences' beliefs look like rather than what the agency does to change them. A recent measurement strategy that addresses this issue and makes use of survey instruments has been devised by Wood, Overman, and Busuioc (2020), who produce a detailed and multidimensional estimate of agency reputation based on a systematic sampling of key stakeholders. Yet the results are still limited to one agency – the EU Chemicals Agency – at one specific time, and in a single political system – the European Union. Furthermore, surveys about technical issues such as bureaucratic agencies tend to induce answers (Zaller 1992). When individuals are asked what they think about the Pensions Regulator, they might not even know that it exists.

Clearly, there is a trade-off between nuance and multidimensionality, on the one hand, and time-, country-, and agencies-coverage, on the other hand. In this paper I address this gap by proposing a new quantitative measure of reputation that trades nuance with coverage, while at the same time focusing on a key audience, politicians, that represent a good synthesis of what a diverse set of audiences perceive the bureaucracy to be. The main assumption invoked is that politicians' electoral incentives to align with key stakeholders when engaging with bureaucratic agencies make them a good - though partial - source of information to capture agencies' reputation, a source that includes politicians' voice too – one that has been surprisingly neglected.

1.3 The Whereabouts of Reputation: Audiences and Beliefs

The standard definition of bureaucratic reputation in political science is Carpenter's (2010, 45), for which reputation is "a set of symbolic beliefs about the unique or separable capacities, roles, and obligations of an organization, where these beliefs are embedded in audience networks." The key words of this definition are *beliefs* and *audience networks*. Every attempt at measuring reputation will therefore have to deal with two questions: what are the beliefs that convey information about reputation?, and what are the audiences whose beliefs give shape to bureaucratic reputation?

The first question is trivial. Beliefs are the value judgements or perceptions about various traits of an organisation. The literature generally clusters beliefs into four main facets or reputation, performative, moral, procedural, and technical (Carpenter 2010), but I follow the general conceptualisation of empirical work and interpret reputation as an aggregate unidimensional measure of the perceptions of multiple audiences about an agency that spans between a positive and negative extreme.

The question "what audiences should we care about?" conversely, is a hard one. Ideally, measures of reputation would start from an accurate mapping of the various constituencies that qualify as audiences, and would then measure how positive or negative each audience's beliefs about the agency are. This exercise – the one generally used in surveys – besides being very costly, is also inherently arbitrary, for a decision to consider farmers' associations an audience of the Department of Agriculture would itself end up being a stand-alone empirical question. This is why scholars have incorporated the identification of the agency's key audiences in the research question itself, using historical analysis, elite interviews, or secondary sources to map an agency's audiences, their beliefs, and various policy outcomes (Busuioc 2016; Carpenter 2001a, 2010; Gilad and Yogev 2012; Maor and Wæraas 2015). An alternative solution to the audience identification problem advanced in the literature is content analysis of the news. Yet newspapers may not be the most appropriate venue to look for perceptions and beliefs about administrative bodies. In addition, media coverage is mediated by editorial concerns and likely to suffer from selection bias, for agencies are more likely to end up on the news when problems, scandals, or clear inefficiencies afflict their related sector.[1]

The challenge is then to devise a measurement strategy that retains the attention to the multiplicity of audiences of qualitative works, that continues the attempt of time-varying measurements of early quantitative works, and that also allows for cross-country, -policy, and -agency comparisons. We need an alternative venue where the perceptions of multiple audiences are voiced and can be condensed into a unidimensional measure without deciding which audiences are part of the sample. I propose that legislatures are close to this ideal venue. Legislatures are the right place to capture what audiences – as mediated by their representatives – think of an agency while at the same time letting politicians decide which

[1]Parliamentary debates too can suffer from selection bias, but the fact that politicians have to regularly discuss a wide range of policy issues results in more uniform coverage.

these audiences are.

Reputation exists at audience-level but only key organised interests qualify as legitimate audiences. Politicians have an interest – because of genuine policy/ideological motivation, strategic re-election interest, or both – to represent these groups and voice their beliefs about the agency. What politicians say during legislative debates can then be a good source of information to measure reputation. It seems plausible to think of politicians as the messengers of external audiences such as business actors, trade unions, consumers associations, or non-for-profit organisations. Free-market parties can report the complaints of businesses towards the alleged over-regulation of the Environmental Protection Agency, whereas social-democratic parties can lament the loose regulation of the Financial Services Authority.[2] What politicians say is then a good proxy not only of what the audiences say, but what politically salient audiences say. It is rational for politicians to align or to report what salient groups think about an organisation. Legislative speeches might even be seen as a device to sort audiences' perceptions by political salience. Irrelevant audiences will be less likely to be dedicated attention by politicians who seek the support of the electorate.

Clearly, legislators are not neutral messengers able to precisely map all the relevant audiences of an agency. They might give voice disproportionately to some audiences or mischaracterise some audiences' beliefs. However, the diversity of their professional and personal background, as well as the interests and geographical areas they represent make legislators a rich and inclusive source of information about an agency's reputation, which minimises researchers' discretion and ensures multiple audiences are voiced.

Parliamentary speeches are frequently used as data input for empirical constructs. Beyond ideological scaling, speeches have been used to estimate the political agenda (Quinn et al. 2010), political influence of MPs (Blumenau 2019), and speeches' complexity (Spirling 2016), to name but a few. Tapping into what politicians say can thus be used in bureaucratic politics too to understand how politicians and elites talk about agencies and therefore estimate bureaucratic attributes such as reputation.

[2] Barberá et al. (2019), for instance, find that US legislators are more likely to follow, than to lead, discussion of public issues.

1.4 Agencies as Word Embeddings

Recent advancements in machine learning and natural language processing allow researchers to devise new measurement tools to study the bureaucracy (on delegation, see e.g. Anastasopoulos and Bertelli 2019; Vannoni, Ash, and Morelli 2021). In particular, a new set of techniques called "word embeddings" – first developed in computational linguistics to learn about semantics (Pennington, Socher, and Manning 2014) – offer new frontiers to measure bureaucratic attributes from text data. The core idea at the basis of word embeddings is that we can "know a word by the company it keeps" (Firth 1957, 11), and we can therefore derive its meaning from the context in which the word is used.

Word embeddings are technically the coefficients from neural network models that predict the occurrence of a word by the surrounding words in a textual sequence. A word of interest is represented as a dense, real-valued vector of numbers, whose length is informative about the complexity of the multidimensional space in which the word is embedded, and whose elements convey information about the semantic meaning of the word, with distances between such vectors capturing how similar the words are (Spirling and Rodriguez 2019). For instance, if the distance between the vector representation of the words "market" and "inequality" is smaller for social-democratic parties than for conservative parties, we can learn the views of market economy of different party families. Similarly, by looking at the vector representation of words that are most similar to the vector representation of word "women," we can examine how individuals, groups, or parties think about the role of women in society. The key innovation of word embeddings is that the meaning of the words is something that is learned from the text and is not exogenously given as in other text analysis approaches that look at word frequency.

Word embeddings have recently entered published work in political science. Preotiuc-Pietro et al. (2017), for instance, use word embeddings to estimate ideology based on tweets and to identify politically moderate and neutral users, and Rheault and Cochrane (2019) fit models of word embeddings augmented with political metadata to estimate the ideology of parties and politicians. The flexibility of word embeddings has been also used to address some limitations of sentiment analysis (Rice and Zorn 2019), and to study how ethnic and gender stereotypes change over time (Garg et al. 2018). The novel measurement strategy I propose builds on these recent trends and represents the first attempt to use large parliamentary corpora and word embeddings in the study of bureaucratic politics.

1.4.1 Countries, Speeches, and Agencies

The textual corpus from which I estimate 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.[3] I decided to focus on the US and the UK because while they both have highly competent bureaucracies in charge of the implementation and administration of policies, the UK civil service is deemed to be neutral and merit-based, while US agencies are on average more politicised (Hood 1991; D. E. Lewis 2008). This makes the study of reputation in these two cases informative about the relationship between bureaucratic traditions and reputation. As for more practical reasons, both the UK and US speeches are easily accessible and in the same language, making the estimation procedures less complicated. I set the time-frame from 1980 because the meaning of the words I use in the estimation has not changed since then, and it is therefore possible to compare estimates over time. Finally, a time coverage of about 40 years is a good balance between allowing reputation to change over time while ensuring an accurate mapping of all the agencies and their multiple denominations.

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.[4]

1.4.2 GloVe

To estimate word embeddings, 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

[3] I exclude the speeches from the House of Lords because the House of Lords has different functions from the House of Commons, and Lords are unelected, therefore the assumption for which politicians have an incentive to represent key constituencies does not hold.

[4] 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.

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

Table 1.1: 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 dog and policy cancels out, so that large values correlate well with words associated with FED, and small values correlate well with words associated with EPA.

vocabulary consisting of all the unique tokens that appear in the corpus. 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. For example, if word $i = FED$, word $j = policy$, and the window size is symmetric and equal to six, X_{ij} is the number of times FED (target word) co-occurs within six words to the left and right of the word *policy* (context word). Let X_j be the sum of the co-occurrences of any word i with the context word $j = policy$ (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.

While the technical aspects of Glove are complicated, the main idea is not. The intuition 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. To do so, we compare the probabilities of these two words happening with various probe words k. We might expect word $k = independence$ to be related to the word FED more than to the word EPA, and word $k = critic$ to be related more to the word EPA than to the word FED. Similarly, we expect the word $k = dog$ to be related to neither, and the word $k = policy$ to be related to both. Table 1.1 represents these expectations in terms of hypothetical probabilities. The probability ratio 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 approximates 1, because they do not help discriminate between which word is related to which. This is why, compared to the raw probabilities, co-occurrence ratios are better able to encode relevant semantic relations and to understand which words are related to the words FED and EPA.

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). The innovation of Glove compared to other algorithms is that the model is trained on all the non-zero entries of the matrix rather than on the entire sparse matrix or on individual context windows (Pennington, Socher, and Manning 2014). 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,

$$J = \sum_{i=1}^{V} \sum_{j=1}^{V} f(X_{ij})(w_i^T w_j + b_i + b_j - log(X_{ij}))^2 \tag{1.1}$$

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. For instance, in the 2018 corpus of UK speeches, the five most similar word vectors to the *Home Office* embedding – one of the most mentioned agencies – are *immigration*, *department*, *official*, *windrush*, *minister*, *ask*. From just these six words, we can learn that *immigration* was a key issue for the 2018 post-Brexit Home Office, and that MPs frequently *asked* the *minister* about the *Windrush* scandal.

1.4.3 Application

I train the GloVe algorithm on a local corpus of parliamentary speeches for every year and every country. I follow standard practice in text-analysis and I lemmatise the tokens, remove punctuations, digits, capitalisation, stopwords, and all tokens with two characters or fewer to increase the precision of the estimation. 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 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 $w_i \in [0, 1)$.[5]

I then 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 reputation of 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:

$$\begin{aligned} \vec{positivity} = \vec{successful} + \vec{effective} + \vec{great} + \vec{excellent} \\ - \vec{poor} - \vec{negative} - \vec{terrible} - \vec{bad} \end{aligned} \tag{1.2}$$

where the arrows signify the words are vectors. The selection of words followed four criteria: (1) the meaning of the words should be uncontroversial (i.e., positive or negative), (2) stable over time, (3) similar across countries, (4) present in every local corpus of speeches in any given year and country. The precise words I used are similar to the seed words chosen by Rice and Zorn (2019) to set the benchmark for positivity and negativity dictionaries. In the Appendix (Section A.3) I show that the reputation estimates produced with alternative positivity vectors are highly correlated with the estimates derived from this vector. I finally measure the cosine similarity between the word embeddings of each agency and the $\vec{positivity}$ vector. The reputation score will thus be the angular distance between the two embeddings. Formally,

$$\theta_j = \theta_{(\vec{a},\vec{p})} = \frac{\vec{a} \times \vec{p}}{||\vec{a}|| \times ||\vec{p}||} = \frac{\sum_1^n a_i \times p_i}{\sqrt{\sum_1^n a_i^2} \times \sqrt{\sum_1^n p_i^2}}$$

where $\theta_{(\vec{a},\vec{p})}$ is the cosine similarity between the agency vector $\vec{a}$ and the positivity vector $\vec{p}$, namely the ratio between the sum of the products of the i^{th} elements of the two vectors (the nominator) and the product of the square root of the vectors to the power of two (the denominator). For instance, if the embedding $\vec{FED}$ is semantically very similar to $\vec{p}$, it will have a very high reputation, whereas if the $\vec{EPA}$ embedding is semantically distant, it will have a lower reputation. The resulting metric is normalised to take up values

[5] 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). Estimation implemented with the *text2vec* R package.

between 0 and 1, where greater values signify better reputation.

1.4.4 Uncertainty

Deriving uncertainty measures from neural network models is an area of research still under development (Rheault and Cochrane 2019). Since I cannot estimate uncertainty based on the variance of θ, I produce upper and lower bounds for every agency-year estimate based on the number of mentions of the agency in any given year. The reputation of agencies mentioned 2,000 times per year will be less uncertain compared with that of agencies that barely happen to be mentioned. I therefore model uncertainty as a reciprocal exponential function of the number of mentions, so that agencies with fewer mentions are penalised but in a non-linear fashion. Mathematically, the upper and lower bounds will be given by $\theta_j \pm |1 - \exp \frac{1}{m_j}|$, where m_j is number of mentions of agency j.[6]

This is clearly a mathematical artefact that nonetheless allows me to estimate uncertainty based on the reasonable assumptions for which the more politicians talk about an agency, the more we can learn about its reputation. For instance, since the vocabulary V consists of words appearing at least five times in the corpus, the least mentioned agency will be mentioned at least 5 times, and it will have upper and lower bounds estimates equal to $\theta_j \pm |1 - \exp \frac{1}{5}| = \theta_j \pm 0.221$. Conversely, for agencies mentioned 1,000 times, upper and lower bounds will be equal to $\theta_j \pm 0.001$.

1.5 Results

The final dataset consists of reputation estimates for 465 bureaucratic bodies – 217 in the UK and 248 in the US – and over a period of 39 years.[7] Table 1.2 reports some descriptive statistics for the dataset as a whole and by country. Full lists of agencies are reported in Section A.1 of the SI. Agencies are mentioned on average 179 times per year and have a reputation of 0.50, with a standard deviation of 0.15. The average reputation in the two countries is about the same, 0.48 in the UK and 0.52 in the US. The Environmental Protection Agency and the Department of Homeland Security are the agencies with the highest average number of mentions in the US (1,514 and 1,582 times per year, respectively),

[6] Given $\frac{1}{m_j} \in (0,1]$ for any $m_j > 0$, and $\exp(x) > 1$ for any $x \in (0,1]$, then $\exp(\frac{1}{m_j}) > 1$ and $\lim_{\frac{1}{m_j} \to 0} |1 - exp(\frac{1}{m_j})| \approx 0$ (when m_j is large) and $\lim_{\frac{1}{m_j} \to 1} |1 - exp(\frac{1}{m_j})| = |1 - e| \approx 1.7$ (when $m_j \to 0$).

[7] 465 and not 636 as the initial sample because not all the agencies included in the initial lists are mentioned in the speeches.

	Total		UK		US	
Total Agencies	**465**		217		248	
Observations	**7,067**		2,272		4,791	
Time Coverage (Years)	**40**		40		36	
Variable	Mean	SD	Mean	SD	Mean	SD
Reputation	0.50	0.15	0.48	0.16	0.52	0.15
Mentions	179	344	143	302	196	362

Table 1.2: Descriptive statistics for all the agencies in the dataset and split by country. The table reports the number of agencies for which reputation estimates are produced, the number of observations and time coverage, and the average reputation both for the total dataset and by country.

whereas the Treasury and the Home Office are the most mentioned agencies in the UK (1,829 and 944 times per year on average, respectively).

Figure 1.1 and Figure 1.2 show the reputation of 8 of the most mentioned bureaucratic bodies with the largest year coverage in both countries. At first glance, the comparatively high reputation of the military in the US is remarkable, with the Air Force, the Navy, and the Department of Defense being among the public bodies with the highest reputation over the entire time-frame considered. Though significantly fluctuating, the reputation of most of these eight agencies seem to be stationary, as suggested by the loess function in the plot. There are nonetheless important exceptions. After the peak in the early 1980s, the reputation of the EPA, for instance, experienced a gradual but constant decrease. Similarly, the reputation of the Department of Homeland Security drops rapidly after its establishment in 2002.

As for the UK, it is possible to see how the reputation of Ministry of Defence drops in 1991, the year of the so-called Options for Change – the dramatic manpower cut to the British Armed Forces after the end of the Cold War – and how the reputation of Network Rail increased in the immediate years after its establishment in 2002 for then rapidly decreasing from 2012 to 2018, possibly as a result of the uninterrupted criticism about delays and service disruptions for commuters.

Figure 1.3 shows the reputation of the central banks of the two countries, the Bank of England and the Federal Reserve, with dashed vertical lines representing some critical junctures. The FED is overall mentioned more often than the Bank of England, the latter nonetheless enjoying on average a better reputation throughout all the almost 40 years covered by the data (0.52 for the Bank of England and 0.45 for the FED). There is a jump

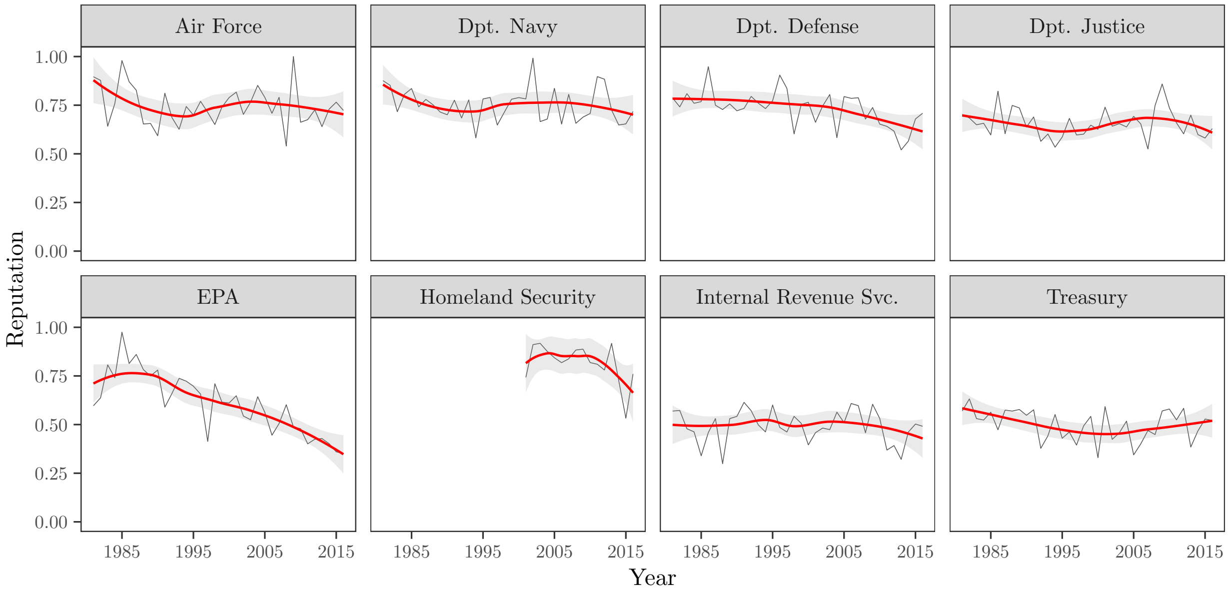

Figure 1.1: Reputation estimates of US agencies over time with loess approximations superimposed. Estimates are cosine similarity between the "agency" embedding and the positivity vector in every year.

in the reputation of the Bank of England when it became independent, in 1997, while the reputation drops with the 2008 and 2011 financial crises. A similar pattern is followed by the FED. For instance, its reputation falls with financial crises (the one in the early 90s, the sub-prime crisis, as well as the Asian Crisis of 1997), while it increases with the new competences delegated by the Dodd-Frank Act in 2010.

1.6 Validation

In this section, I address the validity of the measurement with three tests. I assess face and predictive validity with qualitative information about six different agencies to test whether the estimates follow what we would expect the reputation of an agency to be after some critical events. The second test draws from standard convergent validity tests and looks at the relationship between reputation and public opinion (Collier and Adcock 2001). The third test assesses the criterion validity of the measure and compares the reputation estimates derived from legislative speeches with alternative estimates derived from a large corpus of newspaper articles.

When in 2005 Hurricane Katrina hit the state of Louisiana, local and national leaders blamed the poor response of FEMA (D. E. Lewis 2008). Yet FEMA's mismanagement was just the second act a of longer play that started in 1992 with Hurricane Andrew, which is "best remembered as an epic bungle by the Federal Emergency Management Agency"

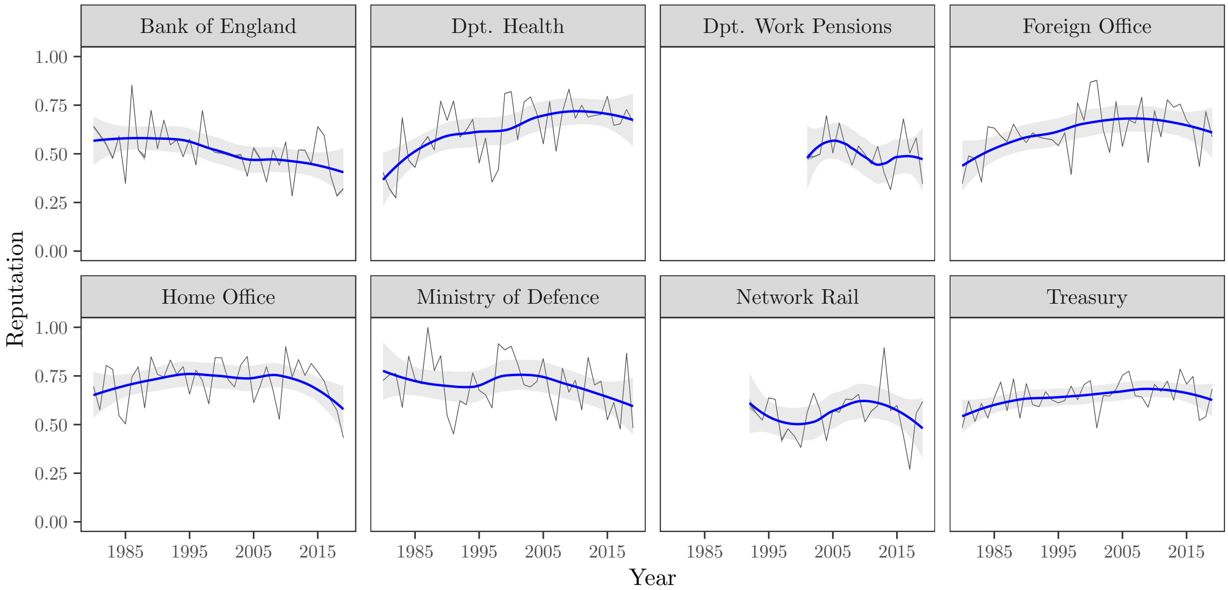

Figure 1.2: Reputation estimates UK agencies over time with loess approximations superimposed. Estimates are cosine similarity between the "agency" embedding and the positivity vector in every year. Data for Network Rail before its establishment in 2002 are from its predecessor Railtrack, established in 1994.

(Timeline 2017). Yet in the period between the two hurricanes, the agency had a rapid renaissance. As Roberts (2006, 56) notes, after its reorganisation in 1992, FEMA *"morphed from a caricature of the ills of bureaucracy into a model of effective federal administration. Politicians who previously blamed the agency for its slow and inefficient response to disasters came to depend on the agency to lend credibility to their own efforts."*.

The Department of Homeland Security, after its establishment following the 09/11 terrorist attacks, has been highly criticised over excessive fraud and lack of transparency. Multiple scandals eroded the reputation of the agency. In 2005, the new personnel system called "MaxHR" was blocked in court for violating collective-bargaining employees' rights (The Washington Post 2008a); in 2008, a Congressional report denounced 15 billion dollars worth of failed contracts (The Washington Post 2008b); and in 2015 the department was found to be operating top secret databases infringing the most basic security procedures (Office of Inspector General (DHS) 2015).

First in 2004, with the dramatic increase in the backlog of pending disability claims, and then in 2014, with the falsified waiting lists, the reputation of the Department of Veterans Affairs has been subject to harsh criticism too (CNN 2014).

Figure 1.4 plots the reputation of these three agencies over time together with the scandals and critical junctures described above. The reputation estimates capture the

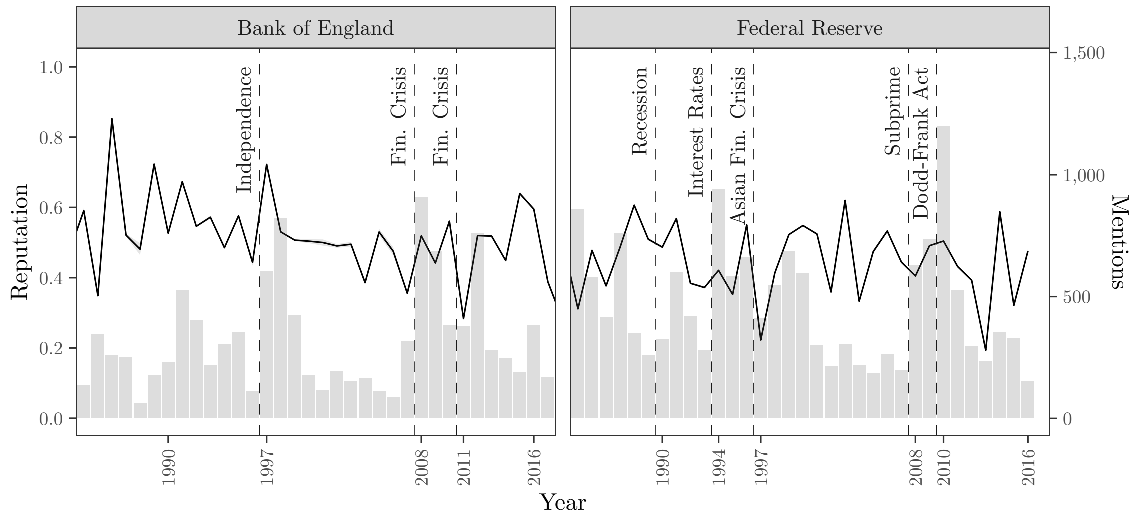

Figure 1.3: Reputation estimates and number of mentions of central banks over time. Vertical dashed lines represent critical junctures.

decrease in reputation of FEMA after 1992 and 2005, as well as the marked increase after its reorganization in 1992. The reputation of the Department of Homeland Security is consistent with the high criticism, with significant drops and an overall decreasing trend up until 2015, and the reputation of the Department of Veterans Affairs drops in 2004-5 and from 2014 follows a decreasing trend.

The reputation of UK agencies too has not been immune to scandals. The Financial Services Authority (FSA) was highly blamed for the 2008 financial crisis, which brought to its abolition later in 2012. The so-called "light-touch" regulatory approach of the authority received cross-partisan criticism (The Daily Telegraph 2008) and was also called into question by the independent review chaired by Lord Turner, which criticised the authority's philosophy for which "markets are in general self-correcting" (FSA 2009, 87). Similarly, the agency's reputation also suffered from the judicial defeat in the *Durant v FSA* case in 2003 – a leading decision with respect to data protection.

Another severe and more recent scandal in British politics involved the Home Office in what has been called the Windrush Scandal. 83 instances were reported in which people were wrongly detained, denied legal rights, and mistakenly deported from the UK (The Times 2018). Although for very different reasons, the reputation of the department was damaged a few years before too, when Jacqui Smith resigned as Home Secretary in 2009 as a consequence of the scandal that involved her husband (The Guardian 2009).

Figure 1.4 follows quite accurately these events, with the reputation of the FSA reaching one of its lowest levels just after 2003, and dropping again during the 2008 financial

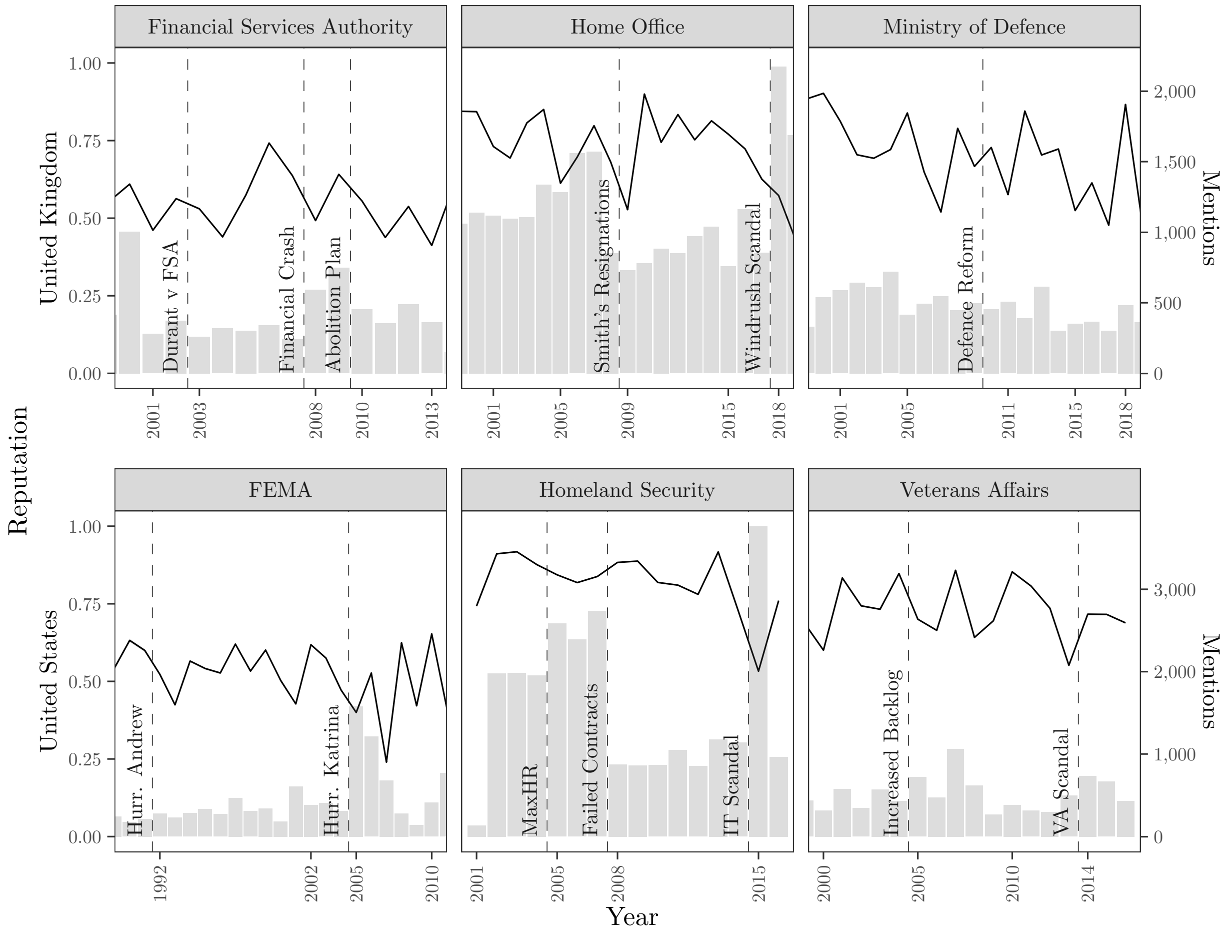

Figure 1.4: Reputation estimates and number of mentions of UK and US agencies and departments. Vertical dashed lines represent scandals or critical junctures.

crisis and after the abolition announcement of the government in 2010. Similarly, the reputation of the Home Office drops in 2009 and 2018 following the scandals. Finally, the graph also shows a rapid decrease in the Ministry of Defence's reputation from 2009 to 2011, the years of the Strategic Defence and Security Review (Ministry of Defence 2010) – highly criticised by Parliament (House of Commons Defence Committee 2011) – and the Defence Reform Report (Ministry of Defence 2011), which assessed the causes of the department's under-performance and proposed changes to prevent it "from getting into such a poor financial position in the future" (Ministry of Defence 2011, 13).

For the second test, I compare the reputation estimates with survey data on public attitudes towards US federal agencies. I assembled a panel of public opinion data from the Pew Research Centre reports (Pew Research Center 2019) and Gallup surveys[8] and matched reputation with public opinion data for 18 agencies over several years, for a total

[8]Data accessed at the following link: news.gallup.com/poll/27286/government.aspx on 10 February 2020.

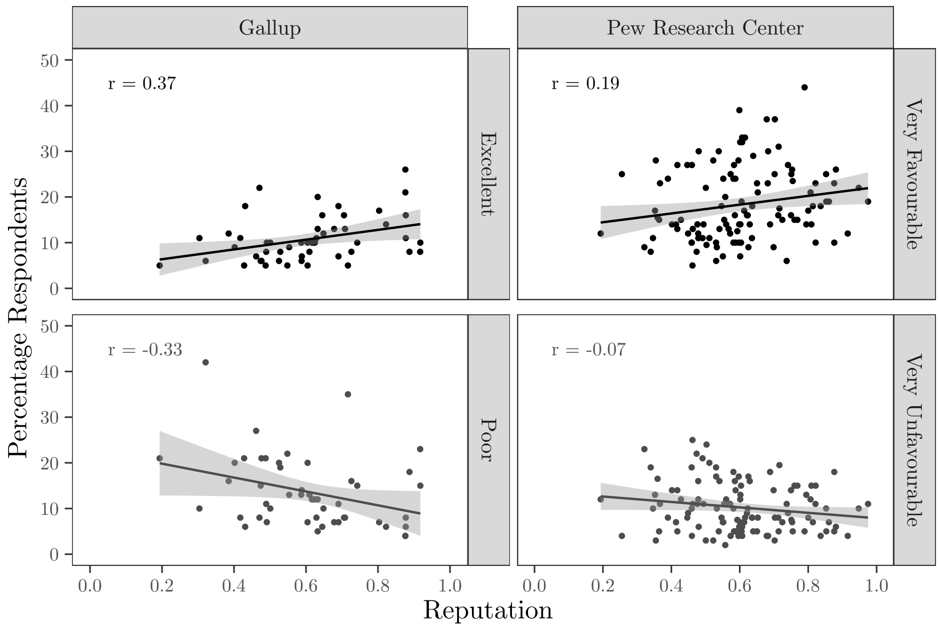

Figure 1.5: Reputation estimates of US federal agencies (x-axis) and percentage of respondents with an opinon about the agency as reported in the panel labels (y-axis). Correlation coefficients reported in each panel. Vertical lines are linear approximation with robust confidence intervals. Panels on the left-hand side use public opinion data from Gallup, whereas panels on the right-hand side use public opinion data from Pew Research Center.

of 51 observations from Gallup and 114 from Pew Research Center. Responses to the survey were recorded on a 4-level scale, from very positive to very negative. Figure 1.5 plots the relationship between the reputation estimates (x-axis) and the percentage of respondents who reported a very positive or very negative opinion about the bureaucracy (y-axis). Overall there is a positive correlation between reputation and positive opinion about the agency and a negative correlation – though weak – between reputation and negative opinions.

The last validity test consists of comparing the reputation estimates with other estimates produced from an alternative corpus. Especially for more salient agencies – and therefore mentioned more often in the news – the estimates derived from speeches can be compare against estimates derived from newspaper articles. I estimated agency reputation from a corpus of more than 1.2 million articles published between 2014 and 2019 in 12 main UK national newspapers. Among the agencies most mentioned in the news – those with the lowest level of uncertainty – there is a positive and high correlation between the reputation estimates derived from the two corpora, thus strengthening the confidence in the criterion validity of the estimates.[9]

[9] In the Appendix (Section A.5) I report more information on the corpus of articles and the correlation

In the next section, I present an empirical application, showing how this measurement strategy can open new research agendas in the study of political-administrative interactions.

1.7 Application 1: Reputation and Autonomy

In *The Forging of Bureaucratic Autonomy*, Carpenter (2001a) identifies reputation as the key driver to bureaucratic autonomy, namely the ability of bureaucratic agencies to resist political control and "change the agenda and preferences of politicians and the organized public" (15). However, reputation-based accounts of bureaucratic autonomy have been pushed back by the lack of a general and dynamic measure of reputation that would allow scholars to test this relationship more systematically on multiple agencies and over time. The measure I introduced in this chapter offers a first attempt at solving the stalemate.

The concept of autonomy that Carpenter presents is the result of a reputation-building enterprise which results into the ability of the agency to freely choose which policy to prioritise, steering the policy process and building coalitions which ultimately enable them to influence politicians. While this definition goes beyond that of early works on bureaucratic discretion (Huber and Shipan 2002; Epstein and O'Halloran 1999) – which looked at the statutory provisions defining the formal authority of the agency – it implies that the agency is not captured by political principals and that it is able to operate *de facto* independently. It is therefore possible to break down Carpenter's conceptualisation of bureaucratic autonomy into two consecutive parts: one that entails autonomy as the process of operating free from political constraints; one that focuses on goal attainment, on policy change as a result of autonomy. While the former centres on the ability of the agency to forgo political influence (i.e., process), the latter includes the substantive changes in policy (i.e., outcome). Quantitative tests of Carpenter's theory are very challenging if the researcher embraces both the process- and outcome-based concept of autonomy. Measuring the extend to which agencies sway the wishes of elected politicians or cause policy to change would be a stand-alone piece of research for each agency and each period of time separately. Yet measures of the process-based concept of autonomy exist and, by combining them with my measure of bureaucratic reputation, it is therefore possible to provide a first - though partial - systematic test of Carpenter's theory.

between estimates.

1.7.1 Carpenter's Model of Bureaucratic Autonomy

Carpenter's theory is arguably the first attempt to move beyond a conception of autonomy that is consequential to expertise and information alone. Contrary to what posited by principal-agent models of autonomy and delegation, it is not the knowledge of the bureaucrats which earned her autonomy, it is rather the capacity of mezzo-level managers to practice a politics of legitimacy and cultivate a reputation among multiple networks which – in equilibrium - put Congress and Presidents in a position where they have to accept the agency's will regardless of whether they like it or not. The path to autonomy therefore starts with bureaus which manage to build "political legitimacy, or strong organisational reputation embedded in an independent power base" (Carpenter 2001b, 14). Building coalitions and cultivating a good reputation therefore leads to autonomy which, entails policy-innovation and decisions that are not overturned by the overseers. The process-based definition of autonomy outlined above is a step in-between building a unique reputation and political influence.

From this theoretical account, I derive the following testable hypothesis.

HYPOTHESIS: Agencies enjoying a better reputation are more likely to report higher levels of perceived autonomy when performing their job.

1.7.2 Empirical Test

I propose a first quantitative test of Carpenter's model by combining my measure of reputation with Bertelli et al. (2013) dataset on the autonomy of 71 US federal agencies from 1998 to 2010. Autonomy is measured from bureaucrats' attitudes as captured by several waves of surveys commissioned by the Office of Personnel Management (OPM) and the Merit Systems Protection Board (MSPB). In particular, the survey items used to produce the autonomy estimates aim to measure the flexibility given to bureaucrats in performing their job, whether they feel encouraged to come up with new and better ways of doing things, or the extent to which they are satisfied with their involvement in decisions that affect their work. Through Bayesian item response theory models, the authors derive agency-level estimates of the posterior means for each agency and over several years.

After combining the two datasets, I obtain reputation and autonomy estimates for

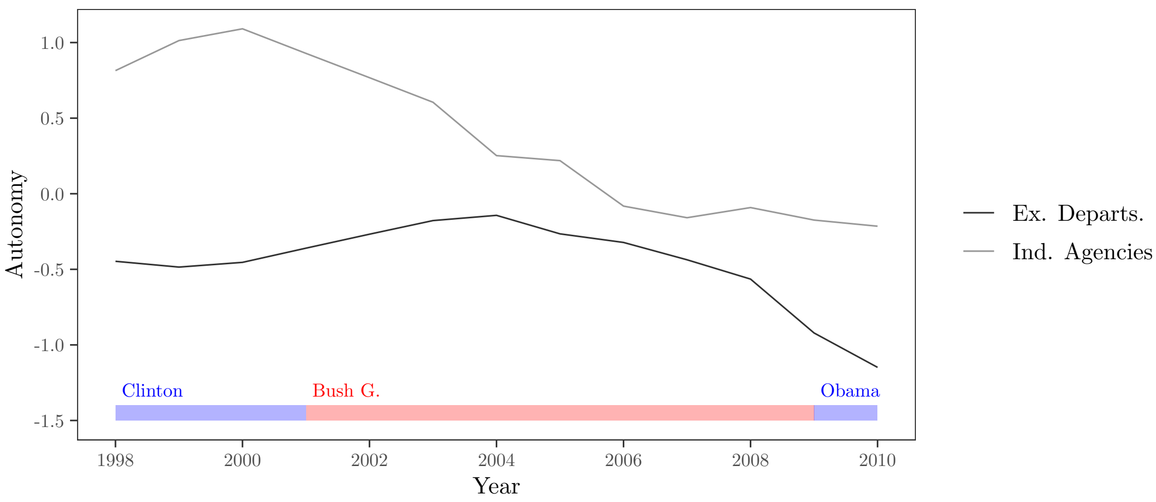

Figure 1.6: Trend of agencies' autonomy across presidencies. Pooled data from 1998 to 2010.

56 agencies for a time-frame that spans over 13 years, for a total of 428 observations. Interestingly, although decreasing, the trend is u- and bell-shaped for executive departments and independent agencies, respectively (see Figure 1.6). While both type of agencies are less autonomous in 2010 compared with 1998, executive departments had higher levels of autonomy during the Bush presidency, whereas independent agencies enjoy higher levels of autonomy during democratic presidencies.

To formally test whether agencies with a better reputation are on average more autonomous, I exploit the panel data structure of the dataset and estimate weighted least square distributed lag models, where lagged values of the treatment estimate dynamic causal effects. Unit and time fixed effects account for unobserved agency-specific characteristics and common shocks, respectively. To account for the varying degrees of uncertainty of the autonomy estimates, I construct regression weights based on the uncertainty of each observation's autonomy estimate. Using weights will thus pull the regression towards matching the data with the lowest levels of uncertainty. In particular, the larger the uncertainty (i.e, the difference between the upper and lower bounds of the estimate), the smaller the weight. Let $\boldsymbol{\theta}$ be a vector of uncertainty of the autonomy estimates of length equal to the number of observations, where each element is the variance of the estimate derived from its 95% confidence interval. The weight assigned to each observation will be the inverse of the variance ($1/\sigma_Y$). I estimate dynamic effects with the following distributed lag model of reputation, where

$$\text{Autonomy}_{it} = \alpha_i + \delta_t + \sum_{\phi} \beta_\phi \text{Reputation}_{i,\phi} + \epsilon_{it} \tag{1.3}$$

where *Reputation* is the agency reputation, α_i and δ_t are agency and year fixed-effects, and ϵ_{it} is the stochastic component. In particular, the vector of coefficients β_ϕ estimates dynamic treatment effects. Because changes in perceived autonomy might take time to realise, I estimate up to four-year later effects and report alternative specifications in the Appendix (see Section A.6). Setting the number of lags equal to four is both a practical and theoretical choice. Given the short time coverage of the dataset (about 13 years on average) it would be very costly in terms of statistical power to estimate longer term effects, and it seems reasonable to expect perceived autonomy to slowly adjust with increasing/decreasing reputation. It is therefore possible to measure how fast a change in reputation affects autonomy. If β_ϕ is distinguishable form 0, it means that the effect of reputation still persists or only realises ϕ years later.

Figure 1.7 shows the results. The coefficients of the lead value of autonomy ($t = -1$) is a falsification test, for it estimates the effect before treatment actually occurs. Overall, there is support for Carpenter's model of bureaucratic autonomy. There is overall a positive association between reputation and autonomy. There is evidence of a large effects of reputation on autonomy in time $t = 0, 1$, with a one-unit increase in reputation accounting for more than one unit increase in autonomy (1.3-1.4). However, because of the strong assumptions on which these models rest, the results must be interpreted with caution. In fact, the main causal identification assumptions of the two-way fixed effects estimator (i.e., no time-varying omitted confounders, as well as no effect of autonomy on future values of reputation (Imai and Kim 2019)) are quite heroic in this context, for many things could change over time and affect simultaneously reputation and autonomy.[10]

1.8 Application 2: Bureaucratic Polarisation

Polarisation is a key characteristic of contemporary politics (Hacker and Pierson 2006). However, we do not know whether polarisation is an appropriate lens through which studying the bureaucracy. If reputation is ultimately a set of beliefs among audiences, does partisanship contribute to the formation of these beliefs?

One major advantage of measuring reputation with text-analysis methods is that, by meaningfully splitting the initial corpus, it is possible to break down the estimates

[10] In the Appendix I report results with SE clustered by agency and for specifications estimating different number of lags (see Section A.6).

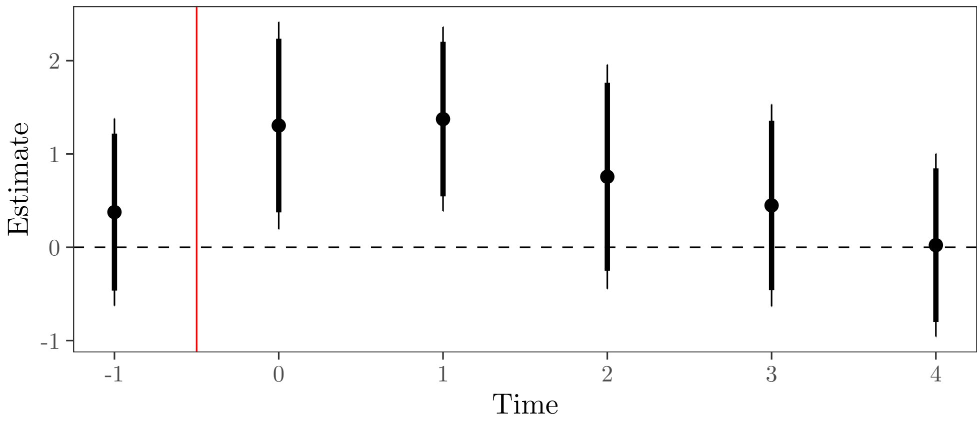

Figure 1.7: WLS regression estimates of the dynamic effect of reputation on autonomy. 90 and 95% confidence intervals estimated with heteroskedasticity robust SE.

by audiences. One way of doing it is to decompose the estimates along the partisan divide and measure the reputation agencies enjoy among different political parties (and arguably different audiences). Do agencies enjoy a different reputation among liberal and conservative parties? If so, how does this difference change over time and across agency type? To answer these questions, I replicated the estimation and trained the models on two different corpora, one for each of the two main parties in each country. The absolute value of the difference between the two partisan measures of reputation can thus represent a measure of "bureaucratic polarisation."

Figure 1.8 plots the average polarisation with respect to government/executive departments and non-departmental bodies over time. Surprisingly, polarisation is on average higher in the UK than in the US. Polarisation about government departments follows a slightly increasing trend in the US, whereas the trend is decreasing for independent agencies/non-departmental bodies, particularly in the 80s-90s. The trend for the UK is less clear. Polarisation about government departments increases during the Thatcher governments of the 80s and spikes again in the early 2000s with the Blair governments, while non-departmental bodies follow a decreasing trend.

The varying level of politicisation of bureaucratic agencies may let us expect that their reputation depends on who leads the executive. Figure 1.9 shows the average reputation by party and by party in government (the label of the panels). In both countries, "partisan" reputation is higher when the party is in power compared to when it is at the opposition, with changes being particularly marked for liberal parties (+.04 and +.02 for the Labour and Democratic Party, respectively).

The last layer to this snapshot of partisan measures of reputation is agency structure.

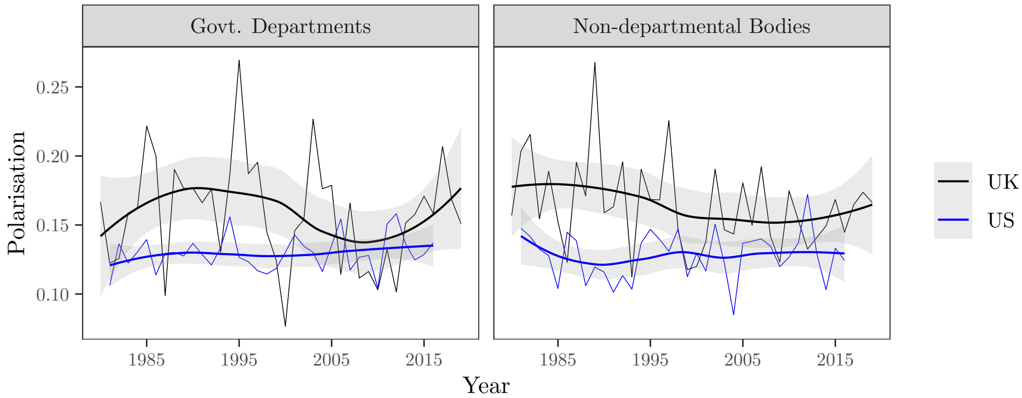

Figure 1.8: Average polarisation with respect to government departments and non-departmental bodies in the US and the UK. Each data point is the absolute difference of the average reputation among the two main political parties for all agencies in every year. For the US, government departments are executive departments, executive agencies, and agencies within the Executive Office of the President. Non-departmental bodies are independent agencies. Non-for-profit public organisations and government-owned corporations are excluded from the figure.

Partisan differences might be more pronounced if the agency is under tight government control compared to more independent agencies. As a result, high levels of independence – and therefore low levels of politicisation – might be associated with lower bureaucratic polarisation.

To test this proposition more rigorously, I match the estimates of partisan reputation with the dataset on the structural independence of US federal agencies assembled by Selin (2015), which captures agency independence along two dimensions: independence as the ability of an agency to make policy decisions without political interference, and independence as statutory limitations to appointment/removal and qualification requirements placed on agency officials with key decision-making roles. The indicators are derived by modelling 50 structural features about the agencies with a Bayesian latent variable model. They range between 0 and 4, with higher values signifying higher independence.

As shown in Table 1.3, independence is negatively associated with bureaucratic polarisation, although only with respect to the requirements placed on the officials who manage the agency. Models from (1) to (4) report OLS estimates, whereas Models from (5) to (8) report the results of WLS regressions, with weights equal to 1 over the average of uncertainty estimates across the two parties (see Section *Uncertainty*), so that most mentioned agencies are assigned larger weights. I also include agency type fixed effects, therefore accounting for the differences between departments, executive agencies, independent agencies, non-for-profit public organisations, and agencies within the Executive Office of the President. Far

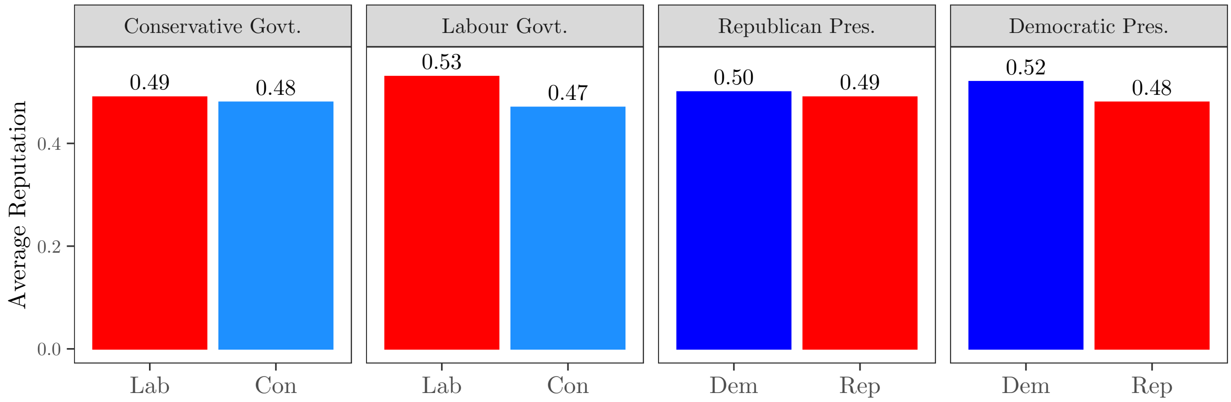

Figure 1.9: Average reputation by party (x-axis) and party in government (panel labels) for UK and US agencies. Y-axis reports the average reputation for every agency and across the entire period covered by the dataset.

from causal interpretation, when pulling the regression towards matching the data with the lowest levels of uncertainty (WLS estimates), a one unit increase in independence as requirements imposed on agency officials is associated with a decrease in polarisation by .04, which is equal to a decrease by 29% with respect to the average polarisation across the sample (.13). The distance between bureaucratic reputation among Republicans and Democrats is smaller when appointment/removal limitations and qualification requirements exist on agency officials in key decision-making position.[11]

1.9 Discussion and Limitations

Despite the significant advantages of a dynamic measure of reputation, it is worth emphasising some limitations of the proposed measurement strategy.

First, like every quantitative measure of agency attributes (e.g., discretion, autonomy, or politicisation), the proposed measure of reputation is a simplified picture of a conceptually rich and multifaceted attribute. Although word embeddings encode rich semantic features of terms, they are simply not able to match the deep observation of qualitative work. The construct validity of the measure could be enhanced by focusing on a smaller sample of agencies and making additional theoretically informed decisions about the estimation procedure. For instance, researchers could limit the textual corpus to a sub-set of speeches given by certain legislators or committee members, or about a pre-defined set of topics.

[11] In the Appendix (Section A.6) I show the estimates are robust to using heteroskedasticity-consistent standard errors, an alternative coding of the type of agencies, and to limiting the dataset to 2014, the year when the data collection in Selin (2015) ended.

DV:	Polarisation about Bureaucracy							
	OLS				WLS			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Independence: Decision-Makers	-0.01		-0.01	-0.02	-0.01		-0.04*	-0.04*
	(0.01)		(0.02)	(0.02)	(0.01)		(0.02)	(0.02)
Independence: Political Review		-0.00	0.01	0.01		0.01	0.02	0.01
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Independent Agency (dummy)			-0.01				0.03	
			(0.03)				(0.02)	
Fixed-effects								
Agency Type				✓				✓
Observations	102	102	102	102	102	102	102	102
R^2	0.01	0.00	0.01	0.03	0.00	0.01	0.06	0.08
Within R^2				0.01				0.05

IID SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 1.3: OLS and WLS estimates, standard errors in parenthesis. DV is distance between reputation estimates measured from speeches given by Republican and Democratic legislators. Independence data collected from agency statutes in 2013-2014, therefore the dependent variable is average reputation in 2013-2014. Agency type coded in the same way as listed on the institutional website of the US government, usa.gov/branches-of-government.

Second, and relatedly, modelling the total population of politicians' speeches might increase measurement error, for not everything said about bureaucracy should contribute to the reputation estimates. Let us consider speeches praising war veterans while at the same time mentioning the Department of Veterans Affairs. These speeches, arguably very positive, could inflate the reputation of the department even though they convey little information about its reputation. In the Appendix (Section A.4) I show that the reputation of US military agencies decreases when speeches mentioning both the name of military agencies and military values (e.g., integrity, honour, courage, etc.) are excluded from the corpus. However, while removing speeches mentioning military agencies and military values might "de-bias" the estimates, we could in fact be cancelling out an important dimension of their reputation, namely the fact that they operate in a salient and respected political domain. Decisions about the textual corpus should therefore be driven by strong theoretical reasons. Other sources of heterogeneity at the agency level that cannot be handled by sampling speeches can be addressed empirically too. For instance, looking at changes in reputation over time can account for the different (and constant) probability of agencies being mentioned alongside more positive concepts.

Third, politicians' speeches – while giving voice to agencies' audiences – might also include an undesired partisan component. This can be problematic for the reputation estimates of more politicised agencies, which may be subject to more partisan debates. As shown in the application, there seem to be partisan differences in the way politicians talk about bureaucracy. Yet the partisan divide is just one out of the many ways we could split the speeches. Other differences could be detected based on the legislative roles, education, or previous political experience of politicians. While the partisan divide is the most intuitive way to group legislative speeches, future research could study other cleavages too, and use them to explain how audiences form their beliefs about bureaucracy as well as audiences' behaviour *vis-à-vis* administrative agencies.

Similarly, while parliamentary speeches are a rich source of information about agencies' reputation, they might not fully or precisely capture key agencies' audiences. Neglecting some audiences or mischaracterising their beliefs would bias the measure of reputation, which would result in reputation estimates diverging the agency's "true" reputation. While this might be a source of concern for researchers willing to use this measure, it can also represent an opportunity to investigate how reputation is portrayed across different venues. As I show in the Appendix (Sectio A.5), for highly salient agencies, newspaper articles might be a good alternative and complementary source of information about agencies' reputation.

Finally, despite the new opportunities for quantitative analysis offered by this dynamic measure of reputation, comparative research should be guided by robust theory, especially when defining the sample of agencies to study. As argued by Carpenter (2020), comparing administrative bodies across policy domains is often likely to lead to flawed conclusions. This is particularly relevant for the study of reputation, which underscores the importance of agency's reputation for uniqueness, and hence focuses on agencies' reputation *within* their own field.[12] Depending on the research question at hand, looking at within-agency variation might provide a solution to the "comparative trap" if the features of the policy domain remain constant over time, but researchers should be careful when reaching conclusions based on comparative analyses that pool many agencies with different tasks, missions, and organisational features.

[12] In the Appendix (Section A.2) I give two examples of within-domain comparisons, looking at the reputation of US agencies overseeing financial institutions and UK independent regulators of network industries.

1.10 Conclusions

Structural accounts of bureaucratic politics have long claimed that formal decisions about the agency are so incisive that they allow politicians to define the balance between agency discretion and political control. Yet on the other hand, it is thanks to reputation that agencies accrue their ability to "sway the wishes of elected officials on particular matters of policy and to secure deference from these elected officials" (Carpenter and Krause 2012, 30). Bureaucratic reputation has been used to study important bureaucratic outcomes, from strategic communication, to autonomy, and accountability, yet empirical work has not been able to match the rich theoretical and conceptual ground of reputational theories of bureaucracy, partly because we lack a dynamic measure of reputation.

I first claimed that bureaucratic reputation can be aptly captured from what politicians say, for they have electoral incentives to align with key agencies' constituencies. By employing word-embedding techniques, I estimated the reputation of 465 bureaucratic bodies from more than 4.5 million speeches over a period of almost 40 years in two major western democracies, the US and the UK. I performed multiple validation tests and presented two applications of this method to the study of bureaucratic autonomy and bureaucratic polarisation. First, I provide a systematic analysis of Carpenter's reputation-driven account of bureaucratic autonomy. By combining estimates of agency autonomy with my dataset, I show that, on average, agencies enjoying a better reputation work in a more autonomous manner. While this test is limited to a process-based definition of autonomy, it yields empirical support to the positive relationship between reputation and bureaucratic autonomy. Second, I split the corpus of speeches by the two main political parties in both countries and show through visual representations and correlations how these measures can capture partisan differences in politicians' beliefs about the bureaucracy, opening new research agendas that combine key questions in bureaucratic politics, such as delegation of authority and political control, with theories of political polarisation.

This paper makes two main contributions. One methodological, showing how recent advances in natural language processing can be employed to study bureaucratic attributes and political-administrative interactions; one substantive, offering an original and validated measure of reputation able to capture variation across agencies and over time. Overall, this measure will be able to advance our understanding of key questions in administrative politics, from the causes and effects of bureaucratic reputation, to more nuanced questions about delegation of authority, political control, and bureaucratic accountability, while simultaneously opening new uncharted research agendas that bridge bureaucratic politics with other subfields of political science.

Chapter 2

Is Partisanship Bad for Bureaucratic Accountability?

Abstract

Bureaucratic accountability rests on legislators' ability to objectively evaluate the performance of bureaucracies. Yet partisanship can trigger selective accountability, whereby government legislators selectively evaluate and acquire information about bureaucracy in order to protect the image of their party. I test this argument with two studies. First, I analyse the sentiment of partisan statements about 336 agencies over 40 years in the US and the UK, estimating word-embedding models from millions of legislative speeches. I find that statements are on average 3 percentage points more positive when the legislator's party is in power. A difference-in-differences design further shows that co-partisan legislators react more positively to scandals affecting bureaucracies. Second, I find that government legislators are less likely to acquire information from bureaucracies in congressional hearings and make less frequent use of quantitative evidence in their speeches about bureaucracies. Therefore partisanship may hinder bureaucratic accountability.

2.1 Introduction

A normative tenet of democratic government is that bureaucracies are accountable to elected politicians. To that end, political 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). Accountability, however, does not come without costs, and politicians might lack the ability or willingness to acquire the information necessary to oversee bureaucracies. In this paper I show that partisanship triggers *selective accountability* and hinders legislators' ability to objectively evaluate and acquire information on bureaucracy.

The thrust of principal-agent models applied to the interactions between elected politicians and bureaucracies is that principals' choices *vis-à-vis* a bureaucratic agent are always the product of a trade-off. Consider for instance the trade-off faced by principals when deciding between credibly delegating discretion to the agent or ensuring responsiveness to the principal's directives (Epstein and O'Halloran 1999), or between the appointment of neutral agents who will provide expertise or political allies who will push policy in the direction favoured by the principal (D. E. Lewis 2008). Accountability, commonly understood as a process of information-acquisition and evaluation, is no exception and principals' decision to hold bureaucracy to account comes with benefits and costs. On the one hand, accountability is necessary if principals want to redress the actions of a drifting agent and citizens favourably evaluate principals who take politically costly oversight initiatives (S. M. Miller and Ruder 2020); on the other hand, oversight and information acquisition is a resource-intensive activity and principals have to prioritise some agencies over others (McCubbins, Noll, and Weingast 1987). A big portion of the costs inherent to principals' decision affecting the bureaucracy take the form of political constraints: legislators' deciding to cut budgets, alter the organisation of bureaucracies, or withdraw delegated authority are all choices that require political coordination and agenda-setting power (Moe 1984). One political constraint which is understudied in the literature is principals' partisanship.

To preview the argument, consider the literature on partisanship and blame/credit attribution, which frequently finds that voters selectively attribute blame and credit to the government based on the party in power (for a review of the argument, see Ashworth and Bueno De Mesquita 2014). Faced with the same negative information about the performance of a Republican president, Republican voters are less likely to attribute blame to the president compared to Democratic voters. According to some accounts, the partisan-induced bias in blame and credit attribution poses a threat to electoral accountability, for it inhibits voters' ability to sanction and reward politicians based on factual information (Healy and Malhotra 2013; Little, Schnakenberg, and Turner 2021). If we allow partisan selectivity to factor into legislators' decisions to hold bureaucracy to account, it becomes clear how partisanship has the potential to hinder bureaucratic accountability, for legislators will evaluate and oversee agencies insofar as it does not undermine the image of their party.

This argument is general and rests on two uncontroversial assumptions. First, legislators care about their electoral consensus, and the image of the party is important for legislators' support among voters. Second, because the government is responsible for administering public policies through bureaucratic bodies, a bureaucracy doing a poor job has negative implications for the electoral approval of the government and its party (James and John 2007).[1] When their party is in power, there is no space in legislators' welfare for criticism towards bureaucracy, for uncovering negative information about bureaucracy would undermine the image of the government. If evaluations of bureaucracy ought to be positive irrespective of bureaucratic performance, government legislators have few incentives to oversee bureaucracy and bear the costs of such a resource-intensive exercise. Partisanship hence triggers selective accountability. When co-partisan with the government, legislators have a lighter touch when holding bureaucracy to account: both evaluations and oversight of bureaucracy receive a partisan discount.

I test this argument with two novel studies: one on the effect of partisanship on legislators' selective evaluation of bureaucracy, and one on the effect of partisanship on legislators' selective information-acquisition.

In the first study, I introduce new data on partisan statements about 336 bureaucracies in the US and the UK estimating word embedding models from million of legislative speeches. I recover the semantic meaning legislators attach to the bureaucratic bodies they mention in their speech by estimating word vector representations for every bureaucracy 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

[1] I use the term "government" to refer to the body holding executive power, the presidency in the US or the prime minister in the UK.

about the sentiment of statements about bureaucracy for each party, agency, and year. I find that statements are approximately 3 percentage points more positive when there is partisan alignment between the party and the government. The effects of government co-partisanship are robust to including measures of agency partisan and ideological alignment with the party. However, since selective evaluation along partisan lines could could be confounded by other factors, I increase the credibility of the test with a difference-in-differences design where I compare how US congresspersons reacted to three major scandals involving three bureaucratic bodies and find that, faced with the same exogenous shocks about the reputation of agencies, co-partisans with the government are between 13-19 percentage points more likely to give a positive statement about the agencies involved in the scandal.

In the second study, I test the implications of co-partisanship with the government for the bedrock of bureaucratic accountability: legislators' willingness to acquire information from bureaucracy and use it to hold bureaucracy to account (Lupia and McCubbins 1994; Gailmard and Patty 2013). Models of congressional oversight assert that congressional committees "possess sufficient rewards and sanctions to create an incentive system for agencies" (Weingast and Moran 1983, 768). In fact, committees are the venues where accountability is most energetic. In particular, it is in committees where bureaucracies are asked to report on their performance. Using original data on the identity of witnesses heard before Senate Congressional Committees and the partisan composition of committees from the 106th to the 116th Congress, I find that when there is partisan alignment between the committee chair and the president the probability of a bureaucracy appearing before the committee as a witness decreases by -21 percentage points. Finally, I show that co-partisans are also less likely to use analytical language and quantitative evidence when arguing about bureaucracy in legislatures, suggesting that partisanship decreases legislators' incentives to focus on what bureaucracies actually do. I find that the frequency of statistical facts in legislators' speeches about bureaucracy decreases with co-partisanship with the government (by about 4% compared to the average value for the US and by 9% compared to the average for the UK).

I summarise the studies, findings, and implications in Table 2.1. While accountability deficits have been documented in the literature (Gailmard 2009; Schillemans 2011; Schillemans and Busuioc 2015b), this is the first attempt at unveiling a "partisan" obstacle

Study	Design	Finding	Implication
Study 1: Evaluation	Panel Data	Co-partisans give more positive statements about bureaucracy.	Partisan selectivity in evaluating bureaucracy.
	Difference-in-Differences	Co-partisans react more positively to negative information about bureaucracy.	
Study 2: Information	Panel Data	Bureaucracies are less likely to appear before committees when there is a co-partisan chair.	Partisan selectivity in acquiring information on bureaucracy.
	Two-Way Fixed Effects	Co-partisans use statistical facts less frequently when debating about bureaucracy.	

Table 2.1: Design, findings, and contibutions of the two studies.

to bureaucratic accountability. I find large support for several empirical implications of the argument. Importantly the fact that I find similar effects of partisanship on selective evaluation and information-acquisition in two countries with very different administrative traditions (politicised in the US and neutral civil service in the UK) is convincing evidence of the importance of partisanship for bureaucratic accountability.

2.2 Selective Accountability

There is a vast literature in political science characterising 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 party-team (Iyengar and Westwood 2015; Kahan et al. 2017). Evidence for this form of partisan selectivity has been found in many countries and levels of government. For instance, recent work on partisan evaluation of former president Trump's management of the COVID-19 pandemic shows that as Democrats increasingly blamed Trump for the pandemic, Republicans assigned him little responsibility (Graham and Singh 2021). Outside the US, Bisgaard (2015) finds that in the UK, despite both Labour and Conservative supporters acknowledging the worsening of the economy, voters attributed blame in a highly partisan fashion. Labour supporters were hesitant to condemn the then-Labour government, whereas Conservatives had no doubt about the government's responsibility for the economic catastrophe following the 2008 financial crisis. These partisan differences, I argue, emerge also when legislators hold bureaucracy to account.

In democratic government, bureaucracy administers and implements policy, but responsibility for positive or negative outcomes rests with the elected government. Despite varying level of autonomy, bureaucratic bodies respond to the political will of the executive, and an under-performing bureaucracy has detrimental consequences for the consensus of the incumbent party. James and John (2007) and Boyne et al. (2009), for instance, show that the publication of low performance information about local public services in UK local authorities decreases the incumbent's aggregate vote share at the election following publication. Similarly, Malhotra and Kuo (2008) study voters' responses to Hurricane Katrina, showing that both Republicans and Democrats attributed most blame for the loss of life and property damage in New Orleans to political leaders – namely President Bush and Mayor Nagin – rather than to Federal Emergency Management Agency Director Michael Brown. These ideas connect bureaucratic performance to the broader scholarship on retrospective voting, which shows that the incumbent party is rewarded for good economic performance and sanctioned for bad economic performance (Ferejohn 1986; Erikson 1989), and that this occurs across all levels of government (De Benedictis-Kessner and Warshaw 2020). The performance of the bureaucracy can be interpreted as a narrower dimension of economic performance, which nonetheless triggers similar responses in voters' support for the incumbent.

Government legislators – who care about their electoral consensus and their party's – are not happy about the reputation of the party being sullied by a bureaucracy doing a poor job. Faced with the potential threat of under-performing bureaucracies, government legislators can choose between two alternative strategies. They can tighten oversight in the attempt to prevent bureaucratic failure, or they can give up oversight and – irrespective of performance – portray bureaucracy under a positive light. These two strategies come with different payoffs. Preemptively increasing bureaucratic oversight is costly and can backfire if legislators unveil poor-performing bureaucracies. Coupled with criticisms from the opposition, it would resemble a self-declaration of failure. Furthermore, uncovering the poor job of bureaucracy would come to uncertain benefits, which would be conditional on successfully remedying poor performance. Conversely, lightening up oversight and disregarding negative information about bureaucracy has no immediate electoral costs, and *a priori* appreciations of bureaucracy would contribute to sustaining the good image of the government. This logic should apply to good performance too. Even though negative information about bureaucracy has been found to have a larger effect on incumbent's electoral consensus compared to positive information (James and John 2007; James and Moseley 2014), legislators will not miss the opportunity to highlight the good performance of bureaucracy when their party is in power. Just as they sweep negative information under the carpet, they also amplify the policy successes of bureaucracy. The first observable implication of this argument is hence that, regardless of actual performance, government legislators evaluate bureaucracy more positively compared to when they are at the opposition.

The distortions created by partisanship have a second observable implication: when co-partisan with the government, legislators deliberately lighten monitoring of agencies. Let us recall that a necessary condition for accountability to be sustained is legislators ability to acquire information about bureaucracies (Busuioc 2009; Gailmard and Patty 2013). In fact, information acquisition was the main gist of early theories of bureaucratic accountability, ensured through constituents raising their voice (i.e., "fire-alarm" mechanism) or politicians' direct monitoring (i.e., "police patrol" mechanisms) (McCubbins and Schwartz 1984). Legislators were ultimately responsible for establishing procedures that would create incentives for bureaucracies to disclose information in order to prevent that delegation of authority to bureaucratic bodies led to abdication of power (Lupia and McCubbins 1994; Moe 2012). 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 (Schillemans and Busuioc 2015b; Busuioc and Lodge 2017). The underlying assumption to accountability being conceived as an information-acquisition process is that politicians care about agency characteristics and behaviour and do not want them to clash with their own preferences (G. J. Miller 2005). However, if partisanship detaches legislators' evaluation of bureaucracies from what bureaucracies actually do, then government legislators have little interest in acquiring and assessing factual information on the performance of agencies.

Moving from evaluation to information-acquisition, this account brings partisanship inside theories of bureaucratic accountability. The empirical predictions I will be testing are therefore two. First, co-partisanship with the government makes evaluations of bureaucracy more positive. Second, co-partisan legislators are less likely to acquire and use information on bureaucracy. Partisanship may ultimately hinder legislators' ability to hold unelected officials to account.

A divisive element of theories of partisan selectivity in accountability behaviour is the underlying mechanism. In fact, despite clear evidence for partisan selectivity, the literature is not unanimous on the interpretation of the nature of partisan responses, which could be sincere – and therefore affected by some form of cognitive bias (Bisgaard 2015, 2019) – expressive – reflecting the value of sustaining the good image of the party (Bullock et al. 2015 and references therein) – or rooted in performance beliefs – suggesting that supporters of the incumbent party believe their party performs better (Sirin and Villalobos 2011; Graham and Singh 2021). Even though recent scholarship identifies several challenges to observational studies claiming to tease out the mechanism (Fowler 2020; Little 2021), when we move the focus from voters to politicians, it seems more reasonable to consider legislators highly strategic actors who will try to protect the image of their party to the detriment of bureaucratic accountability. It is reasonable to expect legislators' selective evaluation and information acquisition to be driven by the electoral gains that would derive from the party, government, and bureaucracy enjoying a good reputation among the public. The account I present here therefore builds on work that interpret partisan selectivity as an expressive response, while nonetheless acknowledging that it is not possible to provide conclusive evidence in support of this mechanism. The consequences of both cognitive or strategic forces are nonetheless equivalent. Partisanship triggers selective accountability: stronger when at the opposition and weaker when in government.

2.3 Data

2.3.1 Statements about Bureaucracy

To test the selective-evaluation argument I analyse the sentiment of legislators' statements about bureaucracy. Legislators express evaluation of bureaucracies in multiple venues. To allow for large time, agency, and party coverage, I focus on legislative speeches. The measurement strategy builds on the one presented in Chapter 1 (and Bellodi 2022), which uses word embedding models to produce validated estimates of bureaucratic reputation in the US and the UK. Here I use a similar approach but I estimate word embeddings from separate corpora of speeches for the two main political parties in the US and the UK. I ultimately produce party-agency-year estimates for about 336 government departments and bureaucratic agencies (197 for the US and 139 for the UK) over a time frame of approximately 40 years.

The intuition behind word embeddings is that we can learn about the relationship between words by looking at the frequency with which words co-occur with one another. Suppose we want to compare the overall sentiment in statements given about the Federal Reserve in a year when the Federal Reserve is praised for its monetary decisions and in a year when it is blamed. To do so, we compare the probabilities of the word FED to co-occur within a selected window of words with some context words. We might expect the word FED to co-occur more often with the word *good* than with the word *bad* when it is praised, and more often with the word *bad* than with the word *good* when it is blamed. Word embeddings are vectors of numbers which encode these co-occurrence relationships between words. The word embedding of FED estimated from a corpus in which the FED is praised, will be "similar" to the embeddings of words capturing positivity (e.g., good, sound, excellent, effective), whereas it will be very different from it if estimated from a corpus when the FED is blamed for its policies. Therefore, the key innovation of word embeddings is that the meaning of the words is not a given, for it is learned from the text and represented as a dense, real-valued vector of numbers, whose length is informative about the complexity of the space in which the word is embedded, and whose elements convey information about the semantic meaning of the word, with distances between such vectors capturing how similar the words are (Rodriguez and Spirling 2022; Bellodi 2022).

Suppose we have a list of unique words V appearing in a corpus from which we build a matrix of dimensions $V \times V$. We then compute the co-occurrence probability for every pair of words. Word embeddings are ultimately the coefficients of statistical models which reduce the matrix and capture the relationship between the ratio of co-occurrence probabilities of each pair of words. Like the genetic information encoded in a strand of DNA, the elements of such vectors carry semantic information about the word. 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; Rodriguez and Spirling 2022). Therefore, by comparing the word embeddings 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 FED vector and a positivity vector. Similarly, if we do this separately for Democrats and Republicans, we can see how positive statements about the FED are for a different set of actors.

I estimate word embeddings on party-year corpora by pooling all the legislative speeches given by the two main parties between 1981-2016 for the US and 1980-2018 for the UK. 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. Once I have word embeddings for every word in the corpus, I exploit the arithmetic properties of vector representations of words and build a vector that combines some unambiguously positive and negative embeddings that will act as benchmark to measure the positivity of statements about bureaucracy. 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:[2]

$$\begin{aligned} \vec{positivity} &= \vec{great} + \vec{excellent} + \vec{successful} + \vec{effective} \\ &\quad - \vec{bad} - \vec{poor} - \vec{negative} - \vec{terrible} \end{aligned} \tag{2.1}$$

I finally measure the cosine similarity between the word embeddings of each agency and the $\vec{positivity}$ vector, producing estimates of positivity for party p, agency a, in year t. The resulting metric is normalised to take up values between 0 and 1, where greater values signify more positive statements.

While this measurement strategy does not allow me to estimate statements of each single legislator, it leaves me with sufficiently large corpora to estimate word embeddings for each year and each party separately, capturing variation across party and over time.[3] Importantly, a key advantage of this strategy is that it allows for the meaning of "positivity" to change over time and across parties, for $\vec{positivity}$ is estimated separately from every party-year corpus. Republicans and Democrats will have different concepts of positivity and will use different words to communicate positive and negative statements. By anchoring the estimates of positivity to time-changing positivity vectors, I improve the validity of the measure.

[2]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 and are the same words used in Chapter 1.

[3]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.

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

2.3.2 Acquisition and Use of Information

To test the argument about selective information acquisition I use two types of complementary data that capture both the acquisition and use of information on bureaucracy.

First, I web-scraped original data on the identity of the witnesses appearing in congressional Senate hearings and the partisan composition of committees through the govinfo.gov API.[5] Congressional committees are a primary accountability forum where elected officials can acquire information on bureaucracy through questions and interrogations, and where they can express criticism about the performance of bureaucracy. Committees are also the central institutions which is supposed to ensure congressional dominance over bureaucracy (Weingast and Moran 1983; Moe 1984). I collect data on witnesses, partisan composition of the committee members and the chair for the universe of Senate hearings from the 106th to the 116th Congress, for a total of 9,281 hearings. I then create a dichotomous variable equal to 1 if a member of bureaucratic body is heard as a witness and 0 otherwise. Witnesses are heard before committees in 62 of the total number of hearings. The three bureaucracies heard most often are the Department of Homeland Security, the Environmental Protection Agency, and the Department of Justice (215, 197, and 185 times, respectively).

If government legislators are less likely to acquire information on bureaucracies, they are also likely to use that information when arguing about bureaucracy in legislative debates. As a proxy of legislators' use of information about bureaucracy, I measure legislators' use of statistical facts and quantitative evidence when arguing about bureaucracy through a targeted dictionary-based analysis of legislative speeches, focusing on sections of text near the name of the bureaucracy. 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,"

[4]Full lists of agencies and the average sentiment across party are reported in Tables B.1 and B.2. Detailed information about source of speeches and the estimation of word embeddings is reported in the Online Appendix (see Sections B.1, 1.4.2).

[5]Because of data availability, I am able to collect this data only for the US.

"percentage," "twice," "total," as well as all numbers used to express quantities.[6]

I estimate the use of statistical facts for more than 500,000 speeches mentioning the name of a bureaucracy (196,689 for the UK and 288,756 for the US) given by a total of 3,833 unique legislators. Text pre-processing steps are reported in Section B.3.1 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. I then compute for every speech the absolute frequency of the words of the speech also appearing in the dictionary.[7]

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. To support this claim, in Section B.3.2 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 a different context (i.e., online medical blog posts). I find a positive and significant correlation between the number of sentences labelled as reporting facts and the estimates of the dictionary method, thus strengthening our confidence in the low context-dependence of the dictionary (see Table B.3).

2.3.3 Agency Partisanship & Ideology

To ensure that the positivity of statements resulting from co-partisanship with the government is not driven by other agency-level characteristics, I gather data on agency ideology and partisanship to account for a plausible source of omitted variable bias in the

[6] The full list of words is reported in Table B.4 in the Appendix.

[7] Results are robust to using the term-frequency inverse-document-frequency of facts-words, therefore reducing the importance of words that appear very often and in many speeches (see Table B.15 in the Appendix).

estimation that I present later. In fact, when there is a republican president, Republicans' statements about bureaucracy might be more positive because the are more bureaucrats appointed by the same party/president.

There are three main challenges to measuring latent agency attributes. 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 donations). 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 to republican or democratic beneficiaries by employees of one of the bureaucracies in my initial sample, for a total of approximately 11 million donations covering 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 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 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 points, 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.

[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.

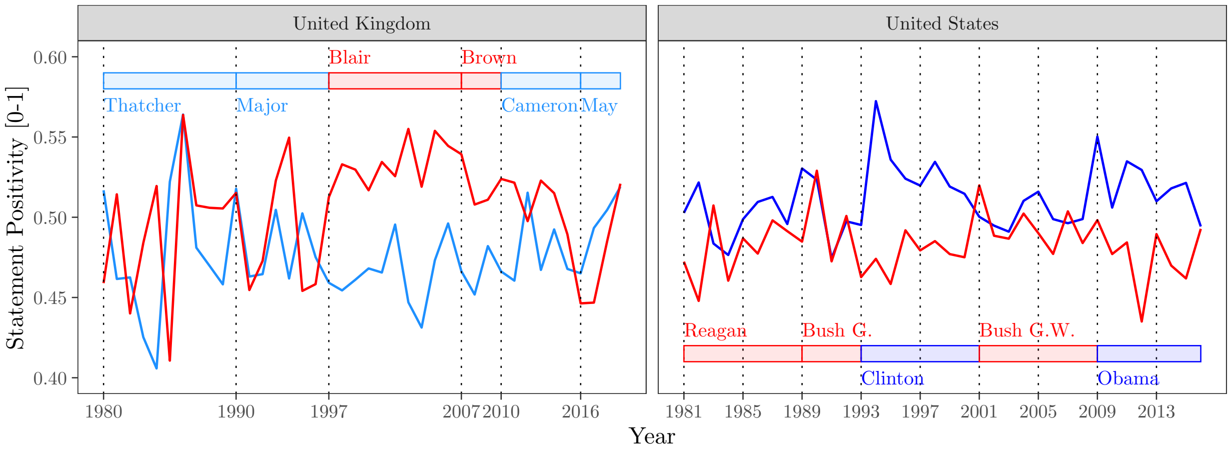

Figure 2.1: Positivity of parties' statements averaged across all agencies over time and across presidencies. Red for the Republican and Labour, blue for the Democratic party, light-blue for the Conservative party.

2.4 Study 1: Selective Evaluations

By just looking at Figure 2.1, it is clear how partisanship matters for statements about bureaucracy. The figure plots the positivity of statements 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 and the prime minister/president. On average, when there is a Democratic President, statements given by the Democratic party are more positive compared to when there is a Republican President, and *vice versa* for the Republican party's statements. The UK shows a similar trend. When the government changes colour, the positivity of statements changes too. Statements about bureaucracy for the Conservative party are more positive during the Cameron and May governments than during the Blair and Brown premierships. Labour party's statements too, despite being on average more positive than the Conservative party's, follow government cycles, more positive under Labour governments, more negative under Conservative governments.

In order to estimate the effect of partisanship on the positivity of statements about bureaucracy more rigorously, I exploit the panel structure of the data and estimate the following model for both the US and the UK separately:

$$\text{Positivity}_{p[a,t]} = \delta_p + \phi_a + \alpha_t + \beta\text{Party-Govt. Alignment}_{p[t]} + u_p \tag{2.2}$$

where $Positivity_{p[a,t]}$ represents the positivity of statements given by party p, about agency a, and in year t. $Party\text{-}Govt.\ Alignment_{p[t]}$ 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 fixed effects to account for 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]

Who is in charge of government matters for the what legislators say about bureaucracy. For both the US and the UK, party-government alignment increases the positivity of statements about bureaucratic bodies. In Table 2.2 I report the results for the UK. Sample size is large and estimates are highly precise even when including party-agency and year-agency fixed effects, and therefore accounting for differences in party attention to agencies and in time-changing agency characteristics. Because positivity ranges from 0 to 1, the effects can be interpreted as changes in percentage points. All else being equal, partisan alignment is associated with statements on average 3 percentage points more positive, as large as 0.21 times the standard deviation of positivity in the sample. Similarly, the effect is as large as the average difference in positivity between the Conservative and Labour parties (i.e., 0.03).

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 evaluations of the bureaucracy might be confounded by a change in agency ideological and partisan leaning. Republican presidents might appoint conservative bureaucrats and the sentiment of Democrats' statements about the agency might decrease for reasons unrelated to their opposition status. To rule out the risk of confounding posed by time-changing agency partisanship and ideology, I add to Model (2.2) 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.

In Table 2.3 I report the results.[10] The effect of government-party partisan alignment is positive, precisely estimated, and in the expected direction. Being aligned with the

[9]In Tables B.5 and B.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).

[10]The sample of agencies for which there is available data for statements' positivity, party-agency ideological distance and partisan alignment are 21 out of the 197 agencies for which I produce estimates of the positivity of statements. Results hold when limiting the analysis to the sample of agencies for which all the three variables are available (see Table B.7 in the Appendix).

DV:	Positivity [0,1]		
Country:	UK		
Model:	(1)	(2)	(3)
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			✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 2.2: Partisanship and Statements' Positivity, UK Data. OLS estimates. Units are party-agency-year observations.

president is associated with an increase in the positivity of statements 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)).[11]

Models (5) and (6) replicate the estimation from Model (2), but this time I use the measure of party-agency ideological distance and partisan congruence. If partisanship did not matter for the sentiment of statements about bureaucracy, we would expect party-agency ideological or partisan differences to be better predictors than party-government partisan alignment. However, selectivity is not triggered by party-agency characteristics 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 in the sample, and as big as the average difference in positivity between Democrats and

[11]In the Appendix (Table B.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)). Furthermore, 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 B.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.

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

DV:	Positivity [0,1]					
Country:	US					
	Party-Government				Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
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		✓	✓	✓	✓	✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 2.3: Partisanhsip, Ideology, and Statements' Positivity, US Data. OLS estimates. Units are party-agency-year observations.

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 effects in the US and UK are similar suggests that partisanship is able to affect evaluations of bureaucracy in very different administrative systems.

2.4.1 Scandals in the Federal Bureaucracy

These results could still be confounded by unobservable sources of heterogeneity. Because β is identified by comparing over-time changes in partisan statements 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, or with changes in agency characteristics that are not captured by agency ideology and partisan leaning. Here I strengthen causal identification by looking at how legislators react to identical information, namely scandals in the US federal bureaucracy. Absent selective evaluation, I should fail to detect a difference in how majority- and opposition-party legislators respond to scandals.

To ensure that legislators react to identical information, 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 of co-partisanship (either Democratic or Republican). Qualitative information on the scandals is reported in Section B.2 of the Appendix.

I estimate the effect of legislator-government co-partisanship on the positivity 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 1 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-valenced words. I then count the number of words in each sentence contained in 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

$$\begin{aligned}\text{Positive Statement}_{i[l,a,t]} = \eta_l + \phi_a + \alpha_m + \tau\text{Leg.-Govt. Alignment}_{l[t]} \\ + \gamma\text{Post-Scandal}_{i[t]} + \beta\text{Leg.-Govt. Alignment} \times \text{Post-Scandal} + u_i\end{aligned} \tag{2.3}$$

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 government legislators. η_l, ϕ_a, and α_m 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 to identify how co-partisanship shapes legislators' *subjective* reaction to *objective* information (i.e., a clear national-level scandal) without resorting to hypothetical scenarios.[13]

Table 2.4 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 selectively evaluate bureaucracies: more positive if aligned with the government, more negative if at the opposition.

This is a conservative test of the selective-evaluation argument, for it exposes legislators to clearly negative information about the bureaucracy. In fact, selective evaluation of bureaucracies might be more pronounced when the valence of the information leaves space for ambiguity.

2.5 Study 2: Selective Information Acquisition

In this second study I test whether partisanship triggers selective information-acquisition on bureaucracy. I present two tests, one on the acquisition and one on the use of information. First, I compare the probability of a bureaucracy appearing before a Senate committee as a witness when the committee chair is of the same party as the president and when the chair belongs to the opposition party. If co-partisan legislators acquire less information on bureaucracy, they should also be less likely to use the information when holding agencies to account in legislative debates. In the second test, I compare legislators' use of statistical facts in their speeches about bureaucracy when they are co-partisan with the government and at the opposition. The expectation is that partisan selectivity follows from the acquisition to the use of information: when co-partisan with the government,

[13]In Table B.10 in the Appendix, I report falsification tests with placebo post-treatment indicators (i.e., placebo scandal dates) in support of the parallel trend assumption.

DV:	Pr(Positive Statement = 1)			
Months before/after Scandal	4 months	3 months	2 months	1 month
Model:	(1)	(2)	(3)	(4)
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	✓	✓	✓	✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 2.4: 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.

legislators are more likely to disregard what bureaucracy is actually doing, and use an argumentative style less grounded in statistical facts.

Table 2.5 reports estimates of linear probability models from the subset of hearings when at least a witness is heard (i.e., 62% of the total sample). Even when estimating congress and committee fixed effects, and conditioning on the partisan alignment between the committee majority and the president, the effects of president-committee chair partisan alignment is large, negative, and precisely estimate. Importantly, the effect of partisan alignment is larger and more precisely estimated than that of ideological alignment between the chair and the president (Models 3-4), and the estimates in column (5) shows that the effect of chair-president partisan alignment is even larger when conditioning on the chair-president ideological distance.[14]

Finally, the second test 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 use of statistical facts as a function of partisan alignment using a two-way fixed

[14]The chair-president ideological distance variable is computed as the absolute value between the ideal points of the two actors. Ideal points measured with DW-NOMINATE scores from J. Lewis et al. (2020).

DV:	Bureaucracy as Witness [0,1]				
	Partisanship		Ideology		Both
Model:	(1)	(2)	(3)	(4)	(5)
Comm. Chair-President Partisan Align.	-0.066*	-0.063*			-0.207*
	(0.030)	(0.032)			(0.082)
Comm. Majority-Govt. Partisan Align.		-0.011		0.011	0.030
		(0.031)		(0.037)	(0.038)
Comm. Chair-President Id. Dist.			0.026	0.025	-0.110
			(0.057)	(0.057)	(0.074)
Fixed-effects					
Congress	✓	✓	✓	✓	✓
Committee	✓	✓	✓	✓	✓
Observations	5,179	5,179	4,776	4,776	4,776
R^2	0.094	0.094	0.084	0.084	0.085
Within R^2	0.001	0.001	0.000	0.000	0.001

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

Table 2.5: OLS estimates of the effect of committe chair-government partisan alignment on the probability of a bureaucracy appearing as a witness in Congressional Senate hearings. In Table B.11 in the Appendix I show results are robust to clustering SE at the presidency level.

effects estimator, in order to look at the change in use of statistical facts within legislators and legislative debate (i.e., date) and therefore holding constant all unobserved sources of heterogeneity at the legislator and debate level.[15] 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:

$$\text{Facts}_{i[l,a,d]} = \eta_l + \phi_a + \alpha_d + \beta \text{Leg-Govt. Alignment}_{l[d]} + \mathbf{X}'_{i[l,d]}\lambda + u_i \tag{2.4}$$

where $Facts_{i[l,a,d]}$ is the frequency of tokens considered statistical facts and evidence in speech i given by legislator l about agency a in date 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 identifies the effect of being co-partisan with the government.

The results reported in Table 2.6 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.

[15]Results are robust to estimating year fixed-effects instead of date fixed-effects, see Table B.12 in the Appendix.

DV:	Statistical Facts (Abs. Frequency)					
Country:	US			UK		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Leg.-Govt. Partisan Alig.	-0.128***	-0.194***	0.360	-0.089***	-0.158***	0.046
	(0.027)	(0.049)	(0.334)	(0.019)	(0.036)	(0.191)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Observations	247,570	247,570	247,570	171,155	171,155	171,155
R^2	0.218	0.256	0.282	0.235	0.321	0.667
Within R^2	0.139	0.181	0.177	0.083	0.152	0.601

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table 2.6: Partisanhsip, Ideology, and Statements' Positivity, US Data. OLS estimates. Dependent variable if absolute frequency of statistical facts 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.

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 50 words before and after the name of the agency, the frequency of statistical facts when there is partisan alignment decreases by -0.19 points for the US (average frequency among all speeches is 5.69) and by -0.16 points for the UK (average frequency among all speeches is 2.09). If we consider the average use of facts as a baseline, being co-partisan reduces the use of facts by approximately 3% for the US and 8% 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 partisanship does not affect the speech as a whole, but rather the portion of speech about bureaucracy captured by the windows of words.[16]

[16] 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 B.13) and that there is no statistically significant association between legislator-agency partisan congruence or ideological distance and argumentative style (see Table B.14).

2.6 Discussion

Politicians' ability to hold bureaucracies to account is a central topic in political science and a cornerstone of bureaucratic legitimacy. While scholars have shown how legislators design institutions to hold agencies to account, no attention has been paid to the possible partisanship-induced distortions that can systematically affect legislators' incentives to objectively evaluate and oversee bureaucracies. In this paper, I show that partisanship triggers selective accountability. I proposed that when legislators hold bureaucracy to account, they do so selectively, giving more positive statements and exerting a lighter oversight when co-partisan with the government. This poses a threat to effective accountability. I make several empirical contributions too. I use natural language processing techniques to estimate the positivity of statements given by legislators from different parties, collect original data on bureaucracies appearing as witnesses before Senate Congressional Committees, and I provide credible inference with two studies and four tests. The data shows large support for the expectations.

Since government legislators care about the reputation of their party, and an under-performing bureaucracy has negative consequences for the governing party's electoral support, government legislators have an incentive to portray bureaucracy more positively compared to when they are at the opposition. Moreover, if positive statements are a political strategy to sustain the image of the government party irrespective of bureaucracies' actual performance, government legislators are less likely to acquire and use factual information when holding agencies to account. Consequentially, when government legislators argue about bureaucracies they resort less frequently to statistical facts and quantitative evidence about bureaucracy. The decrease in probability of bureaucracies appearing before committees when there is partisan alignment between the chair and the president is strong evidence against the alternative strategy discussed in the theoretical section presented above. In fact, if legislators – motivated by the electoral costs of an under-performing bureaucracy – tightened their oversight activity to prevent bureaucratic failures, we should not observe a drop in the probability of bureaucracies appearing before committees. However, the less frequent appearances of bureaucracies in committees alongside the more positive evaluation of bureaucracies suggest that co-partisanship with the government reduces legislators' incentives to hold bureaucracy to account.

As far as the generalisability of the results is concerned, there are reasons to believe partisan selectivity is able to alter bureaucratic accountability in other countries with strong and rooted political parties, with or without a politicised civil service. In fact, I find similar effects of partisanship in two countries with very different administrative systems, politicised in the US and neutral civil service in the UK. However, limits to the generalisability of the results may arise from other features of the US and UK political systems. Both countries are advanced democracies with high levels of political 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. For instance, absent a strong, direct link between legislators and constituents, legislators' incentive to selectively hold bureaucracies to account might be weaker, and legislators might be better off objectively evaluating the performance of agencies rather than portraying them under a more positive light. Future research could look at such different contexts and at how bureaucracies – knowing legislators evaluate agencies in a partisan fashion – administer policies.

Chapter 3

"Listen to me": Ideological Agreement and Bureaucratic Influence in the Legislative Arena

Abstract

The political control of the bureaucracy remains a classical topic in political science. However, little is known about its reverse: bureaucracies influencing politicians. I conceptualise bureaucratic influence as the extent to which legislators use the information produced by agencies in the legislative process. I introduce a new measurement strategy to estimate legislators' use of bureaucratic information which combines syntactic analysis and dictionary-based approaches and apply it to a corpus of 6.8 million speeches given by US congresspersons in floor and committee sessions. Building on cheap talk models of strategic communication, I argue that legislators make greater use of bureaucratic information when ideologically closer to agencies and that agency independence – operating as a credibility-enhancing mechanism – mitigates the effect of ideological distance. I find strong support for bureaucratic influence being ideology-driven, while the evidence is weaker for the ability of statutory independence to mitigate the effect of the ideological divide.

3.1 Introduction

> *Several Senators from the Upper Midwest insisted that the Office of Management and Budget do a study on the effects of the [Dairy] Compact. The OMB report is called "The Economic Effects of the North-east Interstate Dairy Compact". I will be quoting a lot from that study that those Senators wanted in this floor statement.*
>
> Sen. Patrick Leahy, D–VT

This is one of the opening statements of a speech given by Democratic Senator Leahy, VT on 6th October 1998. In his speech, he cites 13 times what claimed by the OMB in the study. Bureaucratic agencies, due to their expertise, produce a great wealth of information that can be used by politicians to reduce uncertainty over policy outcomes or to strengthen their own argument. However, information can also be a an opportunity for agencies to influence political decisions. In this paper I build on theories of strategic communication and show that bureaucratic influence – namely politicians' use of bureaucratic information – decreases when legislators and agencies have divergent policy preferences.

Couched within the principal-agent framework, the scholarship on bureaucratic politics made important strides to enhance our understanding of how politicians seek to exert control over bureaucracies in order to prevent agency slack and restrain bureaucratic autonomous policy-making (Epstein and O'Halloran 1999; Gailmard 2009; McCubbins, Noll, and Weingast 1987; McCubbins and Schwartz 1984), yet very rarely have scholars looked at the influence that bureaucratic agencies exert on the main political decisions that, in theory, should rest with elected politicians. Indeed, though bureaucrats are legally subordinated in the hierarchy of government, "*they can exert political power over their own superiors*. When this happens [...], they can play major roles in determining [...] what policies the latter pursue once in office" (Moe 2012, 37).

Bureaucracies have always been considered rich sources of information for politicians (Niskanen 1971; Wilson 1989), but empirical scholarship has struggled to document and study influence behind the closed doors of government organisations. One of the first attempts to interpret agency-political relations as unfolding on a "two-way street" emerges from Krause's work (Krause 1996, 1999). Krause depicted politicians' decisions to control the agency through budgetary allocation as the result of the interactions between both the principals and the agency In his study of the US Securities and Exchange Commission (SEC), he shows that the budgetary preferences of the government with respect to the SEC are influenced by the SEC's regulatory performance. Beyond budget preferences, a prominent attempt to theorise the influence of bureaucracy on policy formulation is Carpenter's reputation-based account of bureaucratic autonomy (Carpenter 2001a, 2010). In *The Forging of Bureaucratic Autonomy*, Carpenter argues that bureaucratic reputation – a set of symbolic beliefs about an organisation embedded in a network of multiple audiences – allows agencies to secure their desired policies despite the opposition of elected politicians (Carpenter 2001a, 3–4).

More recent scholarship on bureaucratic politics has started to study role of the bureaucracy in different stages of the legislative process. Nicholson-Crotty and Miller (2012), for instance, find a positive relationship between agency perceived effectiveness and politicians' perceptions of bureaucratic influence on legislative outcomes, and Ingold and Leifeld (2016) find that vertically integrated offices with access to formal decision-making venues are on average perceived as more influential. However, despite few exceptions (e.g., Kroeger 2020), scholarly work has generally relied on perception measures of influence, easily susceptible of social desirability bias, which could both inflate or deflate the actual influence exerted by the bureaucracy. In the attempt to overcome over- and under-reporting, Blom-Hansen, Bækgaard, and Serritzlew (2020) implement a series of experiments simulating the decision-making process and find that bureaucrats are willing to use their information to influence politicians' decisions, who are in turn likely to rely on bureaucrats' expertise depending on how the information is framed. While this scholarship made important advancements in the study of the role of bureaucracies in the policy-making process, we know little about the extent to which they capture real world phenomena.

In this paper I propose an information-based concept of bureaucratic influence, defined as the extent to which legislators' use the information produced by bureaucratic bodies in the legislative process. When legislators form their opinions about policy, they ar exposed to multiple sources of information. Bureaucracies are rich sourced of precious expertise for politicians (Gailmard and Patty 2013). When politicians pass and discuss policy, they use the information and the expertise of bureaucracy to form their beliefs about the expected consequences of certain measures. Similarly, they can use the information produced by bureaucracy to increase the persuasion of their appeals or to maximise the probability of success of their legislative enterprise. In all these instances, bureaucracies enter the legislative arena and affect political outcomes. This information-based definition of influence is consistent with a long tradition of work in political science that conceives bureaucracies as shaping policies through information (Weber 1922; Aberbach, Putnam, and Rockman 1981; Workman 2015). Importantly, it also allows to capture influence as a political practice, rather than a set of perceptions.

The theoretical gist of an informational definition of bureaucratic influence can be found in Crawford and Sobel's (1982) cheap talk model of strategic communication. Although the model is general in its formulation, it can easily be applied to the communication game between bureaucracy and legislators (Gailmard and Patty 2012). An expert bureaucracy sends a signal or information to a legislator who will then make a policy decision. Because the legislator cannot verify the quality or veracity of the information, truthful communication is only achieved when both the bureaucracy and the legislator have similar preferences over policy outcomes. The more ideologically apart, the less likely it is that legislators will use the information produced by agencies. However, when agencies are insulated from political pressures and enjoy a high level of statutory independence, the ideological leaning of agencies loses relevance for politicians, and the role of the ideologicla divide is weaker.

I test this theory in the US context. I present a new measurement strategy that detects when legislators use the evidence and statistical facts produced by bureaucratic bodies in legislative speeches and apply it to a corpus of 6.8 million speeches give by US congresspersons in floor and committee sessions. First, I apply syntactic dependency parsing to the corpus of speeches and extract the information produced by agencies and used by legislators. Second, I measure the frequency of words considered statistical facts and evidence and produce estimates of bureaucratic influence for every speech in which the agency is used as a source of information.

I leverage within-legislator variation in ideological distance from bureaucracies as a result of changes in the political leanings of agency officials and estimate the effect of ideological disagreement on politicians' use of bureaucratic information with a series of two-ways fixed effects estimators. I find strong support for the ideology-driven account of bureaucratic influence. If we consider the average score of bureaucratic influence as baseline, a one-unit increase in ideological distance leads to a decrease in bureaucratic influence by more than 20%. I also find support for the moderating effect of agency independence. Overall, ideological distance plays a larger role for agencies that are more controlled by politicians, but the difference in the effect of ideological distance for more and less independent agencies is distinguishable from zero only at 90% confidence level.

With this paper, I make two contributions to the literature on bureaucratic politics. First, I present the largest attempt to measure the role of bureaucratic bodies in legislative politics, presenting fine-grained data for 237 agencies and approximately 40 years of floor and committee speeches. Second, I show how ideological differences can limit bureaucracies' role in the legislative arena. Methodologically, I introduce a new transparent and objective way of measuring political influence which can be used to answer several questions in political science.

3.2 Bureaucrats and Politicians: A Cheap Talk

A vast literature in political science has studied how the ideological leanings of bureaucracies and politicians affect structural characteristics about the agency as well as informal behaviour of agencies and politicians, with topics spanning from delegation of authority (Epstein and O'Halloran 1999) to bureaucratic oversight (McCubbins and Schwartz 1984; Lowande 2018), executive policy-making (Bolton, Potter, and Thrower 2016; Potter 2019), and performance (Spenkuch, Teso, and Xu 2021).

Epstein and O'Halloran (1994), for instance, show that the discretion delegated to bureaucracy decreases as the preferences of Congress and those of executive agencies move apart, and Bolton, Potter, and Thrower (2016) find that, as ideological disagreement between the President and the Office of Information and Regulatory Affairs – which is in charge of overseeing every rule passed by administrative agencies – increases, the agency's review times increase, thus inhibiting the entry into force of rules when the agency is ideologically apart from the President. Similarly, Potter and Lowande (2020) show how legislators who are ideologically distant from the proposals made by the Environmental Protection Agency are more likely to scrutinise the proposal by filing requests for documents, additional hearings, and extended time for public participation. On the performance side, Spenkuch, Teso, and Xu (2021) find efficiency gains for procurement activities when bureaucrats are ideologically aligned with the President.

This strand of the literature shows that political factors such as the ideological leanings of the actors involved in administrative politics affect both formal (e.g., delegation of authority) and informal (ex post oversight) practices. In this paper I show how ideology is important for explaining bureaucratic influence and legislators' use of bureaucratic information too.

Politicians rely on bureaucratic offices to acquire information about both the nature and the solutions to the problems they face. However, besides reducing uncertainty over policy outcomes, bureaucratic information can also represent a channel for bureaucracies to achieve their own goals, and depart from the policy preferences of politicians (Aberbach, Putnam, and Rockman 1981). While political principals need bureaucrats' expertise, information asymmetries make it hard for politicians to know whether the information provided by bureaucrats is consistent with their own policy preferences (G. J. Miller 2005). As a result, politicians face the moral hazard of trusting bureaucracies who might in fact pursue different policy goals. Politicians will therefore have to decide when to trust bureaucracies and let the information they produce influence policy-making. The canonical model of this strategic form of communication is Crawford and Sobel's (1982) cheap talk communication model. In particular, I follow Gailmard and Patty (2012) and apply this framework to a situation in which an expert bureaucracy produces information in the attempt to shape the decisions taken by legislators.

Crawford and Sobel (1982) present a game in which an actor, the sender, tries to influence the decision of another actor, the receiver, who has the power to make authoritative decisions whose consequences affect the welfare of both actors. Let us consider the following hypothetical scenario. A country is experiencing a harsh economic crisis. Legislator L needs to pass a law aimed at restoring the economy. Bureaucracy B, because of its mission and capacity, is in a strong position to provide L with the necessary information in order to maximise the positive outcome of legislation. Both L and B have known policy preferences. Importantly, L cannot verify the quality or veracity of the information produced by B. As a result, L relies on the policy preferences of B as an heuristic to decide whether using the information produced by B in shaping political decisions. The key prediction of the model is that the probability of truthful communication increases as B and L's preferences over outcomes become more similar.

When ideological disagreement between a bureaucratic agency and a legislator is high, bureaucratic influence is low, and the probability that the legislator uses the information produced by the agency in debating and passing legislation decreases.

Hypothesis 1: Bureaucratic influence decreases with ideological distance.

As a real world example, let us consider the US Environmental Protection Agency administered by Edward Scott Pruitt between 2017 and 2018, considered by most political commentators a climate change denier (Meyer 2017). The influence of the EPA on politicians' opinion about environmental policy during Pruitt's mandate will be larger for members of congress who are conservative with respect to climate action and environmental policy, whereas congresspersons who endorse policies aimed at reducing global warming will likely be reluctant to using the information produced by the EPA when debating environmental policy. The influence of the EPA is therefore political: larger when there are shared policy goals between the agency and legislators, lower when the goals are far apart.

One important assumption of the model is that legislators cannot rely on an independent system which verifies the information produced by agencies, and hence have to count on ideology as a heuristic when deciding whether the information is consistent with their policy goals. However, when agencies are insulated from political pressures, their ideological leaning is less salient and the legislator-agency ideological distance plays a weaker role in legislators' choice over bureaucratic information. Independence, acting as a credibility-enhancing mechanism, tempers the distrust of politicians towards information produced by ideologically distant agencies.

Granting statutory independence to agencies is a powerful signal of credible commitment and can be an effective solution to the policy inconsistency inherent to changing governments (G. J. Miller 2005). By delegating independence and authority to agencies, elected politicians raises the barrier between politics and administration, hence reducing the degree of control that the government exerts on bureaucracy. A clear example of such commitment is the independence of central banks and regulatory agencies for the credibility of monetary policies, for controlling inflationary tendencies, and for ensuring a level-playing field to public and private businesses (Cukierman, Webb, and Neyapti 1992; Keeper and Stasavage 2003; Gilardi 2002). Empirically, agency independence has been shown to improve bureaucratic policy-making, in particular the perceived and objective quality of regulation (Bertelli and Whitford 2009; Koop and Hanretty 2018). "Depoliticising" bureaucratic bodies by isolating them from political pressures has consequences not only for the credibility of the commitment to a specific policy, but also for the actions and preferences of the bureaucracy, which do not respond to the political will of the government of the day. Free from political influence, independent agencies – and the statutory provisions which define their relationship with political officials – make the ideological leaning of the agency less salient for legislators, who are more likely to use the information produced by agencies regardless of their ideological position. Agency independence counteracts the effect of the ideological position of agencies on legislators' use of bureaucratic information.

HYPOTHESIS 2: The negative effect of legislator-agency ideological distance on bureaucratic influence is weaker for more independent agencies.

This account of ideology-driven bureaucratic influence shows how political the use of information is in the legislative arena, and how statutory independence can mitigate the role of ideological differences.

3.3 Measuring Bureaucratic Influence

The influence exerted by bureaucratic bodies in the legislative process has generally been measured either qualitatively (Carpenter 2001a; Page 2012) or through perception measures. While qualitative measures, though benefiting from "deep" observation and multiple sources of data, are limited to few cases, answers to questions like "How influential do you think agency x is?" are easily susceptible of social desirability bias. Both self-reported measures of *received* influence (i.e., legislator attributing an influence score to actors) as well as self-reported measures of *exerted* influence (i.e., actor self-evaluating their influence) could either deflate or inflate the actual influence exerted by the bureaucracy.

Blom-Hansen, Bækgaard, and Serritzlew (2020) address social desirability bias with a set of experimental designs and find evidence in support of the demand and supply of bureaucratic influence. They first show that a minority of bureaucrats are willing to organise the information they pass on to politicians "in a way that makes it easy for politicians to choose the solution that bureaucrats consider the best." They also show that politicians rely significantly on bureaucrats' expertise and are also susceptible to the way bureaucrats frame the information.

Beyond perception measures, Kroeger (2020) exploits the fact that some US states publish the number of bills sponsored by state departments and is therefore able to measure the success rate of department-sponsored bills. She finds that bureaucracy-sponsored bills are more likely to be approved by the legislature when there is unified government and when the capacity of the legislature is weaker compared to that of the bureaucracy. This is a very accurate way of looking at the role of bureaucracy in the legislative process. However, the availability of data on the formal involvement of departments in the legislative process is extremely scarce, and – as acknowledged by the author – just looking at sponsored bills neglect the possibility that agencies stop or postpone bills.

In this paper I present a new large scale measurement strategy that is more flexible at capturing the extent to which an agency is influential in the legislative process by applying natural language processing techniques to a large corpus of floor and committee speeches given by US legislators, detecting when legislators use agencies' information and extracting what type of information they use. This measurement strategy has quantitative and qualitative advantages over existing methods. First, by looking at floor and committee speeches, I am able to trace how legislators use the evidence produced by a large set of bureaucratic bodies over a long period of time and on a daily (or debate) basis. Second, I am able to measure the intensity of influence, capturing the frequency and intensity of politicians' use of bureaucratic information. How many times do legislators use the information produced by the Federal Reserve or the EPA? How does the use of information change over time for every agency? Qualitatively, by focusing on the frequency of specific terms capturing statistical facts and quantitative evidence, I can measure the portion of information that taps into the agency's expertise and which is ultimately grounded in hard evidence produced by the agency. I can therefore isolate the evidence-based part of information produced by agencies and used by legislators.

Floor and congressional committee debates are appropriate venues to look for such information. While key decisions might be taken behind closed doors, public debates and congressional committees remain highly salient venues where policies are made. It is in congressional committees where legislators have detailed discussions about policy, advancing their arguments in support of specific bills, or proposing amendments to existing laws. Similarly, in floor debates legislators can focus on the core parts of the law and make more general considerations about policy. In both venues, legislators might rely on agencies' expertise to debate and make policies. In particular, by looking at both congressional committees and floor debates, I am able to capture both the more political rhetoric of floor speeches as well as the more informal, substantive conversations going on in committees.

3.3.1 Measuring Influence Through Information

The key assumption of the proposed measurement strategy is that legislators' use of bureaucratic information can be detected by parsing the syntactic relations of terms in segments of text (e.g., sentences). Syntactic analysis can in fact identify the action of saying something, the subject carrying out the action, and the object of the action. Let us consider a legislator who convinced about a statement given by the Federal Reserve on interest rates and economic growth. She might say "The Federal Reserve *[subject]* said *[action]* that higher interest rates will strengthen the economy *[object]*." By creating extraction rules that detect certain syntactic relationships, I can therefore match every instance in which an agency is used as a source of information, in order to then measure the quality of information that is being used. Importantly, this method allows me to compare the use of information over time, across agencies, and legislators.

Syntactic analysis and dependency parsing are new frontiers in political science research, but few promising applications show the benefit of retaining dependency relationships between words when analysing text. Van Atteveldt et al. (2017), for instance, show the differences in how English-language Chinese media covered the 2008-9 Gaza war and find how US media underscore Hamas' attacks and Israel's right to defence, whereas Chinese media do not portray Hamas as attacking and focus more on the Israeli military operation and the humanitarian consequences. In a very different context, Vannoni, Ash, and Morelli (2021) apply syntactic analysis to a corpus of US state laws to estimate delegation of powers to governors of US states. By extracting syntactic structures encoding delegations and constrains they create a validated measure of delegated authority and test the classical prediction whereby delegation decreases with divided government (Epstein and O'Halloran 1999; Franchino 2004). Consistently with theory, they show that the number of provisions delegating powers to the governor is associated with government unity. Similarly, Ash et al. (2020) shows how these methods can efficiently extract workers' rights and duties from labour union contracts.

The measurement strategy I propose consists of three steps. First, I split every speech mentioning the name of an agency into sentences and process them using a syntactic dependency parser. The parser tags parts of speech (e.g., subject, verb, predicate, etc.) and detect dependency relations. Second, I extract clauses that match syntactic frames. By pre-defining certain extraction rules (e.g., *subject + say verb + quote*), I can isolate the action of saying something, the source of information, and the content of the quote. I am therefore able to isolate sentences where legislators quote bureaucratic agencies or documents and reports produced by agencies. Third, I isolate the quote, namely the actual piece of information used by legislators, and measure the extent to which the quote reports statistical facts and evidence. Eventually, I will obtain a sample of speeches where agencies are used as source of information, and every speech will receive a continuous measure of bureaucratic influence consisting of the sum of the statistical facts words contained in the quote. Theory predicts that this measure will decrease with ideological distance between the legislator and the agency used as source.

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

First, I tag and parse the sentences with the spaCy parser (Honnibal, Goldberg, and Johnson 2013; Honnibal and Johnson 2015).[1] SpaCy operates as a supervised learning algorithm, with the goal of making predictions based on training and labelled data. It achieves state-of-the-art performance on both accuracy and speed (Choi, Tetreault, and Stent 2015). After splitting speeches into sentences, the parser tags parts of speech and detects dependency relations.

For instance, let us consider the previous example about a congressperson reporting the statement of the FED. In her speech, she says: "The FED said that higher interest rates will strengthen the economy." The tokens – namely each single word – within this sentence have syntactic properties and follow specific dependency relations. For instance, "The" refers to the "FED," which in turns is the nominal subject of the verb "to say." The result of syntactic parsing is displayed in Table @ref(table:dep.par), which reports the token ID, the token (i.e., the word), the part-of-speech, the ID of the head token (namely the "parent" token), and the type of dependency relation. For instance, the head token ID of the token "higher" and "interest" is the token ID 7, namely "rates."

3.3.1.2 Step 2: Extraction Rules

Once the parser has tagged each token of the sentence, I annotate the sentence based on extraction rules that detect when a statement is reported, the source of the statement, and the content of the statement.

[1] Version 2.1.6 implemented through the R package *spacyr*.

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

Table 3.1: Dependency parsing example. Token IDs in bold used as example in text.

I create a comprehensive list of five extraction rules that match who-says-what syntactic structures: one for direct or indirect statements (the FED said; as said by the FED), one for "according-to" structures (according to the FED), one for direct nominal recommendations (the FED's proposal is), and one for indirect nominal recommendations (the FED's proposal to). To match direct and indirect statements, I specify a vector of "say verbs" so that the parser marks the lemmatised version of the verb – therefore capturing verbs declined in every form (active or passive) or tense – and its respective subject or, in case of indirect statement, the agent. For "according-to" structures, the parser detects the lemmatised token "accord" and the object of the preposition, which will be the source of the information. For direct and indirect nominal recommendations, I specify a vector of recommendation-related words for the parser to detect, and their possessive determiner – i.e., the owner of the recommendation – will be labelled as the source of the recommendation.[2] Finally, all the tokens that are dependencies of say verbs, recommendation-related verbs, or according-to structures are labelled as quotes. For instance, for the FED's example, the information about inflation boosting the economy could be used in a speech in five different ways. Table 3.2 below reports the precise tokens and syntactic structures used to compile the extraction rules, as well as the toy sentences in which a legislator could use the information produced by the FED, with the quote in italics.

I then apply the extraction rules to the tagged sentences. Figure 3.1 shows the final output of the syntactic analysis for the example of indirect nominal recommendation extraction rule, one that might seem particularly challenging to extract. Dependency trees

[2] Say-verbs and recommendation-type words are reported in Section C.1 in the Appendix.

Extraction Rule	**Syntactic Structure**	**Sentence Example**
Direct Statement	subject + say verbs	The FED said *that higher interest rates will strengthen the economy.*
Indirect Statement	agent + say verbs	As reported by the FED, *higher interest rates will strengthen the economy.*
According-to Structure	accord + object of preposition	According to the FED, *higher interest rates will strengthen the economy.*
Direct Nominal Recommendation	recommendation + possession modifier	The FED's recommendation is *to increase interest rates.*
Indirect Nominal Recommendation	recommendation + possession modifier	I fully endorse the FED's recommendation *to increase interest rates.*

Table 3.2: Illustrative examples of sentences matching the five syntactic frames.

of other rules are shown in Figure C.1 in the Appendix.

3.3.1.3 Step 3: Analysing Quotes

Bureaucratic influence does not occur by just using the agency as a source of information. In fact, politicians might report what said by agencies with a negative tone (e.g., "The FED said nothing about it!") or they could cite an agency without making any reference to policy (e.g., "The FED said that in the long term we're all dead."). Step three of the proposed measurement strategy addresses this issue by extracting qualitative information from the quote, hence establishing whether the information used by politicians taps into the expertise of the bureaucracy.

For each tagged sentence, I extract the quote – namely the information produced by the agency that has been reported by the legislator – and compute the frequency of statistical facts and evidence words in the quote. Sentences which contain the name of an agency but where the agency is not used as a source of information are removed, for I want to compare the use of evidence when the agency is used as a source of information. I follow Hargrave and Blumenau (2020) and apply a dictionary-based approach to measuring the use of statistical facts and evidence in speeches. This step is important to ensure legislators' are actually using expertise-based evidence and information produced by the agency.

Every quote is assigned a score which equals the absolute frequency of words belonging to a pre-defined dictionary of statistical facts and evidence words.[3] I use the off-the-shelf LIWC dictionary (Pennebaker et al. 2015), which contains a comprehensive list of words

[3] In the Appendix I show the results are robust to using the term-frequency inverse-document-frequency measure of the use of statistical facts and evidence (Table C.5).

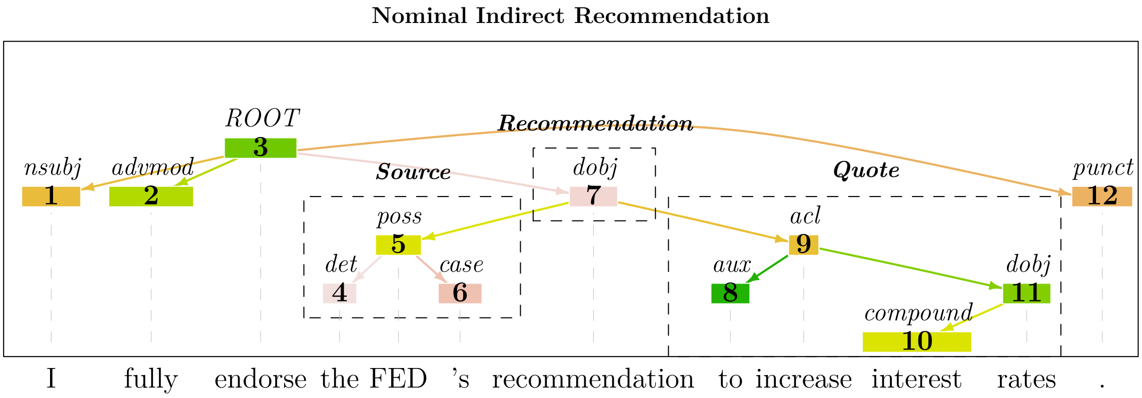

Figure 3.1: Parsed dependency trees of one illustrative example where the FED is used to support a statement. Implemented through the *rsyntax* package in R.

related to quantifiers and numbers, such as “amount,” “approximately,” “average,” “entirety,” “equal,” “less,” “multiple,” “percentage,” “whole,” “twice,” “total,” as well as all digits and numbers used to express quantities.[4]

Once the dictionary analysis has been applied to all the quotes, I merge all the quotes back at the speech level, and sum the statistical-facts scores of each quote. The key quantity of interest is at speech level and combines all the quotes contained in the speech. More formally, it will be determined by the following metric:

$$\text{Use of Bureaucratic Information}_s = I_s = \sum_{i=1}^{N} facts_i, \quad \text{where} \quad facts_i = f_{Qi \in Dict}$$

where the use of information for each speech I_s is given by the sum of quote-level measures of the use of statistical facts and evidence $facts_i$, which in turn is given by the absolute frequency of the tokens of the quote Q_i which appear in the fact dictionary *Dict*. When more agencies are used as source of information in a speech, I consider the agency with the largest statistical-facts score the one used as source of information.

3.3.2 Validation

I present two sets of validation tests, one for the dictionary and one for the measurement strategy as a whole.

Dictionary-based approaches are deemed to be highly context-dependent and therefore

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

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. To support this claim, in Section B.3.2 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.e., use of facts in online medical blogs). I find a positive and large 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 B.3).

To validate the measurement strategy as a whole, I first present the output of the measurement at different levels of intensity of the use of statistical facts captured by the machine. This step should ensure the construct and face validity of the measure. In Table 3.3, I report four examples of sentences in which there are different levels of the use of statistical facts in politicians' quotes of bureaucratic agencies. In particular, I extract sentences whose statistical-facts score equals the first, second, and third quartiles of the distribution.

Finally, I assess the convergent validity of the measure by comparing the output produced by the machine with the outputs produced by human coding. I extract a random sample of 500 sentences where agencies are used as source of information. I then dichotomised the statistical-facts score so that the quote of the agency either contains or does not contain statistical facts. Each of these sentences was manually coded by an independent researcher without knowing the score of the automated analysis nor any details about the automated text analysis performed.[5] Table 3.4 shows the confusion matrix of the classification exercise. The accuracy of the classification is rather high (i.e., 0.76) as well as the precision (i.e., 0.8. While the recall metric is slightly lower (0.63), the F1 score – assessing the balance between precision and recall – is satisfactory, accounting for 0.71.

[5]For instance, the independent researcher did not know the words contained in the dictionary, nor did they know I was performing a dictionary analysis. The unique task given to the researcher was to mark the sentences containing statistical facts and quantitative evidence produced by a bureaucratic body.

Score	Speaker	Year	Party	Sentence
Zero	Sen. Gramm, TX	1994	REP	All of his peers at the Department of the Treasury said he ought to get out of it, he's one of the President's closest friends, he has this long connection with the President from being in college, people are going to say at least there is a potential conflict of interest here with his friend.
Low	Sen. Cohen, ME	1995	REP	The Department of the Navy argues that the risks imposed by consolidating to a single nuclear-capable shipyard outweigh the potential cost savings.
Medium	Sen. Lautenberg, NJ	1990	DEM	According to the Environmental Protection Agency estimates, the number of persons expected to die of lung cancer as a result of radon exposure over their lifetimes can be as high as between 440 and 770 per 1,000 for those exposed to 1.0 WL of radon; as high as 270 to 630 for those exposed to 0.5 WL of radon; and as high as 120 to 380 for those exposed to 0.2 WL of radon.
High	Sen. Durbin, IL	1997	DEM	The US Bureau of Labor Statistics reports that the number of major strikes in the United States has been reduced by more than 90 percent since the middle of this century, from 470 major strikes in 1952 to 37 in 1996, and the number of workers involved in these strikes has been reduced by 90 percent, from three million workers involved in strikes in 1952 to fewer than 300,000 in 1996.

Table 3.3: Example of sentences where a federal agency is used as a source of information with various levels of statistical facts in committee speeches.

3.3.3 Limitations

This method has several advantages. First, it is fully scalable and transparent. Second, once the extraction rules are defined, it does not resort to subjective measures, while it allows to produce estimates of bureaucratic influence for a large set of agencies, legislators, and over a long period of time. Third, it allows to produce micro-level estimates, detecting when the agency is mentioned as a source of information, and the extent to which the information is about statistical facts and evidence. Fourth, it is a highly general strategy, which can be used to estimate influence of multiple actors thus contributing to other sub-fields in political science. However, despite the validation exercises reported above, a discussion of the strategy's limitations is in order.

First, this measurement strategy does not distinguish between politicians genuinely using bureaucratic information to form their opinions about policy and politicians who strategically deploy bureaucratic expertise to confirm a pre-existing argument. This dif-

		Manual Coding	
		Statistical Facts and Evidence	
		No	Yes
Automated Coding	No	234	37
	Yes	84	145

Table 3.4: Confusion matrix of computer and manual coding of 500 random sentences using a bureaucratic body as a source of information.

ference, though subtle, makes the proposed measure silent about the supply and demand of information, for it is not possible to detect whether influence is exerted by agencies or strategically crafted by politicians. This limitation arguably applies to all observational studies of influence. Let us consider researchers studying how influential scholars are. One intuitive way of measuring scholars' influence would be to look at citation patterns. Yet citations too are an intentional choice of other researchers who decide to use the work of someone else in their own. The demand of influence, namely legislators' choice of using bureaucratic influence, is itself a proof of the existence of influence.

Another important limitation pertains to the multiple ways legislators can use bureaucratic information. Legislators' can deploy bureaucratic information both explicitly and implicitly. Explicitly, for legislators could use the information produced by bureaucracies while at the same time acknowledging the source. Implicitly, for legislators could be exposed to the information produced by agencies and act without acknowledging the ownership of the information. By anchoring the quote to the name of the agency, the proposed method is only able to capture explicit ways of using bureaucratic information.

Nevertheless, perfect measures of bureaucratic influence are hard to produce, for researchers often face issues related to social desirability, experimental realism, or unobservable behaviours. The proposed measurement strategy, relying on observational data, does not suffer from social desirability bias or simulated reality issues, but it captures just a partial picture of the process of bureaucratic influence.

3.4 Sample & Data

3.4.1 Speeches and Agencies

I apply the proposed method to a corpus of 2,501,900 floor (1981-2016) and 4,545,416 (1990-2019) committee speeches. I downloaded transcripts of all floor speeches from the

Social Science Data Collection of Stanford University (Gentzkow, Shapiro, and Taddy 2018) and I obtained transcripts of congressional committee sessions through ProQuest.[6] I removed about 15,000 committee speeches given during oversight hearings to exclude information used by politicians to hold agencies to account rather than to exhibit influence.[7] After replacing the various ways in which agencies are mentioned with a standardised name, I subset all speeches mentioning at least one agency. The list of agencies combines large samples of bureaucratic bodies from Bertelli et al. (2013), Chen and Johnson (2014), and Selin (2015), integrated with information on the type of agency directly obtained from the US government website (usa.gov/federal-agencies), for a total of 302 agencies.

A total of 335,445 floor speeches and 473,478 committee speeches mention the name of at least one agency, namely 13.4% and 10.6% of the total sample, respectively. I parse these speeches into sentences and remove all sentences which do not contain the name of an agency. I then apply the extraction rules described in Table 3.2 to each sentence and subset the sentences in which the agency is used as a source of information. I extract the quotes from each sentence and apply the dictionary analysis to the quote, measuring the number of words that belong to the statistical facts and evidence LIWC dictionary.[8] Finally, I merge the quotes with their statistical-facts score at the speech level, for a total of 32,300 floor and 24,871 committee speeches. The average outcome is about the same in both floor and speeches, equal to 1.09 and 1.15 statistical-facts words, respectively. The sample of speeches on which I will perform the analysis therefore consists of all the speeches mentioning the agency as a source of information in floor and committee speeches, so that I can compare politicians using bureaucratic information when they are ideologically close (control group) or far (treatment group) from the agency. Descriptive statistics about the sample of speeches are reported in Table 3.5.

3.4.2 Ideological Distance and Agency Independence

Ideological distance is the key predictor of bureaucratic influence. I build a time-changing measure of ideological distance between legislators and the agency as the absolute

[6] A note on the quality of congressional committees' data and the speech parsing steps are reported in Section C.3 of the Appendix.

[7] I adopt a very conservative exclusion criterion, removing every speech which contains the word "oversight" either in the speech corpus, in the short description of the meeting, or in the list of topics produced by ProQuest.

[8] I remove sentences citing multiple agencies as source of information, 1.8% of the total for I am not able to attribute the quote to a single agency.

Descriptive Statistics	**Floor Speeches**	**Committee Speeches**
Initial Sample of Speeches	2,501,900	4,454,416
N. Speeches mentioning agencies	335,445	531,668
N. Speeches with agency used as source	32,300	35,127
Oversight Hearings		$-11,712$
N. Speeches with agency used as source	**32,300**	**23,415**
N. Agencies used as source	211	222
Average use of facts and evidence	1.09	1.15
Total N. Speeches	**55,814**	
N. Agencies used as source	**237**	
Average use of facts and evidence	**1.11**	

Table 3.5: Sample of speeches and final sample size.

value between the DW-NOMINATE score of each legislator and each agency used as source of information in the speech. Data on legislators' ideal point are from J. Lewis et al. (2020). For agency ideology, I use the dataset assembled by Chen and Johnson (2014), who produce donation-based ideology estimates for 79 federal agencies across five presidencies, from the first Clinton Presidency to the first Obama Presidency (1993-2012). Chen and Johnson (2014) use federal bureaucrats' campaign contributions to individual politicians as input to estimating agency ideology, and produce estimates comparable with the DW-NOMINATE Common Space scores. Because of the limited availability of data on agency ideology – both with respect to time and sample of agencies – the analysis will be limited to speeches given by legislators between 1993 and 2012 for which an agency ideology estimate is available, accounting for approximately one third of the sample of speeches.

I measure independence with three different indicators. First, I use the information on the type of agency, distinguishing from agencies listed on the website of the US government usa.gov/federal-agencies as independent agencies *vis-à-vis* government departments or executive sub-agencies. Second and third, I use the two indicators of agency statutory independence produced by Selin (2015), which captures agency independence along two dimensions: independence as the ability of an agency to make policy decisions without political interference; and independence as statutory limitations and requirements placed on the officials who manage the agency. The indicators are derived by modelling 50 structural features about the agencies with a Bayesian latent variable model. The estimates range between 0 and 4, with higher values signifying higher independence.

All these three measures of independence do not capture variation over time. However, structural features of agencies or the overall categorisation of agencies as independent or not are likely to remain fixed. In fact, as shown in Selin (2015, 983–84), the estimates about the independence of agency as the ability to take policy decisions without political interference derived from the initial statute establishing the law and from the US Code used by the author in 2013-14 are not different from each other. The second dimension of independence, however, displays some temporal change that needs to be taken into account when interpreting the results.

3.5 Methods

I am interested in two relationships: the effect of ideological distance on the use of bureaucratic information (*Hypothesis 1*), and how the effect of ideology varies for more or less independence agencies (*Hypothesis 2*).

There are three methodological concerns for identifying these effects. First, on the legislator side, there could be many individual characteristics that are correlated with ideology, their engagement with bureaucratic agencies, and the extent to which they use bureaucratic information in speeches. Education, socio-economic background, but also their level of interest towards bureaucratic policy-making. Second, on the agency side, some policy domains – e.g., financial or environmental regulation – might be more salient and therefore legislators might mention and use the information of the Federal Reserve or the Environmental Protection Agency more often than that produced by other agencies. Third, the salience of agencies/sectors can also change over time, and the fact that legislators use frequently the information produced by one agency might be the result of the agency being highly salient in that particular period of time, rather than being the result of more similar ideological positions.

I address these sources of omitted variable bias with a two-way fixed effects estimator. I leverage within-legislator variation in the use of bureaucratic information holding constant all time-invariant characteristics of legislators as well as yearly shocks that could affect the use of the information produced by agencies. I also include agency fixed effects to account for all time-constant agency differences. In particular, I estimate the following model:

$$I_{s[l,t,a]} = \gamma_l + \phi_t + \alpha_a + \beta Distance_{l[a,t]} + \mathbf{X}'_{[l,t]} + u_s \tag{3.1}$$

where $I_{s[l,t,a]}$ is the use of bureaucratic information (number of statistical-facts words) in speech s, year t, and given by legislator l, quoting agency a. γ_l and ϕ_t are legislator and year dummies, α_a agency fixed effects, and $Distance_{l[a,t]}$ is the ideological distance between legislator l and agency a in time t. $\mathbf{X}'_{[l,t]}$ is a vector of legislator-level covariates: whether the legislator is a subcommittee chair, majority-party member, majority-party leader, minority-party leader, their seniority and legislative effectiveness score[9] (data from Volden and Wiseman 2020). In particular, β estimates the marginal effect of ideological distance on the use of fact-and-evidence when quoting what said by a bureaucratic agency in the speech.

Despite Model (3.1) accounting for all the agency, year, and legislator-level characteristics that remain fixed, as well as time-changing characteristics of legislators capturing their seniority, membership to Congress, and law-making activity, legislators might use more or less information produced by agencies as a result of the changing salience of the agency. In times of financial instability, politicians will be more likely to use the information of the Federal Reserve, whereas in times of environmental disasters, they might count more often on what said by the Environmental Protection Agency. To account for this source of unobserved heterogeneity, I include in all the specifications the number of mentions the agency receives in speeches every year, which is a good proxy of the time-changing nature of agency salience. Furthermore, I show how the results are robust to including agency-year fixed effects to account for time-changing agency-level confounders which might affect the credibility of the agency's expertise and hence legislators' use of agencies' information. Standard errors are clustered at the legislator level.

This specification has several advantages. First, legislator dummies sweep out all the variation at legislator level. Second, agency dummies account for all time-invariant features of agencies such as the history, culture, mission, and policy sector. Third, the changing saliency and agency characteristics are captured by agency-year fixed effects. Fourth, common shocks or reforms affecting the bureaucracy as a whole are captured by year fixed effects. Fifth, conditioning on time-varying characteristics of legislators reduces the risk of bias stemming from omitted confounders at the legislator-level.

To estimate the moderating effect of agency independence (*Hypothesis 2*), I compare β across different samples based on whether the agency falls above or below the mean value

[9] The legislative effectiveness score synthesises several indicators about the proven ability of a legislator to advance her agenda items through the legislative process and into law (Volden and Wiseman 2020).

of the three independence indicators described in Section 3.4.2. I cut the data in three different ways: *i*) independent and non-independent agencies, *ii*) agencies above/below the mean value of independence from political review, *iii*) agencies above/below the mean value of independence as limitation and requirements on agency officials. We should expect the effect of ideological distance in the below-mean samples to be larger than the effect estimated on the above-mean samples. β estimated from the above-mean samples should instead approximate 0 and be a highly imprecise predictor of bureaucratic influence.

In Section C.4 I present a series of robustness tests. First, because the dependent variable is a count variable ≥ 0, I report regression estimates after log-transforming the dependent variable in order to downgrade the important of extreme values when estimating β (Table C.1). I also report the results of estimates of β produced using the Poisson estimator, which is generally employed for count data (Table C.2). Finally, results are robust to limiting the analysis to speeches quoting only one agency (Table C.5) and to allowing for heterogeneous effects based on the speech type, whether given in floor or committee debates (Table C.3).

3.6 Results

Democratic Senator Leahy's floor speech reported in the introduction is one of the speeches that makes the largest use of bureaucratic information. The extraction strategy outlined above matched 13 instances in which the Office of Management and Budget (OMB) was used to by the Senator to form his argument about dairy industry in New England. His statement is dense of evidence produced by the OBM. For instance, the Senator claims that "during the first 6 months of operation, the OMB reported that New England's dairy farm income rose by an estimated $2,227 million," or again that "the OMB reports that New England suffered a 20% decline in the number of farms with milk cows from 1990 to 1996." and that "The OMB's report states that the compact could support a small increase in participation during the demonstration period." Overall, 20 terms belonging to the "facts" dictionary appears in the 13 quotes of the OBM. This example shows the precision of the extraction rules, which are able to capture both the OBM's action of saying something, as well as the type of information reported.

The general pattern of legislators using bureaucratic information is far from being uniform over time, across parties, and partisan affiliation of legislators. Figure 3.2 shows

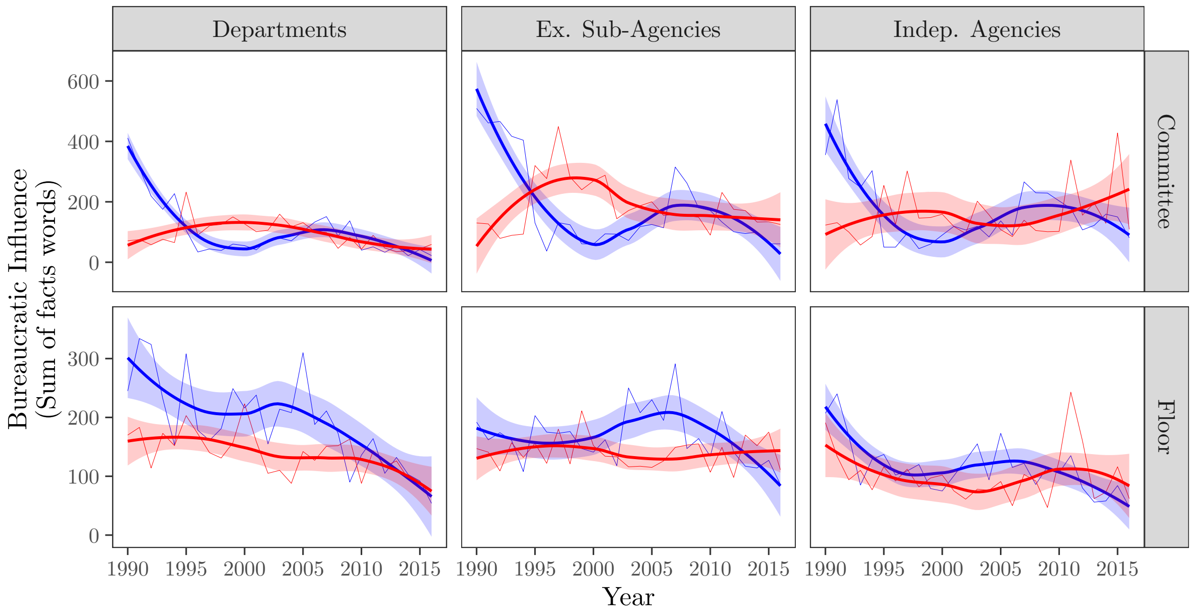

Figure 3.2: Sum of use of bureaucratic information across all agencies for every year and across three types of bureaucracy, with superimposed loess approximation. Blue line for Democrats, red line for Republicans.

the time trend of the use of bureaucratic information for every year, across parties, and three types of agencies: government departments, executive sub-agencies, and independent agencies, for both floor and committee speeches. The most striking pattern in the data is a constant drop in the use of bureaucratic information, possibly suggesting the decreasing role of bureaucracies in shaping legislators' opinions about policy. The y-axis shows the sum of the number of facts-words used in politicians speeches using agencies as a source of information. Interestingly, while Republicans' and Democrats' average use of bureaucratic information is roughly identical both in floor and committee speeches, the use of information increases with majority-party status, but only for Republicans. When there is a Republican president, Republican's use of bureaucratic influence is 7% higher compared to when they are at the opposition (p.value = 0.0005), whereas the difference is not statistically distinguishable from 0 for Democrats. The EPA, followed by the Treasury and Office of Management and Budget are the three bureaucracies whose information is most often used by legislators.

Moving from description to estimation and inference, Table 3.6 shows the regression results. Across various specifications, ideological distance has a negative effect on the use of bureaucratic information, even when conditioning on agency-year and legislator-year dummies (Models (4) and (5)). A one-unit increase in ideological distance leads to a decrease in the frequency of statistical facts and evidence by -0.23. If we compare the change in

DV:	Use of Bureaucratic Information (N. facts-words)				
Model:	(1)	(2)	(3)	(4)	(5)
Ideological Distance	-0.252**	-0.236**	-0.235**	-0.212*	-0.244*
	(0.085)	(0.080)	(0.080)	(0.084)	(0.123)
Log Agency Mentions	-0.093***	-0.075*	-0.075*		
	(0.015)	(0.032)	(0.032)		
Legislator Covs.			✓	✓	
Fixed-effects					
Legislator	✓	✓	✓	✓	
Year	✓	✓	✓		
Agency		✓	✓		
Agency-Year				✓	✓
Legislator-Year					✓
Observations	20,578	20,578	20,548	20,548	20,578
R^2	0.063	0.080	0.080	0.119	0.350
Within R^2	0.003	0.001	0.001	0.001	0.000

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

Table 3.6: OLS estimates. DV is frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

terms of the average use of information (i.e., 1.07), the marginal effect of ideological distance is associated with a decrease equal to 22% of the average.

Table 3.7 shows the estimated effect of ideological distance for agencies below and above the mean values of the three independence indicators. These are respectively a dummy variable which captures whether the agency is an independent commission and Selin's indicators about independence as requirements imposed on agency officials (Models *Decision Makers*) and as the degree of insulation of the agency's decision-making process from political review (Models *Political Review*).[10] As expected, the estimates are smaller and indistinguishable from 0 at 95% level for more independent – namely above-mean – agencies, whereas the effects are large, negative, and statistically significant for less independent – namely below mean – agencies. While the negative and large coefficient for less independence agencies might seem to lend support to *Hypothesis 2*, none of the difference between the estimated effects is distinguishable from 0 at 95% level. Only when splitting the sample between agencies with high levels of independence for agency officials

[10]For the first indicator, whether the agency is and independent commission or not, agencies are considered below mean if they are not independent commissions, and above mean if they are.

DV:	Use of Bureaucratic Information (N. facts-words)					
	Political Review		Decision Makers		Agency Type	
Above/Below Mean:	Below	Above	Below	Above	Below	Above
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Ideological Distance	-0.283**	-0.102	-0.290*	0.013	-0.274*	-0.063
	(0.106)	(0.153)	(0.122)	(0.116)	(0.119)	(0.154)
Floor Speech (Dummy)	-0.065	-0.035	-0.058	-0.065	-0.000	-0.150**
	(0.039)	(0.063)	(0.041)	(0.048)	(0.036)	(0.056)
Legislator Covs.	✓	✓	✓	✓	✓	✓
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency-Year	✓	✓	✓	✓	✓	✓
Observations	13,974	6,529	11,985	8,518	12,710	7,838
R^2	0.121	0.237	0.132	0.192	0.139	0.180
Within R^2	0.001	0.002	0.001	0.001	0.001	0.002

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

Table 3.7: OLS estimates across different samples of units below and above the mean value of independence indicators. DV is frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

(i.e., Models *Decision Makers*), does the difference in the estimated effects (0.21) approaches statistical significance (p.value = 0.086). Overall, there seems to be little support for the moderating effect of agency independence.

3.7 Discussion and Conclusions

The role of unelected bureaucracies in democratic government has received increasing attention in the last decades (Tucker 2018). Empirical evidence on the conditions under which this happens has two important normative implications. The first one concerns the role unelected officials exert on democratic processes In the quest for bureaucratic legitimacy, it is important to understand the extent to which the information and evidence produced by bureaucratic bodies penetrate the legislative process. The second one concerns the use of evidence in decision-making aimed at increasing the quality of policies. Bureaucratic expertise can in fact enhance the quality of legislation and, as a result, improve policy outcomes.

In this chapter I framed the influence of bureaucracy over legislators as a strategic communication game, which predicts that actors are less likely to undertake constructive communication when they have divergent preferences over policy outcomes. I further proposed that the institutional independence of agencies can counteract the negative effect of ideological distance on the quality of communication. Operating as a credibility enhancing mechanisms, agency independence represents a pledge of non-political information, which is more likely to be employed by legislators despite being ideologically far from the agency.

Methodologically, I introduced a new large scale measurement strategy that employs natural language processing techniques and syntactic analysis to detect when legislators use the information produced by agencies in floor and committee debates and measure the extent to which the information consists of statistical facts and evidence. I presented fine-grained data at speech level for over 40 years and more than 200 agencies. The findings support the key theoretical expectations. Bureaucratic influence decreases with ideological distance. Legislators ideologically apart from agencies make less frequent use of statistical facts and quantitative evidence produced by bureaucracy. Despite the weaker test and identification claims, there is suggestive evidence that this effect can be mitigated by agency independence.

This chapter makes three contributions to the literature on bureaucratic politics. Theoretically, it combines rational choice models of inter-institutional communication with politician-bureaucracy interactions, and statutory accounts of bureaucracy. Methodologically, it introduces a new strategy to measure the role of bureaucracy in legislative politics, which can be used by other scholars in other sub-fields to estimate the influence of a multitude of actors. Researchers could study how influential legislators are based on how often parliamentarians cite one another. The agenda setting scholarship could also benefit from knowing the extent to which politicians cite media outlets, interest groups, or trade unions. Researchers could also estimate the influence of political or religious leaders analysing the citation patterns of their statements in the media, in speeches, or political campaigns.

Bibliography

Aberbach, Joel, Robert D. Putnam, and Bert Rockman. 1981. *Bureaucrats and politicians in Western democracies.* Cambridge, MA: Harvard University Press.

Anastasopoulos, L. Jason, and Anthony M. Bertelli. 2019. "Understanding Delegation through Machine Learning: A Method and Application to the European Union." *American Political Science Review*, 1–11. https://doi.org/10.2139/ssrn.3207821.

Anastasopoulos, L. Jason, and Andrew B Whitford. 2019. "Machine Learning for Public Administration Research, With Application to Organizational Reputation." *Journal of Public Administration Research and Theory* 29 (3): 491–510. https://doi.org/10.1093/jopart/muy060.

Ash, Elliott, Jeff Jacobs, Bentley Macleod, and Dominik Stammbach. 2020. "Unsupervised Extraction of Workplace Rights and Duties from Collective Bargaining Agreements." *International Workshop on Mining and Learning in the Legal Domain*.

Ashworth, Scott, and Ethan Bueno De Mesquita. 2014. "Is voter competence good for voters?: Information, rationality, and democratic performance." *American Political Science Review* 108 (3): 565–87. https://doi.org/10.1017/S0003055414000264.

Bach, Tobias. 2012. "The involvement of agencies in policy formulation: Explaining variation in policy autonomy of federal agencies in Germany." *Policy and Society* 31 (3): 211–22. https://doi.org/10.1016/j.polsoc.2012.07.003.

Barberá, Pablo, Andreu Casas, Jonathan Nagler, Patrick J. Egan, Richard Bonneau, John T. Jost, and Joshua A. Tucker. 2019. "Who Leads? Who Follows? Measuring Issue Attention and Agenda Setting by Legislators and the Mass Public Using Social Media Data." *American Political Science Review* 113 (4): 883–901. https://doi.org/10.1017/S0003055419000352.

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. 2022. "A Dynamic Measure of Bureaucratic Reputation: New Data for New Theory." *American Journal of Political Science*, 1–18. https://doi.org/10.1111/ajps.12695.

Benoit, Kenneth, Kohei Watanabe, Haiyan Matsuo 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.

Bertelli, Anthony M., and Andrew B. Whitford. 2009. "Perceiving credible commitments: How independent regulators shape elite perceptions of regulatory quality." *British Journal of Political Science* 39 (3): 517–37. https://doi.org/10.1017/S0007123409000623.

Bisgaard, Martin. 2015. "Bias will find a way: Economic perceptions, attributions of blame, and partisan-motivated reasoning during crisis." *Journal of Politics* 77 (3): 849–60. https://doi.org/10.1086/681591.

———. 2019. "How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning." *American Journal of Political Science* 63 (4): 824–39. https://doi.org/10.1111/ajps.12432.

Blom-Hansen, Jens, Martin Bækgaard, and Søren Serritzlew. 2020. "How Bureaucrats May Shape Political Decisions: The Role of Policy Information." *Public Administration*. https://doi.org/10.1111/padm.12709.

Blumenau, Jack. 2019. "The Effects of Female Leadership on Women's Voice in Political Debate." *British Journal of Political Science*, 1–22. https://doi.org/10.1017/s0007123419000334.

———. 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.

Bolton, Alexander, Rachel Augustine Potter, and Sharece Thrower. 2016. "Organizational Capacity, Regulatory Review, and the Limits of Political Control." *Journal of Law, Economics, and Organization* 32 (2): 242–71. https://doi.org/10.1093/jleo/ewv025.

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.

Boyne, George A., Oliver James, Peter John, and Nicolai Petrovsky. 2009. "Democracy and government performance: Holding incumbents accountable in english local governments." *Journal of Politics* 71 (4): 1273–84. https://doi.org/10.1017/S0022381609990089.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory Huber. 2015. "Partisan bias in factual beliefs about politics." *Quarterly Journal of Political Science* 10 (4): 519–78. https://doi.org/10.1561/100.00014074.

Busuioc, E. Madalina. 2009. "Autonomy , Accountability and Control the Case of European Agencies." *European Law Journal* 15 (5): 599–615.

———. 2016. "Friend or foe? inter-agency cooperation, organizational reputation, and turf." *Public Administration* 94 (1): 40–56. https://doi.org/10.1111/padm.12160.

Busuioc, E. Madalina, and Martin Lodge. 2016. "The Reputational Basis of Public Accountability." *Governance* 29 (2): 247–63. https://doi.org/10.1111/gove.12161.

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

Busuioc, E. Madalina, and Dovilė Rimkutė. 2019. "Meeting expectations in the EU regulatory state? Regulatory communications amid conflicting institutional demands." *Journal of European Public Policy*. https://doi.org/10.1080/13501763.2019.1603248.

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

———. 2001b. "The Political Foundations of Bureaucratic Autonomy: A Response to Kernell." *Studies in American Political Development* 15 (1): 113–22.

———. 2002. "Groups, the Media, Agency Waiting Costs, and FDA Drug Approval." *American Journal of Political Science* 46 (3): 490–505. https://doi.org/10.2307/3088394.

———. 2010. *Reputation and Power. Organizational Image and Pharmaceutical Regulation at the FDA*. 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.

Carpenter, Daniel P., and George A. Krause. 2012. "Reputation and public administration." *Public Administration Review* 72 (1): 26–33.

———. 2015. "Transactional authority and bureaucratic politics." *Journal of Public Administration Research and Theory* 25 (1): 5–25. https://doi.org/10.1093/jopart/muu012.

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/10.1177/0951629813518126.

Choi, Jinho D., Joel Tetreault, and Amanda Stent. 2015. "It depends: Dependency parser comparison using a web-based evaluation tool." *ACL-IJCNLP 2015 - 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, Proceedings of the Conference* 1: 387–96. https://doi.org/10.3115/v1/p15-1038.

CNN. 2014. "Bad VA care may have killed more than 1,000 veterans, senator's report says." edition.cnn.com/2014/06/24/us/senator-va-report/index.html.

Coen, David, and Mark Thatcher. 2005. "The new governance of markets and non-majoritarian regulators." *Governance* 18 (3): 329–46. https://doi.org/10.1111/j.1468-0491.2005.00279.x.

Collier, David, and Robert Adcock. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3): 529–46.

Crawford, VP, and J Sobel. 1982. "Strategic Information Transformation." *Econometrica* 50 (6): 1431–51.

Cukierman, Alex, Steven B. Webb, and Bilin Neyapti. 1992. "Measuring the independence of Central Banks and Its Effects on Policy Outcomes." *The World Bank Economic Review* 6 (3): 353–98.

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.

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. 1994. "Administrative Procedures, Information, and Agency Discretion." *American Journal of Political Science* 38 (3): 697–722.

———. 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.

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public Choice* 50 (1-3): 5–25. https://doi.org/10.1007/BF00124924.

Fiorina, Morris P. 1981. "Congressional Control of the Bureaucracy: A Mismatch of Incentives and Capabilities." In *Congress Reconsidered*, edited by Dodd C. Lawrence and Bruce I. Oppenheimer, 2nd ed. Washington, DC: Congressional Quarterly Press.

Firth, John Rupert. 1957. *Studies in linguistic analysis*. Wiley-Blackwell.

Fowler, Anthony. 2020. "Partisan Intoxication or Policy Voting?" *Quarterly Journal of Political Science* 15 (2): 141–79. https://doi.org/10.1561/100.00018027a.

Franchino, Fabio. 2004. "Delegating Powers in the European Community." *British Journal of Political Science* 34 (2): 269–93. https://doi.org/10.1017/S0007123404000055.

FSA. 2009. "The Turner Review. A regulatory response to the global banking crisis." March. papers2://publication/uuid/3CA20D05-3397-4609-92CC-40C8E0AB881E.

Gailmard, Sean. 2009. "Multiple principals and oversight of bureaucratic policy-making." *Journal of Theoretical Politics* 21 (2): 161–86. https://doi.org/10.1177/0951629808100762.

Gailmard, Sean, and John Patty. 2012. "Formal Models of Bureaucracy." *Annual Review of Political Science* 15: 353–77. https://doi.org/10.1146/annurev-polisci-031710-103314.

———. 2013. *Learning While Governing: Expertise and Accountability in the Executive Branch.* Chicago: University of Chicago Press.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word embeddings quantify 100 years of gender and ethnic stereotypes." *Proceedings of the National Academy of Sciences of the United States of America* 115 (16): E3635–44. https://doi.org/10.1073/pnas.1720347115.

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.

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

Gilad, Sharon, Moshe Maor, and Pazit Ben Nun Bloom. 2013. "Organizational reputation, the content of public allegations, and regulatory communication." *Journal of Public Administration Research and Theory* 25 (2): 451–78. https://doi.org/10.1093/jopart/mut041.

Gilad, Sharon, and Tamar Yogev. 2012. "How Reputation Regulates Regulators: Illustrations from the Regulation of Retail Finance." In *The Oxford Handbook of Corporate Reputation*, edited by Timothy G. Pollock and Michael L. Barnett, 1–24. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199596706.013.0016.

Gilardi, Fabrizio. 2002. "Policy credibility and delegation to independent regulatory agencies: A comparative empirical analysis." *Journal of European Public Policy* 9 (6): 873–93. https://doi.org/10.1080/1350176022000046409.

———. 2007. "The Same, But Different: Central Banks, Regulatory Agencies, and the Politics of Delegation to Independent Authorities." *Comparative European Politics* 5: 303–27. https://doi.org/10.1057/palgrave.cep.6110098.

Graham, Matthew H, and Shikhar Singh. 2021. "Partisan Selectivity in Blame Attribution: Evidence from the COVID-19 Pandemic." SocArXiv. https://doi.org/10.31235/osf.io/t8xar.

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.

Hacker, Jacob S., and Paul Pierson. 2006. *Off Center: The Republican Revolution and the Erosion of American Democracy.* New Haven, CT: Yale University Press.

Hammond, Thomas H., and Jack H. Knott. 1996. "Who controls the bureaucracy? Presidential power, congressional dominance, legal constraints, and bureaucratic autonomy in a model of multi-institutional policy-making." *Journal of Law, Economics, and Organization* 12 (1): 119–66. https://doi.org/10.1093/oxfordjournals.jleo.a023355.

Hargrave, Lotte, and Jack Blumenau. 2020. "The Declining Importance of Gender Stereotypes for Politicians' Style in the UK." http://www.jackblumenau.com/papers/gender_styles.pdf.

Healy, Andrew, and Neil Malhotra. 2013. "Retrospective voting reconsidered." *Annual Review of Political Science* 16: 285–306. https://doi.org/10.1146/annurev-polisci-032211-212920.

Honnibal, Matthew, Yoav Goldberg, and Mark Johnson. 2013. "A non-monotonic Arc-Eager transition system for dependency parsing." *CoNLL 2013 - 17th Conference on Computational Natural Language Learning, Proceedings*, 163–72.

Honnibal, Matthew, and Mark Johnson. 2015. "An improved non-monotonic transition system for dependency parsing." *Conference Proceedings - EMNLP 2015: Conference on Empirical Methods in Natural Language Processing*, no. September: 1373–78. https://doi.org/10.18653/v1/d15-1162.

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

House of Commons Defence Committee. 2011. "The Strategic Defence and Security Review and the National Security Strategy." July.

Huber, John D., and Charles R. Shipan. 2002. *Deliberate Discretion? The institutional foundations of bureaucratic autonomy.* New York: Cambridge University Press.

Imai, Kosuke, and In Song Kim. 2019. "When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?" *American Journal of Political Science* 63 (2): 467–90. https://doi.org/10.1111/ajps.12417.

Ingold, Karin, and Philip Leifeld. 2016. "Structural and Institutional Determinants of Influence Reputation: A Comparison of Collaborative and Adversarial Policy Networks in Decision Making and Implementation." *Journal of Public Administration Research and Theory* 26 (1): 1–18. https://doi.org/10.1093/jopart/muu043.

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, 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.

James, Oliver, and Alice Moseley. 2014. "Does performance information about public services affect citizens' perceptions, satisfaction, and voice behaviour? Field experiments with absolute and relative performance information." *Public Administration* 92 (2): 493–511. https://doi.org/10.1111/padm.12066.

Jordana, Jacint, David Levi-Faur, and Xavier Fernández-i-Marín. 2011. "The global diffusion of regulatory agencies: Channels of transfer and stages of diffusion." *Comparative Political Studies* 44 (10): 1343–69. https://doi.org/10.1177/0010414011407466.

Kahan, Dan M., Ellen Peters, Erica Canterell Dawson, and Paul Slovic. 2017. "Motivated numeracy and enlightened self-government." *Behavioural Public Policy* 1 (1): 54–86. https://doi.org/10.1017/bpp.2016.2.

Keeper, Philip, and David Stasavage. 2003. "The Limits of Delegation: Veto Players, Central Bank Independence, and the Credibility of Monetary Policy." *American Political Science Review* 97 (3): 407–23. https://doi.org/10.1017/S0003055403000777.

Koop, Christel, and Chris Hanretty. 2018. "Political Independence, Accountability, and the Quality of Regulatory Decision-Making." *Comparative Political Studies* 51 (1): 38–75. https://doi.org/10.1177/0010414017695329.

Krause, George A. 1996. "The Institutional Dynamics of Policy Administration : Bureaucratic Influence over Securities Regulation." *American Journal of Political Science* 40 (4): 1083–1121.

———. 1999. *A Two-Way Street: The Institutional Dynamics of the Modern Administrative State*. Pittsburgh, PA: University of Pittsburgh Press.

Krause, George A., and J. Kevin Corder. 2007. "Explaining bureaucratic optimism: Theory and evidence from U.S. executive agency macroeconomic forecasts." *American Political Science Review* 101 (1): 129–42. https://doi.org/10.1017/S0003055407070074.

Krause, George A., and James W. Douglas. 2005. "Institutional design versus reputational effects on bureaucratic performance: Evidence from U.S. government macroeconomic and fiscal projections." *Journal of Public Administration Research and Theory* 15 (2): 281–306. https://doi.org/10.1093/jopart/mui038.

Kroeger, Mary A. 2020. "Bureaucrats as Lobbyists."

Lee, Soo Young, and Andrew B. Whitford. 2013. "Assessing the effects of organizational resources on public agency performance: Evidence from the US federal government." *Journal of Public Administration Research and Theory* 23 (3): 687–712. https://doi.org/10.1093/jopart/mus050.

Levi-Faur, David. 2011. "Regulatory networks and regulatory agencification: Towards a Single European regulatory space." *Journal of European Public Policy* 18 (6): 810–29. https://doi.org/10.1080/13501763.2011.593309.

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. 2021. "Directional motives and different priors are observationally equivalent." http://andrewtlittle.com/wp-content/uploads/2021/04/little_motives_priors.pdf.

Little, Andrew, Keith Schnakenberg, and Ian R. 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.

Maggetti, Martino. 2009. "The role of independent regulatory agencies in policy-making: A comparative analysis." *Journal of European Public Policy* 16 (3): 450–70. https://doi.org/10.1080/13501760802662854.

———. 2014. "The rewards of cooperation: The effects of membership in European regulatory networks." *European Journal of Political Research* 53 (3): 480–99. https://doi.org/10.1111/1475-6765.12042.

Majone, Giandomenico. 1994. "The Rise of the Regulatory State in Europe." *West European Politics* 17 (3): 77–101. https://doi.org/10.1080/01402389408425031.

Malhotra, Neil, and Alexander G. Kuo. 2008. "Attributing blame: The public's response to Hurricane Katrina." *Journal of Politics* 70 (1): 120–35. https://doi.org/10.1017/S0022381607080097.

Maor, Moshe. 2007. "A scientific standard and an agency's legal independence: Which of these reputation protection mechanisms is less susceptible to political moves?" *Public Administration* 85 (4): 961–78. https://doi.org/10.1111/j.1467-9299.2007.00676.x.

Maor, Moshe, Sharon Gilad, and Pazit Ben Nun Bloom. 2013. "Organizational reputation, regulatory talk, and strategic silence." *Journal of Public Administration Research and Theory* 23 (3): 581–608. https://doi.org/10.1093/jopart/mus047.

Maor, Moshe, and Raanan Sulitzeanu-Kenan. 2013. "The Effect of Salient Reputational Threats on the Pace of FDA Enforcement." *Governance* 26 (1): 31–61. https://doi.org/10.1111/j.1468-0491.2012.01601.x.

———. 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.

Maor, Moshe, and Arild Wæraas. 2015. "Understanding organizational reputation in a public sector context." In *Organizational Reputation in the Public Sector*, edited by Moshe Maor and Arild Wæraas, 1st ed. New York: Routledge. https://doi.org/https://doi-org.libproxy.ucl.ac.uk/10.4324/9781315850825.

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.

Mccubbins, Mathew D. 1985. "The Legislative Design of Regulatory Structure." *American Journal of Political Science* 29 (4): 721–48. https://www.jstor.org/stable/2111178.

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.

Meyer, Robinson. 2017. “Trump’s EPA Chief Denies the Basic Science of Climate Change.”

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.

Miller, Susan M., and Alexander I. Ruder. 2020. “Holding agencies accountable: Exploring the effect of oversight on citizens’ approval of members of Congress.” *Journal of Public Policy* 40 (4): 672–93. https://doi.org/10.1017/S0143814X19000151.

Ministry of Defence. 2010. “Securing Britain in an Age of Uncertainty: The Strategic Defence and Security Review Securing Britain in an.”

———. 2011. “Defence Reform. An independent report into structure and management of the Ministry of Defence.” June. https://doi.org/10.1080/1356788000642.

Moe, Terry M. 1984. “The New Economics of Organization.” *American Journal of Political Science* 28 (4): 739–77.

———. 1990. “Political Institutions: The Neglected Side of the Story.” *Journal of Law, Economics, & Organization* 6: 213–53.

———. 2006. “Political control and the power of the agent.” *Journal of Law, Economics, and Organization* 22 (1): 1–29. https://doi.org/10.1093/jleo/ewj011.

———. 2012. “Delegation, Control, and the Study of Public Bureaucracy.” *The Forum* 10 (2). https://doi.org/10.1515/1540-8884.1508.

Moe, Terry M., and Scott Wilson. 1994. “Presidents and the Politics of Structure.” *Law and Contemporary Problems* 57 (2): 1–44. https://doi.org/10.2307/1141049.

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.

Nelson, David, and Susan Webb Yackee. 2012. “Lobbying coalitions and government policy change: An analysis of federal agency rulemaking.” *Journal of Politics* 74 (2): 339–53. https://doi.org/10.1017/S0022381611001599.

Nicholson-Crotty, Jill, and Susan M. Miller. 2012. “Bureaucratic effectiveness and influence in the legislature.” *Journal of Public Administration Research and Theory* 22 (2): 347–71. https://doi.org/10.1093/jopart/mur054.

Niskanen, WA. 1971. *Bureaucracy and Public Economics.* Chicago: Aldine-Atherton.

Office of Inspector General (DHS). 2015. “Evaluation of DHS ’ Information Security Program for Fiscal Year 2015.” oig.dhs.gov/assets/Mgmt/2016/OIG-16-08-Nov15.pdf.

Page, E. 2012. *Policy Without Politicians. Bureaucratic Influence in Comparative Perspective.* Oxford: Oxford University Press.

Pavón Mediano, Andrés. 2018. "Agencies' formal independence and credible commitment in the Latin American regulatory state: A comparative analysis of 8 countries and 13 sectors." *Regulation & Governance*, no. December 2017. https://doi.org/10.1111/rego.12187.

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.

Pew Research Center. 2019. "Public Expresses Favourable Views of a Number of Federal Agencies." https://doi.org/.1037//0033-2909.I26.1.78.

Potter, Rachel Augustine. 2017. "Slow-rolling, fast-Tracking, and the pace of bureaucratic decisions in rulemaking." *Journal of Politics* 79 (3): 841–55. https://doi.org/10.1086/690614.

———. 2019. *Bending the Rules. Pocedural Politicking in the Bureaucracy.* Chicago: University of Chicago Press.

Potter, Rachel Augustine, and Kenneth Lowande. 2020. "Congressional Oversight Revisited: Politics and Procedure in Agency Rulemaking." *The Journal of Politics*. https://doi.org/10.1086/709436.

Preotiuc-Pietro, Daniel, Daniel J. Hopkins, Ye Liu, and Lyle Ungar. 2017. "Beyond binary labels: Political ideology prediction of twitter users." *ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers)* 1: 729–40. https://doi.org/10.18653/v1/P17-1068.

Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R Radev. 2010. "How to Analyze Political Attention.pdf." *American Journal ofPolitical Science* 54 (1): 209–28.

Rheault, Ludovic, and Christopher Cochrane. 2019. "Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora." *Political Analysis*, 1–22. https://doi.org/10.1017/pan.2019.26.

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.

Rodriguez, Pedro L., and Arthur Spirling. 2022. "Word embeddings: What works, what doesn't, and how to tell the difference for applied research." *Journal of Politics* 84 (1922658): 101–15. https://doi.org/10.1086/715162.

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. 2015b. "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.

———. 2015a. "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.

Sirin, Cigdem V., and José D. Villalobos. 2011. "Where Does the Buck Stop? Applying Attribution Theory to Examine Public Appraisals of the President." *Presidential Studies Quarterly* 41 (2): 334–57. https://doi.org/10.1111/j.1741-5705.2011.03857.x.

Spenkuch, J. L., E. Teso, and G. Xu. 2021. "Ideology and Performance in Public Organizations (No. w28673)." http://www.nber.org/papers/w28673.

Spirling, Arthur. 2016. "Democratization and Linguistic Complexity: The Effect of Franchise Extension on Parliamentary Discourse, 1832-1915." *Journal of Politics* 78 (1): 120–36. https://doi.org/10.1086/683612.

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

Thatcher, Mark, and Alec Stone Sweet. 2002. "Theory and practice of delegation to non-majoritarian institutions." *West European Politics* 25 (1): 1–22. https://doi.org/10.4324/9780203005989.

The Daily Telegraph. 2008. "FSA staff will get bonuses despite overseeing financial crisis." 3. telegraph.co.uk/news/3374017/FSA-staff-will-get-bonuses-despite-overseeing-financial-crisis.html.

The Guardian. 2009. "Jacqui Smith goes - hit by porn claims, but fatally damaged by 42 days." theguardian.com/politics/2009/jun/02/jacqui-smith-resignation.

The Times. 2018. "Home Office knew Windrush scandal was brewing for years." thetimes.co.uk/article/home-office-knew-windrush-scandal-was-brewing-for-years-ws08ghmxx.

The Washington Post. 2008a. "DHS Withdraws Bid to Curb Union Rights." washingtonpost.com/wp-dyn/content/article/2008/02/19/AR2008021902459_pf.html.

———. 2008b. washingtonpost.com/wp-dyn/content/article/2008/09/16/AR2008091603200_pf.html.

Thomson, Robert, and René Torenvlied. 2011. "Information, Commitment and Consensus: A Comparison of Three Perspectives on Delegation in the European Union." *British Journal of Political Science* 41 (1): 139–59. https://doi.org/10.1017/S0007123410000268.

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.

Tucker, Paul. 2018. *Unelected power: the quest for legitimacy in central banking and the regulatory state.* Oxford: Princeton University Press.

Van Atteveldt, Wouter, Tamir Sheafer, Shaul R Shenhav, and Yair Fogel-Dror. 2017. "Clause Analysis: Using Syntactic Information to Automatically Extract Source, Subject, and Predicate from Texts with an Application to the 2008-2009 Gaza War." *Political Analysis* 25: 207–22. https://doi.org/10.1017/pan.2016.12.

Vannoni, Matia, Elliott Ash, and Massimo Morelli. 2021. "Measuring Discretion and Delegation in Legislative Texts: Methods and Application to US States." *Political Analysis* 29 (1): 43–57. https://doi.org/10.1017/pan.2020.9.

Volden, Craig, and Alan Wiseman. 2020. "Centre for Effective Lawmaking." https://thelawmakers.org/data-download.

Weber, Max. 1922. *Economy and Society.* Edited by Roth Guenther and Wittich Claus. 1968th ed. Berkeley: University of California Press.

Weingast, Barry R, and Mark J Moran. 1983. "Bureaucratic Discretion or Congressional Control ? Regulatory Policymaking by the Federal Trade Commission Author." *Journal of Political Economy* 91 (5): 765–800.

Wilson, J. Q. 1989. *Bureaucracy: What government agencies do and why they do it.* New York: Basic Books.

Wood, Matthew, Sjors Overman, and E. Madalina Busuioc. 2020. "A Multidimensional Reputation Barometer for Public Agencies: A Validated Instrument." *Public Administration Review*, 1–11. https://doi.org/10.1111/puar.13158.

Workman, Samuel. 2015. *The dynamics of bureaucracy in the US government: How Congress and federal agencies process information and solve problems.* Cambridge: Cambridge University Press.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511818691.

Appendix A

Appendix to Chapter 1

A.1 Sample of Agencies

In the following tables, I report all the agencies for which I produce reputation 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 reputation.

A.1.1 Sample of US Agencies

Agecy Type	**N. Agencies**	**Obs.**	**Avg. Mentions**	**Avg. Reputation**
Executive Department	18	590	567.97	0.61
Executive Office of the President	5	141	166.43	0.42
Executive Sub-Agencies	147	2,515	124.19	0.51
Federal Partnership	1	5	7.40	0.33
Government-owned Corporation	4	63	55.02	0.48
Independent Agency	68	1,434	183.35	0.51
Non-for-profit Public Organisation	3	51	52.76	0.46
Unclear	2	7	9.71	0.46

Table A.1: US descriptive statistics by agency type.

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Administration for Children and Families	1	6.00	0.42	Executive Sub-Agencies
Administrative Conference of the United States	15	9.07	0.41	Independent Agency
Advisory Council on Historic Preservation	6	12.83	0.41	Independent Agency
Agency for Healthcare Research and Quality	6	8.50	0.53	Executive Sub-Agencies
Agency for International Development	36	109.72	0.55	Independent Agency
Agency for Toxic Substances and Disease Registry	11	15.36	0.47	Executive Sub-Agencies
Agricultural Marketing Service	6	7.83	0.37	Executive Sub-Agencies
Agricultural Research Service	30	25.17	0.51	Executive Sub-Agencies
Air Force	36	1,038.31	0.75	Executive Sub-Agencies

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Alcohol and Tobacco Tax and Trade Bureau	7	11.86	0.39	Executive Sub-Agencies
American Battle Monuments Commission	16	13.44	0.42	Executive Sub-Agencies
AMTRAK	19	8.32	0.42	Independent Agency
Animal and Plant Health Inspection Service	28	25.32	0.55	Executive Sub-Agencies
Appalachian Regional Commission	35	30.29	0.52	Executive Sub-Agencies
Benefits Review Board	2	14.00	0.47	Executive Sub-Agencies
Board of Veterans Appeals	20	22.35	0.40	Executive Sub-Agencies
Bonneville Power Administration	26	20.46	0.44	Executive Sub-Agencies
Border and Transportation Security Directorate	1	5.00	0.47	Executive Sub-Agencies
Broadcasting Board of Governors	14	13.50	0.44	Independent Agency
Bureau of Alcohol, Tobacco, Firearms, and Explosives	31	63.26	0.52	Executive Sub-Agencies
Bureau of Competition	3	9.67	0.35	Independent Agency
Bureau of Economic Analysis	30	7.27	0.47	Executive Sub-Agencies
Bureau of Engraving and Printing	7	7.00	0.38	Executive Sub-Agencies
Bureau of Indian Affairs	35	56.20	0.47	Executive Sub-Agencies
Bureau of Indian Education	3	9.33	0.53	Executive Sub-Agencies
Bureau of Industry and Security	2	5.50	0.23	Executive Sub-Agencies
Bureau of International Labor Affairs	1	11.00	0.48	Executive Sub-Agencies
Bureau of Labor Statistics	36	45.94	0.48	Executive Sub-Agencies
Bureau of Land Management	36	104.92	0.53	Executive Sub-Agencies
Bureau of Ocean Energy Management	6	11.33	0.44	Executive Sub-Agencies
Bureau of Political-Military Affairs	1	6.00	0.42	Executive Sub-Agencies
Bureau of Prisons	35	34.91	0.45	Executive Sub-Agencies
Bureau of Reclamation	36	102.94	0.57	Executive Sub-Agencies
Bureau of Safety and Environmental Enforcement	1	15.00	0.81	Executive Sub-Agencies
Bureau of the Census	36	111.61	0.40	Executive Sub-Agencies
Bureau of the Public Debt	6	8.33	0.42	Executive Sub-Agencies
Centers for Disease Control and Prevention	24	56.08	0.47	Executive Sub-Agencies
Centers for Medicare and Medicaid Services	19	164.47	0.49	Executive Sub-Agencies
Central Intelligence Agency	36	532.47	0.59	Independent Agency
Chemical Safety and Hazard Investigation Board	4	7.25	0.46	Independent Agency
Citizen and Immigration Services	10	31.00	0.51	Executive Sub-Agencies
Civil Rights Division	27	23.37	0.47	Executive Sub-Agencies

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Commission on Civil Rights	31	17.58	0.45	Independent Agency
Commodities Futures Trading Commission	30	121.60	0.55	Independent Agency
Commodity Credit Corporation	27	77.37	0.47	Government-owned Corporation
Community Development Financial Institutions Fund	10	7.50	0.38	Executive Sub-Agencies
Consumer Financial Protection Bureau	7	294.57	0.49	Independent Agency
Consumer Product Safety Commission	35	82.89	0.50	Independent Agency
Corporation for National and Community Service	12	16.33	0.44	Independent Agency
Corporation for National Community Service	3	8.00	0.52	Independent Agency
Corporation for Public Broadcasting	28	78.93	0.46	Non-for-profit Public Organisation
Council of Economic Advisers	35	29.71	0.40	Executive Office of the President
Council on Environmental Quality	23	15.83	0.42	Executive Sub-Agencies
Court Services and Offender Supervision Agency	4	12.25	0.38	Executive Sub-Agencies
Customs and Border Protection	14	65.36	0.61	Executive Sub-Agencies
Defense Acquisition Regulations System	2	6.00	0.38	Executive Sub-Agencies
Defense Advanced Research Projects Agency	31	22.81	0.56	Executive Sub-Agencies
Defense Contract Audit Agency	17	21.18	0.41	Executive Sub-Agencies
Defense Contract Management Agency	4	7.50	0.36	Executive Sub-Agencies
Defense Finance and Accounting Service	9	21.89	0.44	Executive Sub-Agencies
Defense Information Systems Agency	1	9.00	0.63	Executive Sub-Agencies
Defense Intelligence Agency	36	37.33	0.51	Executive Sub-Agencies
Defense Logistics Agency	23	14.04	0.45	Executive Sub-Agencies
Defense Media Activity	5	10.00	0.48	Executive Sub-Agencies
Defense Nuclear Facilities Safety Board	4	7.50	0.36	Independent Agency
Defense Security Service	4	10.00	0.40	Executive Sub-Agencies
Defense Technology Security Administration	2	6.50	0.47	Executive Sub-Agencies
Delta Regional Authority	5	7.40	0.33	Federal Partnership
Department of Agriculture	36	356.19	0.59	Executive Department
Department of Commerce	36	308.39	0.61	Executive Department
Department of Defense	36	1,507.64	0.73	Executive Department
Department of Defense Education Activity	1	5.00	0.35	Executive Department
Department of Education	36	291.83	0.55	Executive Department
Department of Energy	36	567.97	0.63	Executive Department
Department of Health and Human Services	36	361.19	0.58	Executive Department

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Department of Homeland Security	16	1,581.62	0.82	Executive Department
Department of Housing and Urban Development	36	74.44	0.52	Executive Department
Department of Justice	36	1,054.39	0.66	Executive Department
Department of Labor	36	297.00	0.54	Executive Department
Department of State	36	166.64	0.55	Executive Department
Department of the Army	33	24.91	0.50	Executive Department
Department of the Interior	36	290.47	0.57	Executive Department
Department of the Navy	36	994.31	0.75	Executive Department
Department of the Treasury	36	1,486.94	0.50	Executive Department
Department of Transportation	36	288.31	0.58	Executive Department
Department of Veterans Affairs	36	536.75	0.75	Executive Department
Directorate of Defense Trade Controls	3	6.67	0.57	Executive Sub-Agencies
Domestic Nuclear Detection Office	7	17.14	0.55	Executive Sub-Agencies
Drug Enforcement Administration	36	158.92	0.56	Executive Sub-Agencies
Economic Development Administration	33	175.48	0.56	Executive Sub-Agencies
Economic Research Service	18	7.83	0.41	Executive Sub-Agencies
Election Assistance Commission	12	33.67	0.48	Executive Sub-Agencies
Employment and Training Administration	3	6.33	0.50	Executive Sub-Agencies
Employment Standards Administration	1	8.00	0.58	Executive Sub-Agencies
Environmental Protection Agency	36	1,514.19	0.61	Independent Agency
Equal Employment Opportunity Commission	36	73.72	0.50	Independent Agency
Executive Office for Immigration Review	2	7.50	0.35	Executive Sub-Agencies
Export-Import Bank of the United States	15	9.67	0.46	Independent Agency
Farm Credit Administration	18	31.83	0.44	Independent Agency
Farm Service Agency	21	17.71	0.50	Executive Sub-Agencies
Federal Agricultural Mortgage Corporation	8	44.00	0.53	Government-owned Corporation
Federal Aviation Administration	36	87.19	0.54	Executive Sub-Agencies
Federal Bureau of Investigation	36	708.36	0.62	Executive Sub-Agencies
Federal Communications Commission	36	379.06	0.55	Independent Agency
Federal Deposit Insurance Corporation	36	165.72	0.52	Independent Agency
Federal Election Commission	36	117.78	0.48	Independent Agency
Federal Emergency Management Agency	36	363.19	0.52	Executive Sub-Agencies
Federal Energy Regulatory Commission	36	210.89	0.48	Executive Sub-Agencies

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

Agecy Name	N. of Observations	Avg. N. Mentions	Avg. Reputation	Agency Type
Federal Highway Administration	35	50.37	0.49	Executive Sub-Agencies
Federal Housing Administration	36	193.03	0.50	Executive Sub-Agencies
Federal Housing Finance Agency	8	21.25	0.43	Independent Agency
Federal Labor Relations Authority	11	12.36	0.38	Independent Agency
Federal Law Enforcement Training Center	20	19.80	0.55	Executive Sub-Agencies
Federal Maritime Commission	29	40.45	0.52	Independent Agency
Federal Mediation and Conciliation Service	8	10.50	0.41	Independent Agency
Federal Motor Carrier Safety Administration	15	15.53	0.48	Executive Sub-Agencies
Federal Prison Industries	13	53.77	0.47	Government-owned Corporation
Federal Railroad Administration	34	39.85	0.49	Executive Sub-Agencies
Federal Reserve	36	514.78	0.45	Independent Agency
Federal Retirement Thrift Investment Board	3	8.00	0.38	Independent Agency
Federal Student Aid	3	10.67	0.37	Executive Sub-Agencies
Federal Trade Commission	36	237.36	0.57	Independent Agency
Federal Transit Administration	23	12.96	0.48	Executive Sub-Agencies
Financial Crimes Enforcement Network	16	11.81	0.42	Executive Sub-Agencies
Financial Management Service	19	19.89	0.40	Executive Sub-Agencies
Financial Stability Oversight Council	6	43.50	0.52	Executive Sub-Agencies
Fish and Wildlife Service	36	141.31	0.59	Executive Sub-Agencies
Food and Drug Administration	36	705.72	0.66	Executive Sub-Agencies
Food and Nutrition Service	21	12.00	0.49	Executive Sub-Agencies
Food Safety and Inspection Service	21	10.29	0.42	Executive Sub-Agencies
Foreign Agricultural Service	16	9.62	0.50	Executive Sub-Agencies
Foreign Claims Settlement Commission	5	11.20	0.52	Executive Sub-Agencies
Forest Service	36	461.39	0.60	Executive Sub-Agencies
General Services Administration	36	174.44	0.57	Independent Agency
Geological Survey	36	54.64	0.57	Executive Sub-Agencies
Government National Mortgage Association	15	21.73	0.45	Government-owned Corporation
Grain Inspection, Packers, and Stockyards Administration	6	22.67	0.39	Executive Sub-Agencies
Health Resources and Services Administration	29	15.69	0.48	Executive Sub-Agencies
Housing Finance Agency	12	7.33	0.37	Independent Agency
Immigration and Customs Enforcement	22	130.09	0.50	Executive Sub-Agencies
Independent Payment Advisory Board	6	160.67	0.48	Independent Agency

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

Agecy Name	N. of Observations	Avg. N. Mentions	Avg. Reputation	Agency Type
Indian Health Service	36	54.83	0.48	Executive Sub-Agencies
Institute of Peace	15	19.27	0.55	Independent Agency
Inter-American Foundation	14	11.43	0.46	Independent Agency
Internal Revenue Service	36	1,058.50	0.49	Executive Sub-Agencies
International Boundary and Water Commission	15	9.47	0.46	Executive Sub-Agencies
International Trade Administration	24	12.79	0.45	Executive Sub-Agencies
International Trade Commission	35	121.77	0.47	Independent Agency
Maritime Administration	32	27.19	0.48	Executive Sub-Agencies
Marshals Service	36	24.81	0.50	Executive Sub-Agencies
Merit Systems Protection Board	31	24.87	0.39	Independent Agency
Metropolitan Washington Airport Authority	1	9.00	0.46	Independent Agency
Millennium Challenge Corporation	8	16.50	0.43	Independent Agency
Minority Business Development Agency	8	19.88	0.49	Executive Sub-Agencies
Missile Defense Agency	12	26.42	0.51	Executive Sub-Agencies
Mississippi River Commission	1	12.00	0.61	Executive Sub-Agencies
National Aeronautics and Space Administration	36	549.78	0.73	Independent Agency
National Archives and Records Administration	13	7.38	0.42	Independent Agency
National Capital Planning Commission	14	16.07	0.49	Independent Agency
National Cemetery Administration	9	6.22	0.50	Executive Sub-Agencies
National Consumer Cooperative Bank	5	11.60	0.48	Unclear
National Council on Disability	6	8.33	0.38	Independent Agency
National Credit Union Administration	21	22.86	0.44	Independent Agency
National Geospatial-Intelligence Agency	23	14.96	0.44	Executive Sub-Agencies
National Highway Traffic Safety Administration	36	55.14	0.52	Executive Sub-Agencies
National Indian Gaming Commission	9	8.78	0.48	Executive Sub-Agencies
National Institute of Building Sciences	2	5.00	0.40	Unclear
National Institute of Food and Agriculture	2	5.00	0.45	Executive Sub-Agencies
National Institute of Standards and Technology	29	88.83	0.60	Executive Sub-Agencies
National Institute on Disability and Rehabilitation Research	3	7.33	0.46	Executive Sub-Agencies
National Institutes of Health	36	191.25	0.63	Executive Sub-Agencies
National Labor Relations Board	36	110.47	0.43	Independent Agency
National Mediation Board	11	33.45	0.39	Independent Agency
National Nuclear Security Administration	16	42.25	0.51	Executive Sub-Agencies
National Oceanic and Atmospheric Administration	36	46.08	0.58	Executive Sub-Agencies

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

Agecy Name	N. of Observations	Avg. N. Mentions	Avg. Reputation	Agency Type
National Park Service	36	248.75	0.67	Executive Sub-Agencies
National Reconnaissance Office	18	35.00	0.48	Executive Sub-Agencies
National Science Foundation	36	243.19	0.65	Independent Agency
National Security Agency	36	84.53	0.53	Executive Sub-Agencies
National Technical Information Service	9	18.11	0.46	Executive Sub-Agencies
National Telecommunications and Information Administration	31	24.81	0.49	Executive Sub-Agencies
National Transportation Safety Board	36	65.22	0.50	Independent Agency
Natural Resources Conservation Service	19	19.11	0.49	Executive Sub-Agencies
Nuclear Regulatory Commission	36	182.56	0.54	Independent Agency
Occupational Safety and Health Administration	35	228.06	0.49	Executive Sub-Agencies
Occupational Safety and Health Review Commission	2	42.50	0.63	Independent Agency
Office of Acquisition Policy	1	5.00	0.48	Executive Sub-Agencies
Office of Economic Adjustment	8	14.50	0.47	Executive Sub-Agencies
Office of Electricity Delivery and Energy Reliability	1	5.00	0.30	Executive Sub-Agencies
Office of Energy Efficiency and Renewable Energy	6	23.17	0.44	Executive Sub-Agencies
Office of Federal Procurement Policy	14	18.64	0.43	Executive Office of the President
Office of Foreign Assets Control	16	20.31	0.47	Executive Sub-Agencies
Office of Government Ethics	26	32.23	0.46	Independent Agency
Office of Health, Safety, and Security	1	5.00	0.46	Executive Sub-Agencies
Office of Justice Programs	10	14.40	0.46	Executive Sub-Agencies
Office of Labor-Management Standards	2	51.50	0.61	Executive Sub-Agencies
Office of Management and Budget	36	572.86	0.48	Executive Office of the President
Office of Minority Economic Impact	2	8.00	0.34	Executive Sub-Agencies
Office of National Drug Control Policy	24	47.71	0.52	Executive Office of the President
Office of Personnel Management	36	111.06	0.52	Independent Agency
Office of Science and Technology	32	12.44	0.30	Executive Office of the President
Office of Special Counsel	25	21.04	0.51	Independent Agency
Office of Special Education and Rehabilitative Services	2	5.50	0.64	Executive Sub-Agencies
Office of Special Trustee for American Indians	1	5.00	0.42	Executive Sub-Agencies
Office of Surface Mining, Reclamation and Enforcement	2	9.00	0.37	Executive Sub-Agencies
Office of the Comptroller of the Currency	15	9.93	0.44	Executive Sub-Agencies
Office of the Director of National Intelligence	10	11.00	0.51	Independent Agency
Office of Thrift Supervision	20	44.05	0.46	Executive Sub-Agencies

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Office of Vocational and Adult Education	1	8.00	0.57	Executive Sub-Agencies
Office of Workers' Compensation Programs	1	5.00	0.59	Executive Sub-Agencies
Parole Commission	13	13.85	0.45	Executive Sub-Agencies
Patent and Trademark Office	34	68.59	0.48	Executive Sub-Agencies
Peace Corps	36	192.94	0.63	Independent Agency
Pension Benefit Guaranty Corporation	35	56.11	0.42	Independent Agency
Pentagon	36	568.36	0.45	Executive Sub-Agencies
Pipeline and Hazardous Materials Safety Administration	10	25.80	0.52	Executive Sub-Agencies
Postal Regulatory Commission	6	13.00	0.44	Independent Agency
Privacy and Civil Liberties Oversight Board	7	15.29	0.53	Independent Agency
Public and Indian Housing	1	5.00	0.52	Executive Sub-Agencies
Public Buildings Service	5	11.00	0.51	Independent Agency
Public Health Service	36	77.36	0.56	Executive Sub-Agencies
Railroad Retirement Board	17	14.65	0.37	Independent Agency
Rehabilitation Services Administration	5	8.80	0.36	Executive Sub-Agencies
Research and Innovative Technology Administration	1	6.00	0.32	Executive Sub-Agencies
Risk Management Agency	9	14.56	0.41	Executive Sub-Agencies
Rural Housing Service	4	9.25	0.34	Executive Sub-Agencies
Rural Utilities Service	7	11.57	0.45	Executive Sub-Agencies
Saint Lawrence Seaway Development Corporation	3	7.67	0.36	Executive Sub-Agencies
Securities and Exchange Commission	36	410.31	0.68	Independent Agency
Securities Investor Protection Corporation	10	27.60	0.45	Non-for-profit Public Organisation
Selective Service System	17	22.82	0.37	Independent Agency
Small Business Administration	36	424.47	0.64	Independent Agency
Social Security Administration	36	113.72	0.49	Independent Agency
Social Security Advisory Board	2	10.50	0.43	Independent Agency
State Justice Institute	13	15.77	0.47	Non-for-profit Public Organisation
Substance Abuse and Mental Health Services Administration	23	22.04	0.50	Executive Sub-Agencies
Surface Transportation Board	18	17.28	0.46	Independent Agency
Tennessee Valley Authority	35	30.03	0.48	Independent Agency
Trade and Development Agency	11	16.00	0.51	Independent Agency
Transportation Security Administration	17	253.00	0.65	Executive Sub-Agencies
Tricare Management Activity	22	84.18	0.53	Executive Sub-Agencies

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

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
US Postal Service	36	353.03	0.58	Independent Agency
Veterans Benefits Administration	17	16.88	0.49	Executive Sub-Agencies
Veterans Employment and Training Service	12	10.58	0.45	Executive Sub-Agencies
Veterans Health Administration	26	21.77	0.45	Executive Sub-Agencies
Wage and Hour Division	5	10.80	0.43	Executive Sub-Agencies
Western Area Power Administration	13	16.46	0.42	Executive Sub-Agencies
Womens' Bureau	6	8.83	0.40	Executive Sub-Agencies

A.1.2 Sample of UK Agencies

Agecy Type	N. Agencies	Obs.	Avg. Mentions	Avg. Reputation
Government Dept.	23	487	432.98	0.59
Non-dept. Agency	194	1,789	64.19	0.45

Table A.3: UK descriptive statistics by agency type.

Table A.4: Sample of UK agencies and descriptive statistics.

Agecy Name	N. of Observations	Avg. N. Mentions	Avg. Reputation	Agency Type
Advisory Committee on Business Appointments	6	16.33	0.51	Non-dept. Agency
Advisory Committee on Novel Foods and Processes	7	7.14	0.33	Non-dept. Agency
Advisory Committee on Releases to the Environment	6	7.17	0.28	Non-dept. Agency
Advisory Council on the Misuse of Drugs	23	23.87	0.41	Non-dept. Agency
Advisory, Conciliation and Arbitration Service	37	56.89	0.44	Non-dept. Agency
Agriculture and Horticulture Development Board	2	5.00	0.46	Non-dept. Agency
Air Accidents Investigation Branch	9	7.89	0.47	Non-dept. Agency
Animal and Plant Health Agency	4	15.25	0.50	Non-dept. Agency
Armed Forces' Pay Review Body	13	13.69	0.40	Non-dept. Agency
Arts Council England	24	15.67	0.44	Non-dept. Agency
Arts Council of Wales	4	17.75	0.37	Non-dept. Agency
Atomic Energy Authority	25	19.48	0.52	Non-dept. Agency
Attorney General's Office	1	5.00	0.33	Non-dept. Agency
Bank of England	40	293.55	0.52	Non-dept. Agency
Biometrics Commissioner	1	7.00	0.63	Non-dept. Agency
Biotechnology and Biological Sciences Research Council	7	10.71	0.46	Non-dept. Agency
Boundary Commission for England	11	9.64	0.36	Non-dept. Agency
Boundary Commission for Scotland	4	6.00	0.41	Non-dept. Agency
Boundary Commission for Wales	3	9.33	0.43	Non-dept. Agency
British Business Bank	5	26.80	0.61	Non-dept. Agency
British Council	40	61.08	0.60	Non-dept. Agency
British Film Institute	14	8.50	0.38	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
British Transport Police Authority	1	6.00	0.45	Non-dept. Agency
Broads Authority	8	35.75	0.43	Non-dept. Agency
Building Regulations Advisory Committee	1	9.00	0.45	Non-dept. Agency
Cabinet Office	40	142.60	0.59	Government Dept.
Care Quality Commission	16	237.19	0.48	Non-dept. Agency
Central Arbitration Committee	5	12.60	0.32	Non-dept. Agency
Centre for Data Ethics and Innovation	1	10.00	0.49	Non-dept. Agency
Centre for Environment, Fisheries and Aquaculture Science	8	10.00	0.40	Non-dept. Agency
Centre for the Protection of National Infrastructure	1	7.00	0.68	Non-dept. Agency
Certification Officer	5	10.80	0.28	Non-dept. Agency
Children and Family Court Advisory and Support Service	12	9.33	0.42	Non-dept. Agency
Civil Justice Council	2	10.00	0.60	Non-dept. Agency
Civil Service Commission	6	10.17	0.35	Non-dept. Agency
Coal Authority	15	41.60	0.44	Non-dept. Agency
College of Policing	7	48.43	0.64	Non-dept. Agency
Commission on Human Medicines	2	11.00	0.35	Non-dept. Agency
Commissioner for Public Appointments	12	14.58	0.37	Non-dept. Agency
Committee on Climate Change	12	79.08	0.44	Non-dept. Agency
Committee on Radioactive Waste Management	3	13.67	0.24	Non-dept. Agency
Committee on Standards in Public Life	24	31.21	0.46	Non-dept. Agency
C. on Toxicity of Chemicals in Food, Consumer Prod. and the Env.	4	7.75	0.36	Non-dept. Agency
Companies House	28	19.93	0.43	Non-dept. Agency
Competition Appeal Tribunal	2	14.00	0.35	Non-dept. Agency
Construction Industry Training Board	20	30.70	0.43	Non-dept. Agency
Consumer Council for Water	6	10.50	0.43	Non-dept. Agency
Copyright Tribunal	2	5.50	0.43	Non-dept. Agency
Council for Science and Technology	1	10.00	0.34	Non-dept. Agency
Courts and Tribunals Service	9	25.11	0.51	Non-dept. Agency
Criminal Cases Review Commission	21	24.24	0.47	Non-dept. Agency
Criminal Injuries Compensation Authority	9	8.78	0.37	Non-dept. Agency
Crown Commercial Service	1	7.00	0.39	Non-dept. Agency
Crown Prosecution Service Inspectorate	1	12.00	0.63	Non-dept. Agency
Debt Management Office	2	11.50	0.43	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Defence Science and Technology Laboratory	11	10.00	0.41	Non-dept. Agency
Department for Business, Energy and Industrial Strategy	4	51.25	0.49	Government Dept.
Department for Digital, Culture, Media and Sport	21	40.67	0.47	Government Dept.
Department for Education	40	197.57	0.59	Government Dept.
Department for Environment Food and Rural Affairs	19	158.68	0.49	Government Dept.
Department for Exiting the European Union	4	89.75	0.49	Government Dept.
Department for International Development	23	161.43	0.61	Government Dept.
Department for International Trade	4	120.75	0.61	Government Dept.
Department for Transport	19	345.79	0.61	Government Dept.
Department for Work and Pensions	19	318.79	0.51	Government Dept.
Department of Finance for Northern Ireland	2	5.00	0.47	Government Dept.
Department of Health and Social Care	40	395.20	0.63	Government Dept.
Department of Health for Northern Ireland	1	11.00	0.24	Government Dept.
Department of Social Security	20	122.40	0.40	Government Dept.
Disabled Persons Transport Advisory Committee	4	6.25	0.37	Non-dept. Agency
Disclosure and Barring Service	6	11.33	0.47	Non-dept. Agency
Driver and Vehicle Licensing Agency	28	51.11	0.45	Non-dept. Agency
Driver and Vehicle Standards Agency	5	19.20	0.37	Non-dept. Agency
Economic and Social Research Council	9	8.89	0.42	Non-dept. Agency
Education and Skills Funding Agency	3	12.67	0.56	Non-dept. Agency
Engineering and Physical Sciences Research Council	5	9.80	0.46	Non-dept. Agency
Engineering Construction Industry Training Board	3	6.33	0.30	Non-dept. Agency
English Institute of Sport	4	5.75	0.20	Non-dept. Agency
Environment Agency	25	288.92	0.61	Non-dept. Agency
Equality and Human Rights Commission	12	55.58	0.39	Non-dept. Agency
Estyn	2	5.50	0.22	Non-dept. Agency
Export Guarantees Advisory Council	1	17.00	0.38	Non-dept. Agency
Financial Conduct Authority	9	377.00	0.62	Non-dept. Agency
Financial Reporting Council	14	15.93	0.47	Non-dept. Agency
Financial Services Authority	28	250.43	0.54	Non-dept. Agency
Fire Service College	2	22.50	0.48	Non-dept. Agency
Foreign and Commonwealth Office	40	471.72	0.62	Government Dept.
Forensic Science Regulator	1	7.00	0.33	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Forest Research	1	6.00	0.24	Non-dept. Agency
Gambling Commission	11	48.09	0.44	Non-dept. Agency
Gangmasters and Labour Abuse Authority	2	15.00	0.65	Non-dept. Agency
Gas and Electricity Markets Authority	2	8.00	0.37	Non-dept. Agency
Government Communications Headquarters	1	6.00	0.33	Non-dept. Agency
Government Equalities Office	12	16.33	0.37	Non-dept. Agency
Groceries Code Adjudicator	7	31.43	0.50	Non-dept. Agency
Health and Safety Executive	40	156.30	0.58	Non-dept. Agency
Health Education England	7	46.14	0.61	Non-dept. Agency
Health Research Authority	1	7.00	0.27	Non-dept. Agency
Higher Education Statistics Agency	3	8.00	0.34	Non-dept. Agency
Highways England	5	181.80	0.58	Non-dept. Agency
Home Office	40	943.78	0.72	Government Dept.
Horserace Betting Levy Board	6	9.50	0.41	Non-dept. Agency
House of Lords Appointments Commission	5	6.20	0.38	Non-dept. Agency
Human Fertilisation and Embryology Authority	20	33.60	0.42	Non-dept. Agency
Human Tissue Authority	5	19.60	0.43	Non-dept. Agency
Hydrographic Office	3	23.67	0.55	Non-dept. Agency
Independent Anti-slavery Commissioner	4	9.00	0.32	Non-dept. Agency
Independent Commission for Aid Impact	8	21.00	0.41	Non-dept. Agency
Independent Office for Police Conduct	5	19.80	0.56	Non-dept. Agency
Independent Parliamentary Standards Authority	11	237.09	0.48	Non-dept. Agency
Independent Reconfiguration Panel	6	15.17	0.31	Non-dept. Agency
Independent Reviewer of Terrorism Legislation	1	5.00	0.61	Non-dept. Agency
Industrial Development Advisory Board	1	12.00	0.34	Non-dept. Agency
Industrial Injuries Advisory Council	13	10.00	0.41	Non-dept. Agency
Information Commissioner's Office	13	21.54	0.48	Non-dept. Agency
Infrastructure and Projects Authority	4	13.25	0.37	Non-dept. Agency
Insolvency Service	21	15.43	0.42	Non-dept. Agency
Institute for Apprenticeships and Technical Education	3	9.67	0.39	Non-dept. Agency
Intellectual Property Office	6	12.67	0.45	Non-dept. Agency
Intelligence Services Commissioner	3	5.33	0.49	Non-dept. Agency
Investigatory Powers Commissioner's Office	1	8.00	0.47	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Investigatory Powers Tribunal	5	10.60	0.36	Non-dept. Agency
Joint Nature Conservation Committee	4	9.25	0.32	Non-dept. Agency
Judicial Appointments Commission	11	14.64	0.50	Non-dept. Agency
Law Commission	40	113.22	0.49	Non-dept. Agency
Leasehold Advisory Service	5	8.20	0.38	Non-dept. Agency
Legal Aid Agency	7	21.29	0.46	Non-dept. Agency
Legal Services Board	2	10.00	0.30	Non-dept. Agency
Low Pay Commission	21	46.10	0.41	Non-dept. Agency
Marine Accident Investigation Branch	11	19.82	0.37	Non-dept. Agency
Marine Management Organisation	11	13.45	0.44	Non-dept. Agency
Maritime and Coastguard Agency	20	18.10	0.46	Non-dept. Agency
Medical Research Council	38	25.71	0.47	Non-dept. Agency
Medicines and Healthcare products Regulatory Agency	15	41.20	0.46	Non-dept. Agency
Migration Advisory Committee	12	33.67	0.31	Non-dept. Agency
Ministry of Defence	40	496.85	0.71	Government Dept.
Ministry of Housing, Communities and Local Government	4	106.50	0.39	Government Dept.
Ministry of Justice	18	235.72	0.62	Government Dept.
National Data Guardian	2	26.50	0.36	Non-dept. Agency
National Infrastructure Commission	5	31.80	0.46	Non-dept. Agency
National Institute for Health and Care Excellence	7	62.57	0.51	Non-dept. Agency
Natural Environment Research Council	12	10.00	0.43	Non-dept. Agency
Network Rail	28	485.57	0.56	Non-dept. Agency
Northern Ireland Housing Executive	18	18.72	0.39	Non-dept. Agency
Northern Ireland Human Rights Commission	10	15.50	0.37	Non-dept. Agency
Northern Ireland Office	40	80.85	0.45	Government Dept.
Northern Ireland Policing Board	8	7.38	0.44	Non-dept. Agency
Northern Lighthouse Board	2	10.50	0.52	Non-dept. Agency
Nuclear Decommissioning Authority	11	14.91	0.45	Non-dept. Agency
Office for Budget Responsibility	10	191.00	0.30	Non-dept. Agency
Office for National Statistics	24	111.04	0.41	Non-dept. Agency
Office for Students	4	52.25	0.51	Non-dept. Agency
Office of Communications	20	263.35	0.58	Non-dept. Agency
Office of Electricity Regulation	4	6.25	0.45	Non-dept. Agency
Office of Gas and Electricity Markets	21	152.19	0.46	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	**N. of Observations**	**Avg. N. Mentions**	**Avg. Reputation**	**Agency Type**
Office of Gas Supply	14	19.00	0.38	Non-dept. Agency
Office of Manpower Economics	2	8.00	0.57	Non-dept. Agency
Office of Rail and Road	16	53.31	0.41	Non-dept. Agency
Office of Tax Simplification	9	18.33	0.39	Non-dept. Agency
Office of the Children's Commissioner	9	9.78	0.35	Non-dept. Agency
Office of the Immigration Services Commissioner	1	6.00	0.39	Non-dept. Agency
Office of the Leader of the House of Commons	2	5.50	0.50	Government Dept.
Office of the Police Ombudsman for Northern Ireland	3	12.33	0.48	Non-dept. Agency
Office of the Public Guardian	4	17.50	0.42	Non-dept. Agency
Office of the Schools Adjudicator	1	7.00	0.45	Non-dept. Agency
Oil and Gas Authority	4	31.50	0.54	Non-dept. Agency
Oil and Pipelines Agency	1	6.00	0.21	Non-dept. Agency
Parliamentary and Health Service Ombudsman	7	9.71	0.38	Non-dept. Agency
Parole Board	26	33.27	0.36	Non-dept. Agency
Payment Systems Regulator	2	15.50	0.47	Non-dept. Agency
Pensions Advisory Service	3	16.33	0.47	Non-dept. Agency
Pensions Regulator	10	30.70	0.46	Non-dept. Agency
Planning Inspectorate	16	43.94	0.44	Non-dept. Agency
Police Service of Northern Ireland	21	50.90	0.56	Non-dept. Agency
Prison and Probation Service	3	17.00	0.57	Non-dept. Agency
Prison Service	30	113.87	0.53	Non-dept. Agency
Privy Council Office	9	23.11	0.39	Non-dept. Agency
Professional Standards Authority	7	9.43	0.35	Non-dept. Agency
Prudential Regulatory Authority	8	30.25	0.48	Non-dept. Agency
Public Health England	10	112.10	0.61	Non-dept. Agency
Public Health Wales	1	6.00	0.55	Non-dept. Agency
Pubs Code Adjudicator	4	26.00	0.42	Non-dept. Agency
Rail Accident Investigation Branch	1	5.00	0.48	Non-dept. Agency
Rail Safety and Standards Board	5	14.60	0.45	Non-dept. Agency
Regulatory Policy Committee	1	8.00	0.22	Non-dept. Agency
Royal Mint	10	16.30	0.39	Non-dept. Agency
Rural Payments Agency	17	31.29	0.40	Non-dept. Agency
School Teachers' Review Body	8	8.50	0.37	Non-dept. Agency

Table A.4: Sample of UK agencies and descriptive statistics. *(continued)*

Agecy Name	N. of Observations	Avg. N. Mentions	Avg. Reputation	Agency Type
Science and Technology Facilities Council	2	24.00	0.54	Non-dept. Agency
Sea Fish Industry Authority	13	17.31	0.46	Non-dept. Agency
Secret Intelligence Service	26	25.73	0.40	Non-dept. Agency
Security Industry Authority	10	57.10	0.50	Non-dept. Agency
Senior Salaries Review Body	15	14.40	0.32	Non-dept. Agency
Sentencing Council	11	35.36	0.47	Non-dept. Agency
Service Complaints Ombudsman	1	11.00	0.42	Non-dept. Agency
Service Prosecuting Authority	2	9.00	0.46	Non-dept. Agency
Single Financial Guidance Body	2	25.00	0.46	Non-dept. Agency
Single Source Regulations Office	1	8.00	0.45	Non-dept. Agency
Small Business Commissioner	1	7.00	0.42	Non-dept. Agency
Social Mobility Commission	4	45.00	0.34	Non-dept. Agency
Social Security Advisory Committee	31	24.45	0.28	Non-dept. Agency
Space Agency	21	15.33	0.44	Non-dept. Agency
Sports Council for Wales	2	16.50	0.37	Non-dept. Agency
Sports Grounds Safety Authority	2	12.50	0.44	Non-dept. Agency
Stabilisation Unit	1	6.00	0.57	Non-dept. Agency
Standards and Testing Agency	1	9.00	0.38	Non-dept. Agency
Surveillance Camera Commissioner	1	5.00	0.33	Non-dept. Agency
Treasury	40	1,828.97	0.64	Government Dept.
UK Export Finance	7	21.86	0.56	Government Dept.
Valuation Office Agency	16	19.75	0.41	Non-dept. Agency
Valuation Tribunal Service	1	14.00	0.48	Non-dept. Agency
Veterinary Medicines Directorate	3	7.67	0.49	Non-dept. Agency
Veterinary Products Committee	2	5.50	0.32	Non-dept. Agency
Victims' Commissioner	6	20.83	0.52	Non-dept. Agency
Wales Audit Office	4	6.50	0.32	Non-dept. Agency
Welsh Language Commissioner	2	15.50	0.34	Non-dept. Agency
Youth Justice Board	21	28.90	0.51	Non-dept. Agency

A.2 Reputation of Agencies in Similar Policy Domains

An important dimension of theories of reputation is agencies' reputation for uniqueness and therefore researchers wishing to compare the reputation of agencies should think carefully about whether the type of comparison is meaningful given the different policy sectors, organisational field, and the different responsibilities (Carpenter 2020). For instance, in Figure A.1 below I provide some examples of same-field comparisons for US agencies overseeing financial institutions and UK network industries' regulators.

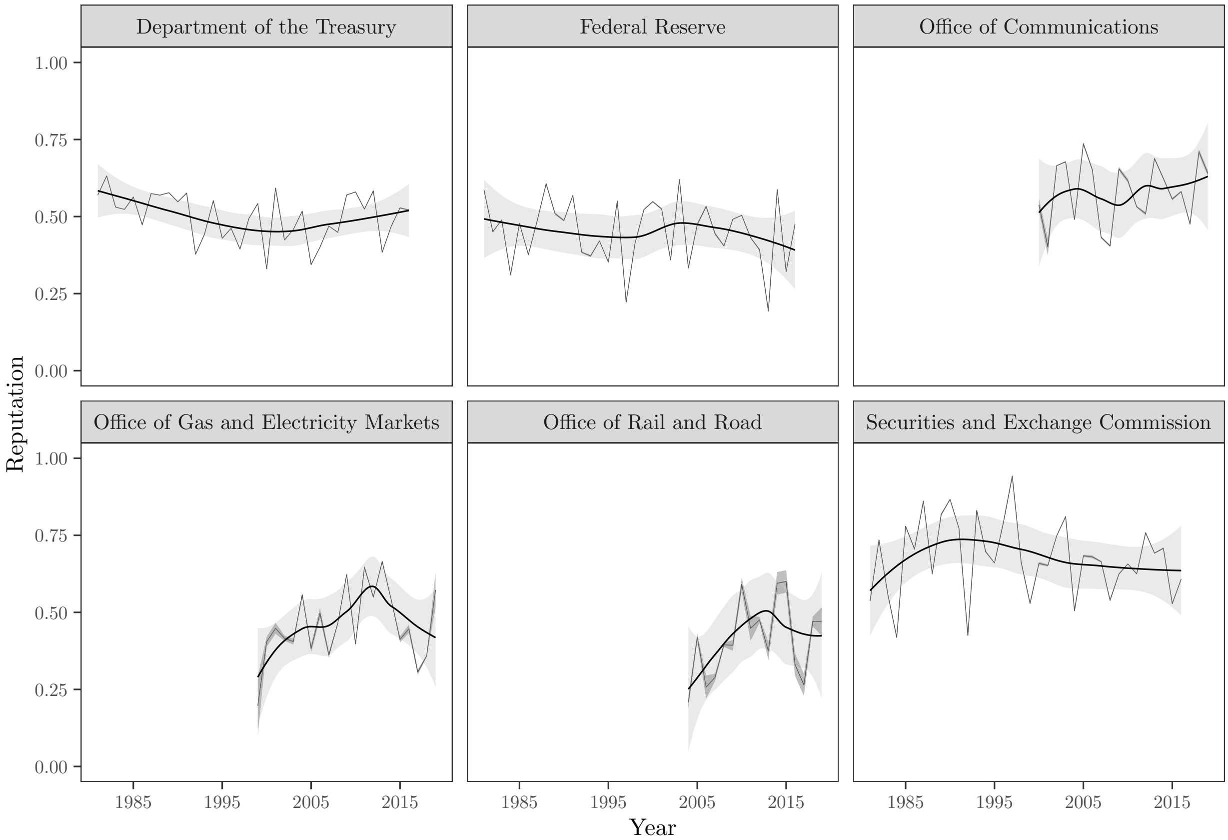

Figure A.1: Reputation estimates for agencies belonging to the same policy domain. Reputation of US agencies overseeing financial institutions (above) and reputation of UK network industries' regulators (below).

A.3 Positivity Vector: Robustness to Different Specifications

The measurement strategy presented in the main text produces estimates of the reputation of agency a in year t by measuring the cosine similarity between the "agency embedding" and a positivity vector in every year (see Section *Agencies as Word Embeddings* in the manuscript). Here I show how the reputation estimates are robust to different specifications of the positivity vector. I build two alternative positivity vectors. One with only a vector consisting of the difference between the vector representation of "good""" and "bad" (Positivity 1), and one with all the positive and negative words of the widely used LIWC dictionary for sentiment analysis. In particular, I create a vector that sums together the embeddings of all the tokens of each local corpus that appear in the positivity dictionary, and then deduct the embeddings of all the tokens that appear in the negativity dictionary (Positivity 2). Figure A.2 reports the relationship between the reputation estimates measured with the baseline positivity vector (on the x-axis) and the two alternative positivity vectors (y-axis) for both the UK and the US. The high and positive correlation suggests that, although the selection of the words used to build the positivity vector is important, alternative positivity vectors produce similar estimates.

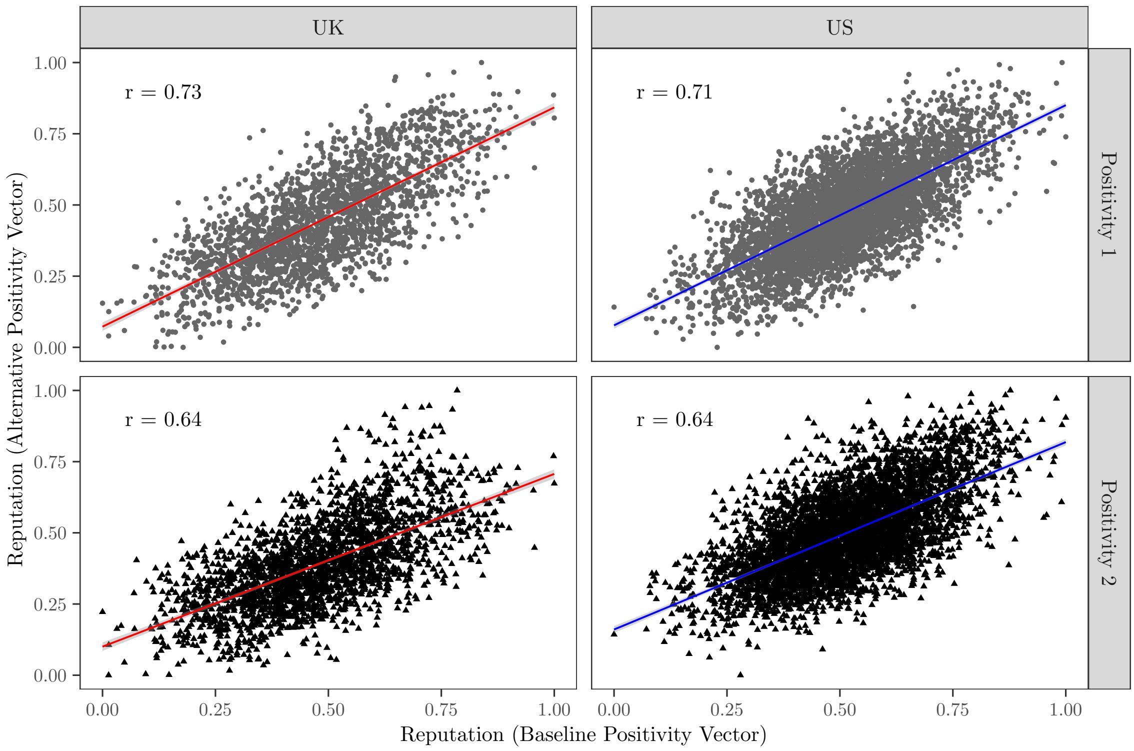

Figure A.2: Reputation estimates derived from different specifications of the positivity vectors. On the x axis, reputation estimates as measured with the baseline positivity vector reported in the manuscript. On the y axis, reputation estimates as measured with two alternative positivity vectors. Correlation coefficients reported in each panel.

A.4 Sub-setting Speeches by Topic

The proposed measurement strategy rests on the assumption for which the way politicians talk about bureaucracy in parliamentary speeches can be meaningfully used to estimate bureaucratic reputation. Sometimes, however, politicians can talk about some agencies whereas in fact they are debating issues that are not strictly related to the agency's activities or characteristics, and hence estimates of reputation based on such speeches might fail to capture the true reputation of the agency.

Let us consider, for instance, military agencies, which might be mentioned often by politicians who are praising military values and the courage and sacrifice of soldiers, while not focusing on the performance, the activities, the budget, or the expertise of those agencies. In such grandstanding speeches, agencies such as the Department of Veterans Affairs, the Air Force, or the Navy are very likely to be mentioned and to be used with a positive connotation. Their vector representation might therefore be similar to the positivity vector and upwardly bias the estimate of the reputation of those agencies.

In this section I show how the reputation estimates are sensitive to such speeches. I identify and remove all the US speeches mentioning *both* a military agency and a pre-defined set of words capturing military values. I then estimate the reputation of agencies from this new corpus of speeches and compare the estimates with the baseline estimates from the total corpus of speeches. Below I report the sample of US military agencies and the words capturing military values. After removing the speeches mentioning at least a military agency and a military value, the corpus size drops from 2,528,833 to 2,501,098 speeches (-1.1%).

Military Agencies: Air Force, Central Intelligence Agency, Department of Defense, Department of the Army, Department of the Navy, Department of Veterans Affairs, National Security Agency, Pentagon.

Military Values: duty, loyalty, sacrifice, honor, courage, selfless, integrity.

In Table A.5 I report the average number of mentions and the average reputation of military agencies from two different corpora. The *All Speeches* corpus contains all the speeches and it is the same used to measure reputation as outlined in the manuscript. For the *No Military Values* corpus, all speeches mentioning (at least) one military agency and (at least) one military-value word were removed. This is clearly a very conservative way of detecting how much speeches which praise military values drive the reputation estimates of military agencies, for the removed speeches can also convey information that should legitimately contribute to the reputation estimates.

	All Speeches		**No Military Values**	
Agency	Avg. Mention	Avg. Reputation	Avg. Mentions	Avg. Reputation
Air Force	1,038.31	0.75	548.28	0.64
Central Intelligence Agency	532.47	0.59	297.86	0.56
Department of Defense	1,507.64	0.73	965.67	0.66
Department of the Army	28.85	0.53	18.15	0.47
Department of the Navy	994.31	0.75	511.03	0.64
Department of Veterans Affairs	536.75	0.75	301.47	0.63
National Security Agency	84.53	0.53	51.86	0.48
Pentagon	568.36	0.45	332.94	0.44

Table A.5: Average reputation and average mentions of military agencies based on estimation from total corpus of speeches and from an alternative corpus where speeches containing military agencies and military values were removed. Averages compute from the same agency-year observations.

The average number of mentions of military agencies in the no-military-values corpus drops significantly compared to the all-speeches corpus. In the all-speeches corpus, for instance, the Department of Defense is mentioned on average 1,508 times, with an average reputation of .73. But once we remove all the military agency-military value speeches, the average number of mentions drops to 966 and reputation to .66. The reputation in the all-speeches dataset is on average 12.6% higher than in the no-military-values dataset. However, the correlation between the estimates is very high, equal to .73, and the time trend of the agencies reputation remains similar (see Figure A.3).

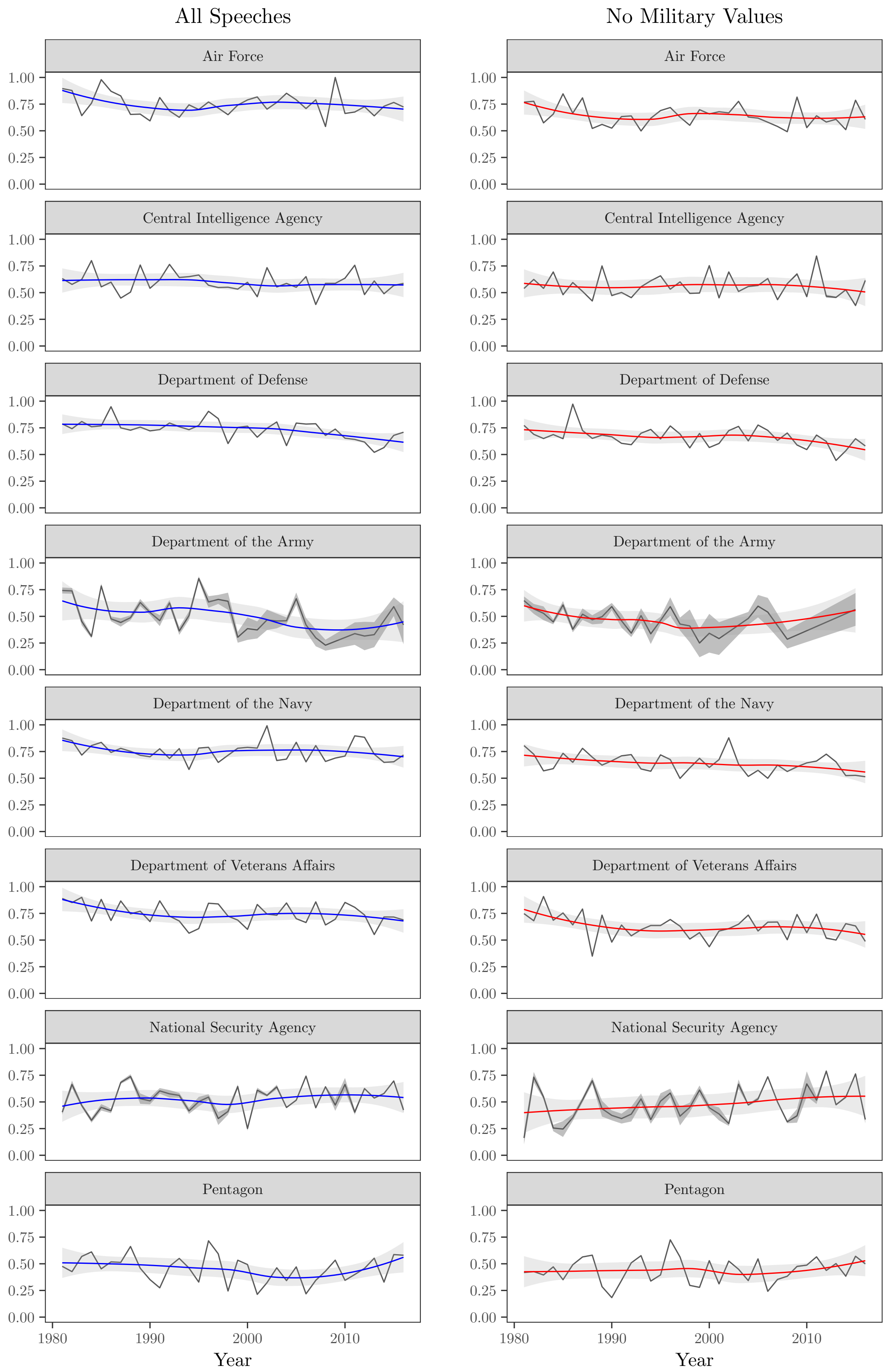

Figure A.3: Reputation of military agencies measured from two alternative corpus, one with all the speeches; one without speeches mentioning at the same time at least a military agency and a word capturing military values.

A.5 Validation: Newspaper Articles

In Table A.6 I report descriptive statistics of the sample of UK newspaper articles used to compare the reputation estimates measured from the corpus of parliamentary speeches with estimates derived from a large corpus of more than 1.2 million newspaper articles. The sample consists of all published articles in 12 major UK newspapers between 2014 and 2019. Articles were accessed through the LexisNexis API.

Newspaper	**N. Articles**	**Avg. Length (N. Words)**
Daily Mail	59,129	501
Daily Mirror	60,675	210
Daily Star	56,964	168
Financial Times	26,887	477
The Daily Telegraph	559,335	366
The Guardian	48,611	837
The Independent	54,806	553
The Observer	67,234	732
The Sun	168,529	206
The Sunday Telegraph	39,698	464
The Sunday Times	48,045	473
The Times	48,929	355
Total	**1,238,842**	**445**

Table A.6: List of UK newspapers and number of articles used to estimate bureaucratic reputation from the news. The articles were obtained through the LexisNexis API. I implemented the same text-preprocessing steps and estimation procedures as those used for estimating reputation from parliamentary speeches.

Figure A.4 shows the correlation between the reputation estimates derived from the two corpora for the agency-year observations with the largest number of mentions in the news. The estimation of word embeddings from this corpus of newspaper articles is identical to the one used for parliamentary speeches, with one exception. When building the vocabulary of unique words for which I estimate word embeddings, I keep all tokens which appear in the speeches at least 5 times, whereas for newspaper articles I keep tokens appearing at least 15 times. Because of the type of language used in the news, the number of unique tokens is much larger than that of speeches and sub-setting words occurring at least 15 times help me work with a more tractable term-co-occurrence matrix.

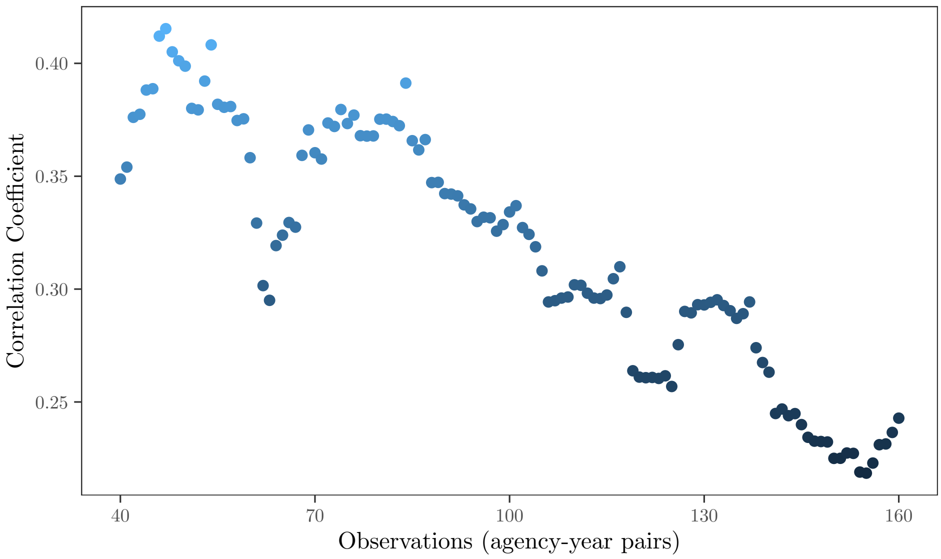

Figure A.4: Correlation coefficients between reputation estimates derived from speeches and newspaper articles. On the x-axis, the first N agency-year pairs (observations) with the highest number of mentions in the news. When agencies appear often in the news, the correlation between the two estimates is higher than 0.4 (Pearson's product-moment correlation). The correlation coefficient decreases when including more agency-pairs that are cited less often in the news.

A.6 Autonomy and Reputation: Robustness Tests

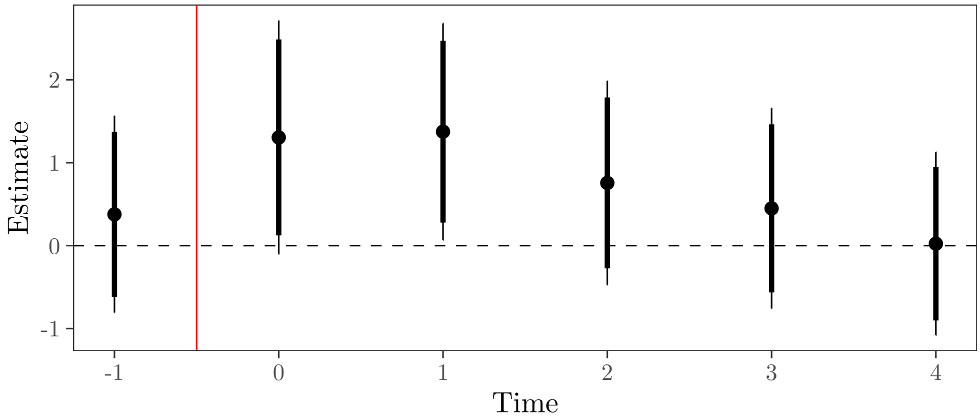

Figure A.5: “WLS regression estimates of the dynamic effect of reputation on autonomy. 90 and 95% confidence intervals estimated with SE clustered by agency.”

DV:	Autonomy				
Model:	(1)	(2)	(3)	(4)	(5)
Reputation	1.42^{**}	1.43^{**}	1.28^{**}	1.27^{*}	1.30^{*}
	(0.46)	(0.49)	(0.47)	(0.56)	(0.57)
Reputation_{t-1}	1.09^{*}	1.41^{**}	1.60^{**}	1.08^{+}	1.37^{**}
	(0.53)	(0.48)	(0.50)	(0.60)	(0.50)
Reputation_{t-2}	1.01^{+}	1.07^{*}	1.32^{**}	0.68	0.76
	(0.52)	(0.53)	(0.48)	(0.63)	(0.61)
Reputation_{t-3}	0.63	0.49	0.45	0.83	0.45
	(0.47)	(0.50)	(0.50)	(0.56)	(0.55)
Reputation_{t-4}		0.08	-0.19		0.02
		(0.46)	(0.44)		(0.50)
Reputation_{t-5}			0.03		
			(0.46)		
Reputation_{t+1}				0.49	0.38
				(0.55)	(0.51)
Fixed-effects					
Agency	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓
Observations	279	236	198	236	198
R^2	0.85	0.89	0.91	0.84	0.89
Within R^2	0.06	0.09	0.14	0.04	0.07

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, +: 0.1*

Table A.7: WLS dynamic treatment effects. Dependent variable is bureaucratic autonomy. Number of mentions in floor debates always included as covariate.

A.7 Bureaucratic Polarisation: Robustness Tests

I present four robustness tests for the results reported in Table 1.3 of the manuscript. In Table A.8 below, I show that the WLS estimates from Model (8) are robust to using heteroskedasticity-consistent standard errors (i.e., HC2 variance estimator) (Model (1)), an alternative coding of the variable agency type which merges agencies within the Executive Office of the President with Executive Departments, and non-for-profit public organisations with independent agencies (Model (2)), a model which uses both robust standard errors and the new agency type fixed effects (Model (3)), and a model which uses robust standard errors, the new type of agency type fixed effects, and only 2014 data, the year when Selin's data collection about agencies' statutes ended (Model (4)). The estimated effect of independence measured as the requirements on key decision-makers is always distinguishable from zero at 90% and 95% confidence level.

	DV: Bureaucratic Polarisation			
	(1)	(2)	(3)	(4)
Ind: Decision-Makers	-0.04^{+}	$\mathbf{-0.03}^{*}$	-0.03^{+}	$\mathbf{-0.05}^{*}$
	(0.02)	(0.02)	(0.02)	(0.02)
Ind: Political Review	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)
Agency Type FE	Yes	Yes	Yes	Yes
R^2	0.08	0.08	0.08	0.11
Num. obs.	102	102	102	93

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table A.8: WLS estimates, standard errors in parenthesis. Model (1): HC2 variance estimator; Model (2): new coding of agency type; Model (3): HC2 SE and new coding of agency type; Model (4): Only 2014 data, HC2 SE, and new coding of agency type.

Appendix B

Appendix to Chapter 2

B.1 Sample of Agencies

In the following tables, I report all the agencies for which I produce sentiments' positivity 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 positivity among both parties.

B.1.1 Sample of UK Agencies

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

			Conservative		**Labour**	
Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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

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

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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
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

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

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Positivity	Avg. Mentions	Avg. Positivity
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
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

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

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Positivity	Avg. Mentions	Avg. Positivity
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
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

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

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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
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

B.1.2 Sample of US Agencies

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

			Republican		Democratic	
Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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

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

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
Citizen and Immigration Services	Ex. Sub-agency	7	21	0.49	14	0.56
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

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

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Positivity	Avg. Mentions	Avg. Positivity
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
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

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

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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
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

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

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Positivity	Avg. Mentions	Avg. Positivity
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
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

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

Agency Name	Agency Type	N. obs.	Avg. Mentions	Avg. Positivity	Avg. Mentions	Avg. Positivity
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
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

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

Agency Name	**Agency Type**	**N. obs.**	**Avg. Mentions**	**Avg. Positivity**	**Avg. Mentions**	**Avg. Positivity**
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
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

B.2 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 (German 2015).

B.3 Dictionary Analysis

B.3.1 Text Pre-Processing

I implement the dictionary-based measurement through the following steps: the speeches are the same used to estimate statements' positivity (see Section 1.4.2 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.

B.3.2 Dictionary 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 B.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
Dictionary Measure	2.760*	0.817***	0.773***	0.540**
	(1.296)	(0.145)	(0.110)	(0.198)
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 SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

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

B.4 Study 1, Selective Evaluation: Robustness Tests

In Tables B.5 and B.6 below I show that the results reported in Tables 2.2 and 2.3 are robust to clustering SE at party-prime minister and party-presidency level.

DV:	Positivity [0,1]		
Country:	UK		
Model:	(1)	(2)	(3)
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			✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

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

DV:	Positivity [0,1]					
Country:	US					
	Party-Government				Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
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		✓	✓	✓	✓	✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

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

In Table B.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:	Positivity [0,1]				
Country:	US				
	Party-Government			Party-Agency	
Model:	(1)	(2)	(3)	(4)	(5)
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	✓	✓	✓	✓	✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

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

In Table B.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:	Positivity [0,1]					
Country:	UK			US		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
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			✓		✓	✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.8: Partisanhsip, Ideology, and Statements' Positivity, 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 B.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:	Positivity [0,1]		
Country:	US		
Model:	(1)	(2)	(3)
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			✓
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) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.9: Partisanhsip, Ideology, and Statements' Positivity, 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 B.10 I report falsification tests of the difference-in-differences strategy of Section *Scandals in the US Federal Bureaucracy.* 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)
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	✓	✓
Observations	1,831	1,950
R^2	0.182	0.187
Within R^2	0.000	0.001

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.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.

B.5 Study 2, Selective Information-Acquisition

B.5.1 Bureaucracy Appearing Before Senate Committees: Robustness Tests

Linear probability models with alternative clustering strategy (at presidency level).

DV:	Bureaucracy as Witness [0,1]				
	Partisanship		Ideology		Both
Model:	(1)	(2)	(3)	(4)	(5)
Comm. Chair-President Partisan Align.	-0.066^{+}	-0.063^{+}			-0.207*
	(0.030)	(0.031)			(0.052)
Comm. Majority-Govt. Partisan Align.		-0.011		0.011	0.030
		(0.038)		(0.052)	(0.051)
Comm. Chair-President Id. Dist.			0.026	0.025	-0.110
			(0.098)	(0.099)	(0.076)
Fixed-effects					
Congress	✓	✓	✓	✓	✓
Committee	✓	✓	✓	✓	✓
Observations	5,179	5,179	4,776	4,776	4,776
R^2	0.094	0.094	0.084	0.084	0.085
Within R^2	0.001	0.001	0.000	0.000	0.001

Clustered (pres) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, +: 0.1*

Table B.11: OLS estimates of the effect of committe chair-government partisan alignment on the probability of a bureaucracy appearing as a witness in Congressional Senate hearings.

B.5.2 Use of Statistical Facts: Robustness Tests

In Table B.12 below I show results in Table 2.6 are robust to replacing date with year fixed-effects.

DV:	Statistical Facts (Abs. Frequency)					
Country:	US			UK		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Leg.-Govt. Partisan Alig.	-0.119***	-0.177***	0.452	-0.095***	-0.173***	0.130
	(0.029)	(0.054)	(0.357)	(0.019)	(0.037)	(0.196)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓	✓
Observations	247,570	247,570	247,570	171,155	171,155	171,155
R^2	0.165	0.201	0.194	0.195	0.283	0.648
Within R^2	0.135	0.176	0.173	0.082	0.152	0.612

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.12: Partisanhsip, Ideology, and Statements' Positivity, US Data. OLS estimates. absolute frequency of statistical facts 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 B.13 below I replicate the estimation of Table 2.6 for the US conditioning on the ideological distance between the agency and the legislator giving the speech.

DV:	Statistical Facts (Abs. Frequency)		
Country:	US		
Window Size:	20	50	Total
Model:	(1)	(2)	(3)
Leg.-Govt. Partisan Alig.	-0.151**	-0.245**	0.199
	(0.047)	(0.078)	(0.531)
Leg.-Agency Id. Dist.	-0.098	-0.053	-0.967
	(0.157)	(0.276)	(1.239)
Fixed-effects			
Legislator	✓	✓	✓
Agency	✓	✓	✓
Date	✓	✓	✓
Observations	95,334	95,334	95,334
R^2	0.200	0.236	0.293
Within R^2	0.120	0.156	0.189

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.13: Partisanhsip, Ideology, and Statements' Positivity, US Data. OLS estimates. Dependent variable absolute frequency of statistical facts 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 B.14 below I replicate the estimation of Table 2.6 replacing legislator-government partisan alignment (treatment) with two alternative treatments: legislator-agency partisan alignment and legislator-agency ideological distance.

DV:	Statistical Facts (Abs. Frequency)					
Country:	US					
	Ideology			Congruence		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Leg.-Agency Id. Dist.	0.032	0.157	-1.137			
	(0.148)	(0.263)	(1.191)			
Leg.-Agency Partisan Align.				0.060	-0.065	-2.665*
				(0.118)	(0.195)	(1.147)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Observations	95,334	95,334	95,334	37,043	37,043	37,043
R^2	0.200	0.236	0.293	0.279	0.313	0.343
Within R^2	0.120	0.156	0.189	0.135	0.173	0.196

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.14: Legislator-Agency Partisan Alignment and Ideological Distance. US Data. OLS estimates. Dependent variable is absolute frequency of statistical facts in US speeches. Controls include legislator's age and speech length (log number of words).

In Table B.15 below I replicate the estimation of Table 2.6 using the *tf-idf* of facts-words as dependent variable, and therefore down-weighting statistical-fact words that appear in many speeches.

To build the *tf-idf* metric, I first build a document-token matrix, with one row for every speech, and one columns for every unique token used in the corpus as a whole. 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 "average" which belong to the dictionary. If "average" appears in more speeches than "approximately," then "average" 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.

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

$$Fact_s = \sum_{t \in Dict} tf - idf_{t,s} \quad \text{with} \quad tf - idf_{t,s} = \frac{f_{t,s}}{|s|} \times log \frac{|S|}{|\{s \in S : t \in Dict\}|} \tag{B.1}$$

where t is each token within the pre-defined windows of words for speech s, $Dict$ the list of words capturing the use of statistical facts, 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 ($|\{s \in S : t \in Dict\}|$).

DV:	Statistical Facts (*tf-idf*)					
Country:	US			UK		
Window Size:	20	50	Total	20	50	Total
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Leg.-Govt. Partisan Alig.	-0.104***	-0.116**	0.221	-0.092***	-0.140***	0.162
	(0.025)	(0.036)	(0.157)	(0.027)	(0.041)	(0.098)
Fixed-effects						
Legislator	✓	✓	✓	✓	✓	✓
Agency	✓	✓	✓	✓	✓	✓
Date	✓	✓	✓	✓	✓	✓
Observations	247,570	247,570	247,570	171,155	171,155	171,155
R^2	0.224	0.269	0.316	0.226	0.312	0.669
Within R^2	0.145	0.192	0.208	0.075	0.142	0.602

Clustered (Legislator) SE in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Table B.15: Robustness Analysis: Argumentative Style, US and UK Data. OLS estimates. Dependent variable is *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.

Appendix C

Appendix to Chapter 3

C.1 Extraction Rules

Say verbs used to match syntactic rules : say, tell, show, claim, report, admit, acknowledge, present, explain, state, indicate, recommend, propose, advance, believe, think, affirm, conclude, propose, advise, encourage, argue, contend, set out, inform, suggest, advise.

Recommendation-type words are: recommendation, advise, suggestion, indication, proposal, attempt, document, idea, project, programme, conclusion, report, program, brief, paper, argument, thesis, statement, survey, study, suggestion, advice.

C.2 Dependency Parsing: Examples

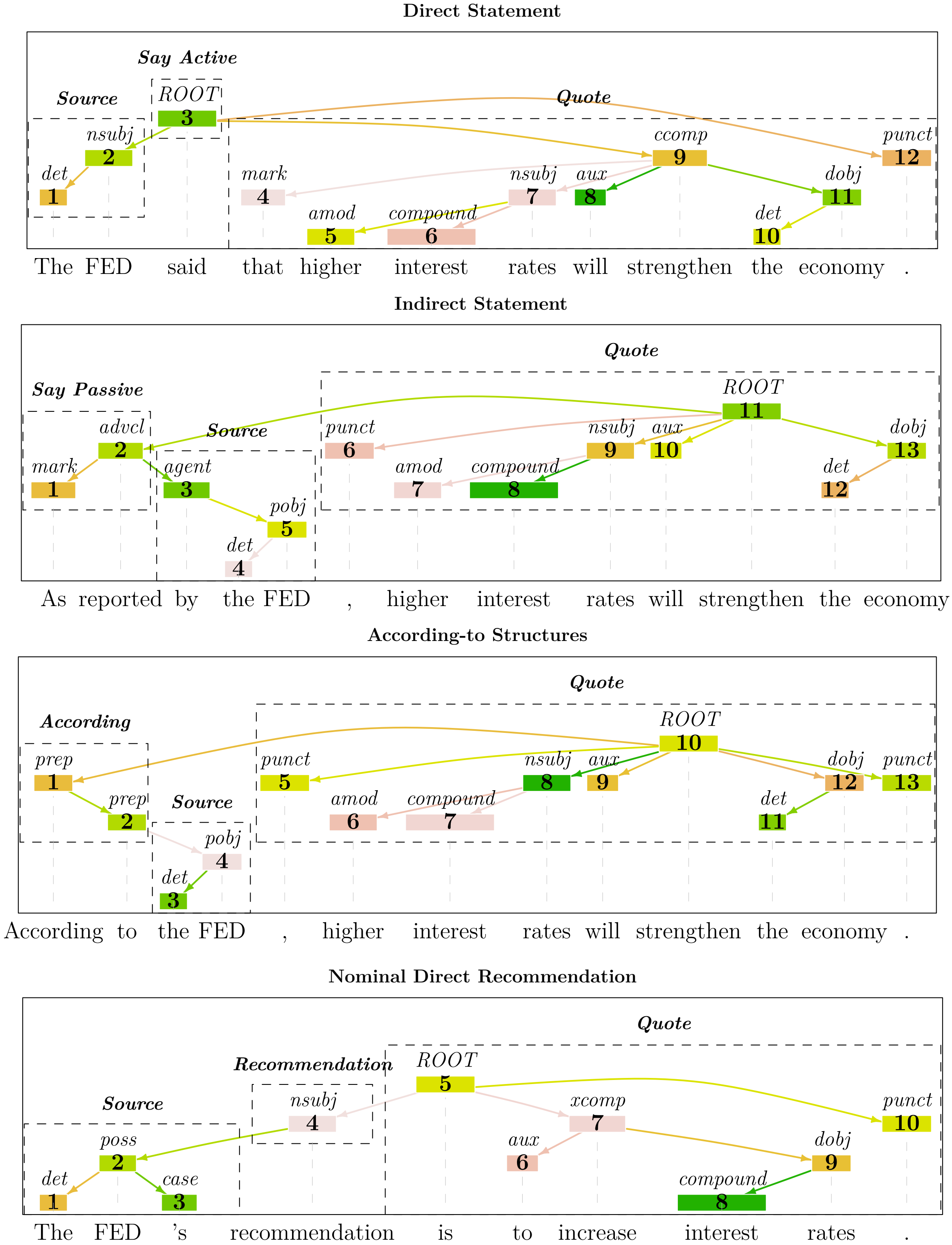

Figure C.1: Parsed dependency trees of the three remaining illustrative examples where the FED is used to support a statement. Implemented through the *rsyntax* package in R.

C.3 Committees' Speeches: Data Quality

I accessed transcripts of 42,277 congressional committee sessions from ProQuest. Each transcript consists of one text file, and no metadata exists to facilitate the extraction of single speeches. Speeches are nonetheless identifiable thanks to the way they appear in the text. The title and SURNAME of the speaker in fact precedes the speech and is reported in capital cases. "Mr. FORD," for instance, marks a new speech. Many individuals are heard in congressional committees. To extract speeches given by politicians, I exploit the fact that at the beginning of each transcript, the names of all congresspersons are reported followed by their home state. From every transcript I therefore extract all the name of politicians with a regular expression that matches the name and surname of individuals followed by the name of their respective state. Only speeches given by any of the extracted names are parsed from the transcript.

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

C.4 Robustness Checks

DV:	log(1+dv)				
Model:	(1)	(2)	(3)	(4)	(5)
Ideological Distance	-0.076*	-0.070*	-0.070*	-0.060*	-0.069
	(0.030)	(0.028)	(0.028)	(0.029)	(0.042)
Log Agency Mentions	-0.042***	-0.026*	-0.026*		
	(0.005)	(0.012)	(0.012)		
Legislator Covs.			✓	✓	
Fixed-effects					
Legislator	✓	✓	✓	✓	
Year	✓	✓	✓		
Agency		✓	✓		
Agency-Year				✓	✓
Legislator-Year					✓
Observations	20,578	20,578	20,548	20,548	20,578
R^2	0.071	0.092	0.093	0.135	0.382
Within R^2	0.004	0.001	0.001	0.001	0.000

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

Table C.1: OLS estimates. DV is log-transformed frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

DV:	Use of Bureaucratic Information (N. facts-words)			
Model:	(1)	(2)	(3)	(4)
Ideological Distance	-0.243**	-0.232**	-0.229**	-0.208**
	(0.079)	(0.075)	(0.075)	(0.078)
Log Agency Mentions	-0.084***	-0.066*	-0.066*	
	(0.013)	(0.029)	(0.029)	
Legislator Covs.			✓	✓
Fixed-effects				
Legislator	✓	✓	✓	✓
Year	✓	✓	✓	
Agency		✓	✓	
Agency-Year				✓
Observations	20,366	20,365	20,337	20,218
Squared Correlation	0.060	0.079	0.080	0.124
Pseudo R^2	0.045	0.059	0.059	0.087
BIC	74,924.774	74,520.666	74,457.741	79,250.949

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

Table C.2: Poisson estimates. DV is frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

DV:	Use of Bureaucratic Information (N. facts-words)	
Model:	(1)	(2)
Ideological Distance	-0.212*	-0.211*
	(0.103)	(0.103)
Floor Speech (Dummy)	-0.057	-0.058
	(0.059)	(0.058)
Ideological Distance × Floor Speech (Dummy)	-0.038	-0.037
	(0.128)	(0.128)
Legislator Covs.		✓
Fixed-effects		
Legislator	✓	✓
Year	✓	✓
Agency	✓	✓
Observations	20,578	20,548
R^2	0.080	0.080
Within R^2	0.001	0.001

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

Table C.3: OLS estimates. DV is frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

DV:	Use of Bureaucratic Information (N. facts-words)	
Model:	(1)	(2)
Ideological Distance	-0.163*	-0.165*
	(0.081)	(0.081)
Log Agency Mentions	-0.072*	-0.073*
	(0.031)	(0.031)
Legislator Covs.		✓
Fixed-effects		
Legislator	✓	✓
Year	✓	✓
Agency	✓	✓
Observations	19,380	19,350
R^2	0.082	0.083
Within R^2	0.001	0.001

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

Table C.4: OLS estimates for sample of speeches quoting only one agency. DV is frequency of of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.

DV:	Use of Bureaucratic Information (tf-idf of facts-words)				
Model:	(1)	(2)	(3)	(4)	(5)
Ideological Distance	-0.089**	-0.083**	-0.082**	-0.074*	-0.085*
	(0.029)	(0.028)	(0.028)	(0.029)	(0.043)
Log Agency Mentions	-0.032***	-0.027*	-0.027*		
	(0.005)	(0.011)	(0.011)		
Legislator Covs.			✓	✓	
Fixed-effects					
Legislator	✓	✓	✓	✓	
Year	✓	✓	✓		
Agency		✓	✓		
Agency-Year				✓	✓
Legislator-Year					✓
Observations	20,578	20,578	20,548	20,548	20,578
R^2	0.063	0.080	0.080	0.118	0.349
Within R^2	0.003	0.001	0.001	0.001	0.000

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

Table C.5: OLS estimates. DV is sum of tf-idf of fact-words of statistical facts and evidence in quotes of agencies mentioned in legislators' speeches.