Forecasting Sunlight: Using structural topic modeling to predict U.S. agency responsiveness to 2018 Freedom of Information Act requests

By: Lanlily M. Nguyen

Department of Political Science, University of California San Diego

Thesis Advisor: Margaret E. Roberts

March 30, 2020

Table of Contents

Acknowledgments

To my advisor Molly Roberts, your unwavering support fanned out the fire on this paper and I am so grateful to have you on my team for both your incredible expertise and compassion. Congratulations on your new baby. Thank you Wumbo and Justin for your help with organizing my data and allowing for this project to happen. To my dad, Tammy, Faith and friends, thank you for your uplifting encouragement, support, and love. To my mom, I remind myself of your resilience every day to keep me going. Thank you.

Abstract

The sparsity of research on the origins of transparency poses a challenge for researchers who seek to identify long-term trends in transparent governance or evaluate the role of macro-level factors that may predict more practices of accountability. This study introduces a means of overcoming this barrier and provides a structural topic modeling approach to understand access to government information using data from Freedom of Information Act (FOIA) requests and scraping Wikipedia textual descriptions across agency home pages. I compile data from 119 agencies and sub-agencies that allow me to estimate topical predictors of transparency. I find that the topical content of mission and description of the work of the agencies is highly predictive of the level of transparency, even when controlling for other factors such as agency size and proactive disclosures. My findings suggest that the type of work that an agency carries out, as well as the institutional legacy of particular types of agencies, is important to understanding how responsive they are to the public. For example, I find that public-facing financial agencies are highly transparent, whereas agencies that serve veterans have extremely low transparency rates. My results and approach should facilitate additional research on the mechanisms through which values and purposes can influence the abandonment of transparent government practices or incentives to be exposed under sunlight.

1 Introduction

Since 1967, the Freedom of Information Act (FOIA) has permitted requests to access any records created, possessed, or controlled by a federal regulatory agency, cabinet and military departments, and other organizations of the Executive Branch of the federal government. Unless covered by a specific FOIA exemption such as threats to national security or foreign policy, anyone— US citizen or not— may make a request.

FOIA's original purpose was to promote an "open society," one where federal agencies were not to withhold information if it were not dire. The law by many scholars and policymakers has been heralded to generally increase government transparency which brings a number of benefits to a society such as trust and legitimacy (Grimmelikhuijsen, 2009). From a public opinion perspective, attempts to reform public sector governance has been integral to establishing confidence in government since at least the early 1950s. Broadly defined, government transparency is the overall degree to which citizens, the media, and financial markets can observe the government's strategies, its actions, and the resulting outcomes (Alt, et al. 2006).

However, while FOIA may have been designed to increase overall government transparency, not all information under law may be revealed by agencies. Over time, the law has been amended repeatedly and has arguably decreased transparency with each revision. Infamously during the Bush administration's in 2006, leakers and the recipients of leaks of government information were targeted for criticism and denunciation, and were threatened with prosecution (Kirtley, 2010). On the other hand, FOIA law also in and of itself requires that agencies make certain records available for public inspection and copying without a formal

request, also known as “proactive disclosures.” Records made to be readily available include unpublished policy statements and agency interpretations and staff manuals that affect the public.

Requests for information through the Freedom of Information Act procedure have substantially increased over the span of several years. There are also more requests to more departments. Though more requests have been submitted, there are many more failures to respond to these requests. As to what may explain this discrepancy there are a few studies that attempt to provide an answer. Scholars have pointed to some reasons such as the increase in FOIA litigation that can extend the process (Chaffetz, 2016) or decreases in overall bureaucratic efficiency (Herz, 2008). To fill this gap as to what are the determinants of government transparency in the U.S. in complying with FOIA law, this study will focus on a broader view of agencies by understanding their goals and purposes and how they respond to FOIA requests. This study is a further step to answering what determines transparency by examining government agency typology through online mission statements.

It is very difficult, if not nearly impossible, to predict the behavior of agencies and governments on whether they will reveal more records and information to the public. However, if we look at some top-priority issues of today— such as health, economy, homeland security, energy, justice— and increasing pressure emanating from the public for the government to be more open, we can say that future behavior will be so complex as to require insights from across all agencies. Research across multiple agencies will help create a theoretical framework to advance fundamental understanding of how governments may operate in how they comply and respond to Freedom of Information Act requests.

In this study, my objective is to use topic modeling, a statistical method used to generate topics independently from sets of documented text, to investigate how the nature of agencies' activities might influence transparency. Of particular interest, structural topic modeling (STM; Roberts et al., 2013) allows me to assess the relationship between the topic (content and proportion) and my dependent variable: Freedom of Information Act request responsiveness. By using textual descriptions of agency activities and relating them to transparency, STM allows me to discover how the agencies' activities can predict the dynamic of government transparency. STM analyses have often been applied in political contexts (Grimmer and Stewart, 2013; Munksgaard and Demant, 2016). Introducing this method of analysis will allow me to generally examine the research area of transparency which is often in and of itself neglected as an outcome variable. To fill this gap, my independent variable will be sourced and inductively created across agency textual data pulled from Wikipedia and federal government websites that will allow topics from agency descriptors to explain agency responsiveness.

First, my results show that the topical content of the mission and description of the work of the agencies is highly predictive of the level of transparency. Topics that capture agencies that are more public-facing are ranked higher in transparency ratings than topics that are more likely to contain classified information, as well as agencies that historically struggle with bureaucratic efficiency and FOIA processing. This suggests that the particular functions of an agency, as well as the institutional design and purpose of particular types of agencies, is important to understand how responsive they are to the public in processing requests.

Secondly, holding constant previously studied variants for transparency, I find that proactive disclosures, agency size, and fiscal pressures do not substantially change the relative

ranking of topics in terms of their transparency rates. Based on this supporting evidence, my model captures the significance of topics and how they reveal predictions on the transparency of the agencies categorized under them.

While this approach may not entirely disentangle whether topics of agencies can determine government accessibility to information, current quantitative measures studying factors transparency face similar challenges. In contrast with existing approaches that test factors of transparency under very specific conditions (e.g. times of crisis), my approach offers new analytical insight by examining transparency at the federal level and sourcing data from the bureaucracies that adjudicate the information that they disclose. That is, I seek to provide a wider theory on determinants of transparency instead of explaining particular instances or rarer phenomena. My approach contributes both to understandings of the role of agency descriptions and missions, and to understandings of the factors behind the topical content that can determine agency responsiveness to the public in disclosing information.

2 Federal Agencies and the Freedom of Information Act

The goal of transparency in governance is to make decisions and behave in a way that is more accountable to citizens and the public to reduce government secrecy. Many scholars claim that it can combat problems of government corruption, and increase overall better governance, accountability, and trust (Banisar 2006; Birkinshaw 2006; Florini 2007) The solutions that come from these ideas often include providing open information to the public by several mechanisms to have accountable officials and decisions. While transparency does not guarantee the eradication of corruption necessarily, it works towards improving the relationship between

government officials and citizens. According to Fung (2013), there are four principles in transparency: (a) information about the operations and actions of public organizations that affect citizens' interests should be rich and readily available to the public; (b) the amount of available information should be adequate to the extent to which those organizations threaten citizens' interests; (c) information should be provided in ways that are accessible to those that use the information; and (d) social, political, and economic structures of society should be organized in ways that allow citizens to take action.

For the purposes of this study, I will be focusing on transparency in the form of how access to information— more specifically, how the Freedom of Information Act— has enabled U.S. agencies to be held more accountable to the public. Access to information laws, in general, are often lauded for their "disinfecting" properties in both raising the stakes for agencies to be ethically compliant and as a relatively successful punishment and reward system for kicking out sources of corruption. Berliner (2014) found that FOI laws make it more difficult for political actors to profit from the control of government information or to use public office for private gain. By increasing the risks of exposure, FOI laws reduce the expected utility of corruption (Berliner, 2014). Yet, as we will see, many agencies are only partially responsive to Freedom of Information Act requests. This raises the question of what types of agencies in the U.S. are most transparent and responsive.

The Freedom of Information Act (5 USC § 552) was passed in 1966 by President Lyndon B. Johnson with the intention to promulgate an "open society" (citation). Since then, further amendments and orders had hindered FOIA's original purpose of greater transparency (Kirtley, 2006). For example, in 2003, President Bush signed a classified directive that eliminated the

former instruction that "[i]f there is significant doubt about the need to classify information, it shall not be classified." In its place, it created a presumption that unauthorized disclosure of "foreign government information" is presumed to damage national security, and declared that classified information "shall not be declassified automatically as a result of any unauthorized disclosure of identical or similar information" (Exec. Order, 2003). The amendment also expanded the authority of the Director of the CIA to veto declassification rulings made by the Interagency Security Classification Appeals Panel (Kirtley, 2010). At the same time, FOIA requests have since then been increased and responses have backed-up in volume. Whether this is due to further constraints of FOIA procedural rules, increases in requests, or other ulterior political factors such as conflicts with national security interests, the causes are uncertain. What is certain is that transparency has decreased over time. FOIA backlogs have drastically increased since 2008, reaching an all-time high for several agencies in 2019.

Transparency in government at the federal level is often put on the shoulders of the presidential administration and less so the agencies themselves (Coglianese, 2009; Kirtley, 2010; Wasike, 2016). This makes sense as executive agencies take their orders from the president and offer these guidelines down the ladder of bureaucracy. While these sub-agencies should often be independent of politics and the presidency, e.g. watchdog departments and inspection agencies, how agencies may operate differently from administration to administration is hard to tell due to many factors that can change across time. Presidents may affect how agencies respond to FOIA requests, however, by changing the law itself. For example, in the post-9/11 era, the Bush administration's amendment to the original FOIA law was enacted in order to conduct counter-terrorism efforts. However, as a side effect, less oversight on agencies enabled for other

unintended activities to go unnoticed, such as prohibiting the previous requirement by the Department of Energy to release reports on toxic waste sites and their contents (U.S. Congress, 2004). The original purpose of this was so that public enemies could not use this information against the U.S. However, as a consequence, the safety of the public was at risk in the respect of the health of citizens and the environment.

Presidents are often blamed for general government performance (Coglianese, 2009; Kirtley, 2010; Wasike, 2016) and public releases of records via FOIA and transparency in general is no exception. Amendments and changes can, after all, greatly alter how agencies may respond to requests with new requirements for reading rooms, new exceptions introduced for requests, etc. Despite changes in the law, there has been an overall trend in decreased responsiveness since the early 2000s. Former President Bush's administration did put a substantial halt to requests being processed (Pack, 2004), however, this was hardly changed even with Obama's "Open Government" initiative where backlogs were still piling up. Trends since 2008 have only gotten worse where 2018 has marked one of the worst years of censorship as claimed by *The Associated Press* and several other major news outlets (Bridis, 2018). While there may be a point of interest in studying presidential conduct of how government transparency is handled, it appears largely different approaches still give in to the same trend. As seen in Figure 1, a sample of some of the largest U.S. federal agencies have an increasing trend of increasing backlogs.

What is largely overlooked is how these individual federal agencies were designed for a purpose that is consistent through time. Agencies themselves can be examined for how they handle requests and how they process them. To fill this research gap, I propose a research design

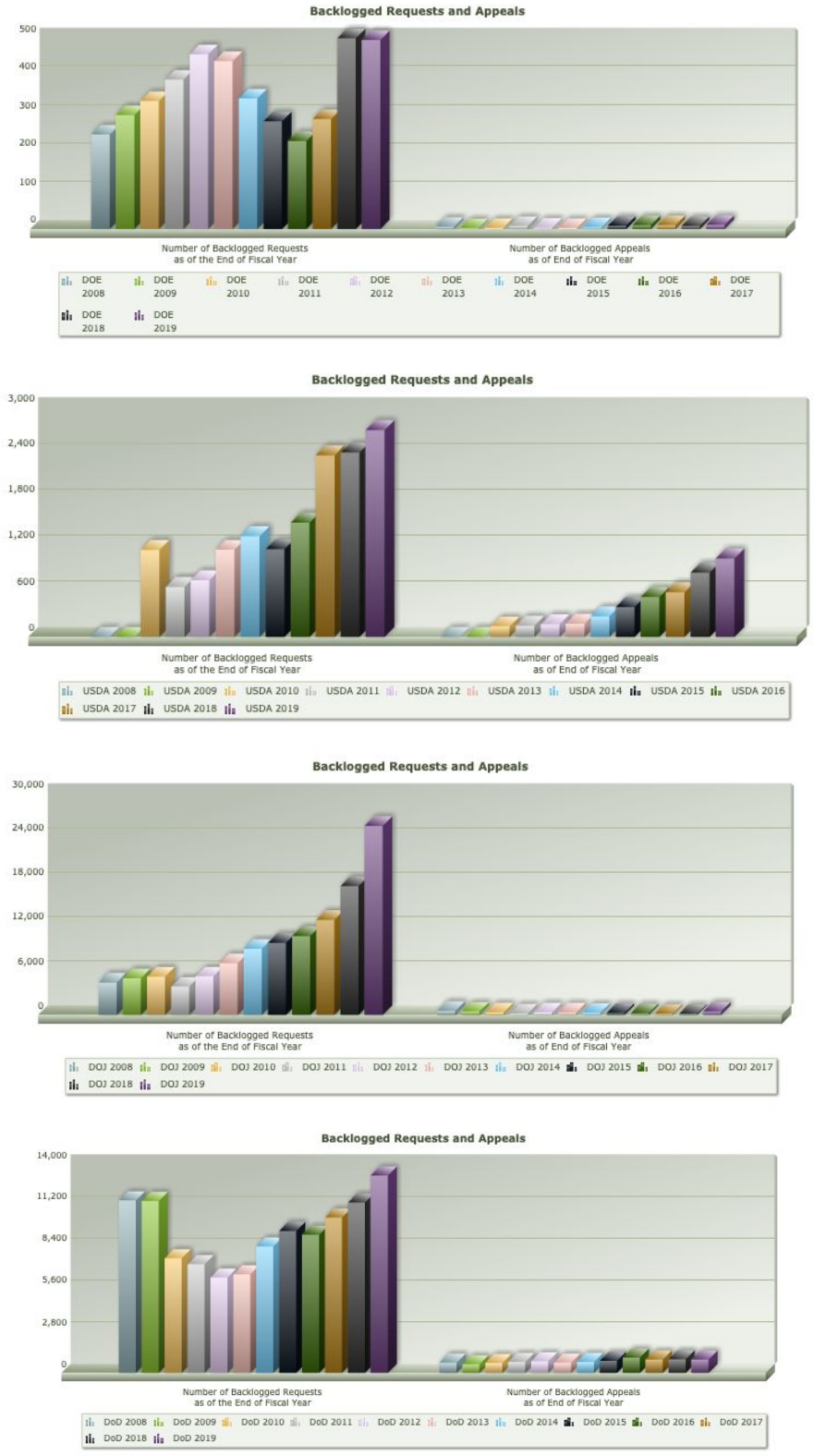

Figure 1. Depart of Energy, Department of Agriculture, Department of Justice, and Department of Defense FOIA backlogs 2008-2019. Source: FOIA.gov.

that analyzes the specific missions and descriptions of these agencies to offer insight into the types of agencies that tend to be most transparent to the public. Achieving an idea of what types of agencies may be more or less responsive to requests can help us determine what leads to government transparency.

3 Topical Predictors of Transparency

The main challenge to transparency research is that this field is overtly studied as an independent variable (Cucciniello, et al. 2016). Academics and policymakers see greater transparency as generally beneficial and leading to better outcomes in government. The specific reasons according to Posen (2002), include trust, predictability, reduced noise in markets, credibility, and coordination. What is less often studied is the origins of transparency and what determines how governments and their institutions are more likely to be open to the public about their operations. By studying what determines openness in institutions and agencies, it can help reveal incentives and drives for these operations to encourage the benefits that transparency has to offer.

Determinants of Access to Information

Public Facing

Some agencies are structurally and politically more public-facing where they are more often under the surveillance of citizens. Each agency has their own reputation to the public along with their own campaigns such as those witnessed in the media. It is often more beneficial for these agencies to have a rather good reputation, one of which can be achieved by being more open and responsive to the public (Geraats, 2005). Moffitt (2010) argues that bureaucrats

actively pursue publicity and public participation for tasks that risk implementation failure. This can be witnessed by the Food and Drug Administration (FDA) that had sought public advice for its riskiest tasks. Such advice is associated with a lower probability of subsequent Congressional oversight and with a greater probability of subsequent agency information campaigns such as safety label provisions.

Not only may agencies be incentivized to be more public-facing by avoiding Congressional scrutiny but also because of the pressure of scrutiny from the public. In a study conducted on bureaucratic discretion under Right to Information laws (RTI),— similar to Freedom of Information laws— Rodríguez (2018) found that only citizens who know the RTI law and invoke its existence have a greater likelihood of obtaining an answer from bureaucrats. These findings suggest that "public campaigns to promote citizens' awareness of RTI laws, not only would increase requests, but also governments' responsiveness regarding RTI requests" (Rodríguez, 2018). Agencies that may be already more open by default may be already more open and transparent. According to these previous studies, reputation and default practices of an agency can be indicative of their transparency levels. The factor will be included in my model to study how it affects the relationship between topics and transparency rates.

Agency Size

Because I am looking at transparency at the federal agency level, it is important to take into account the potential implications of the bureaucratic process. Specifically, what may affect how an agency can respond to FOIA requests may be influenced by the resources an individual agency may have. Bureaucratic efficiency is key in carrying out the essential function and

missions of an agency. In particular, what may be attributed to an agency's efficiency is the number of staff members each agency has. Carroll (1967) notes that the size of an agency can serve as indicators of extra resources and appear important to the innovative process. In a study examining stock exchange scores in Istanbul, Aksu and Kosedag (2006) found that of the determinants of Transparency & Disclosure scores (T&D) was the size of the firm. They argue that larger firms are more visible and are thus more closely followed by financial intermediaries, and are overall more politically sensitive (Watts and Zimmerman 1986). Based on previous findings, the size of an agency is important to take into account to understand how it may influence how topics of agencies predict transparency.

Fiscal Pressure

One of the most pertinent complaints to FOIA law is not in the law itself, but often the monetary expense of its administration (Wagner, 2017). Agencies incur substantial costs with FOIA access mechanisms from the process to administer the final reports, to the litigation fees that come with unsatisfactory appeal responses on the part of the requestor. In 2018, the Federal Bureau of Investigation (FBI) accrued a total of $12,147,220,000.00 just from litigation-related fees alone (foia.gov citation). Moreover, these costs accumulated are hardly counterbalanced with the processing fees incurred by the agency to the requestor. A large part of the procedural costs is for agencies to review reports line by line, costs of which cannot be legally charged to the requestor. Requests overall are rather inexpensive and fees incurred may be waived under circumstances such as holding a student status (*Sack v. U.S. Department of Defense, 2016*).

Many scholars argue that these high expenses may be concerning for the potential detrimental effects on transparency and public access to information. One famous critic was the late Justice Antonin Scalia who denounced FOIA's ineffectiveness, inefficiency, and high expenses in a 1982 essay. He notes that the "'free lunch' aspect of the FOIA is significant not only because it takes money from the Treasury that could be better spent elsewhere, but also because it brings into the system requests that are not really important enough to be there (Scalia, 1982, pp. 17). Justice Scalia may not have been so critical of the idea of providing the freedom of access to information in general, but to have the government pay for this freedom was a concept that was perhaps too idealistic. Cate et al. (1994) concluded that these high procedural costs may pose an "enormous burden on private individuals and organizations, administrative agencies, and the courts." These burdens described by Cate et al. could have unintended consequences such as increasing the backlog of requests and the decrease in transparency on reports could lead to larger political and economic consequences. Furthermore, in examining the Clinton administration's support for FOIA, Sinrod (1994) was skeptical of the financial undertaking, claiming that the FOIA backlog problem will not be solved without a serious commitment of further government personnel, equipment and monetary resources." Based on these speculated consequences for the costs incurred on the part of agencies for FOIA requests, the net costs of each will be added to the model.

Evaluating these past claims are essential to contributing to the STM model. Because my independent variable will be induced based on extant text, my resulting analysis on the topics produced will require context from the larger domain of transparency research. The purpose in adding these aforementioned variables as constants is twofold. First, witnessing consistent

patterns on transparency will offer insight into how my model is appropriate as a theoretical framework for the research field on government transparency. Topics on the agency descriptions of their functions and missions are meaningful in determining how more open they are in accessing information per requested. As mentioned, determinants to transparency are often conditional. Placing these factors in a larger scope will show how broadly we can apply these determinants. Second, the STM model can provide more clarity in determining how these previously studied variables may influence transparency in the context of this broader model. Variables of public exposure, agency size, and fiscal pressure are often studied in the context of public administration in local government. Applying this new approach of STM to transparency research using these variables will provide insight how they may affect transparency at the federal agency level.

Structural Topic Modeling: Establishing a Theoretical Framework

Topic modeling is a statistical and computational technique for discerning information about the contents of a large corpus of documents (Blei, 2012). A topic model uncovers patterns of word co-occurrence across the corpus, yielding a set of word clusters, together with associated probabilities of occurrence, which constitute the 'topics'. Topic modeling is often used to gain insight into a collection of documents, such as online discussion (Mishler, 2015), newspapers (DiMaggio, et al. 2013), and research papers (Wang & Blei, 2011) to name a few. This method of analysis enables the user to understand trends of the corpus by sorting into labeled groups that can be helpful as it does not require reading the original text. The standard topic modeling uses the Latent Dirichlet Allocation technique (LDA). This technique however may have its limits as

LDA makes a statistical assumption that all texts in the modeled corpus are generated by the same underlying process (Blei, 2002). It is not ideally suited to examining differences in topical content that are affected by external variables such as author identity or time of writing.

To address this limitation, structural topic modeling (STM) is a variant of LDA that is designed to represent the effect of external variables on both topical content and topical prevalence (Roberts, et al. 2016). Topical content refers to the probabilities associated with words in each topic, while topical prevalence refers to the proportions of different topics that occur within documents. The external variables can consist of any metadata that distinguishes one text from another such as the year of production or the size of the corresponding agency. STM effectively allows us to estimate a series of regression models that treat the prevalence of each identified topic as an outcome variable whose variation is then modeled as a function of the FOIA response rate variable and other explanatory variables.

Using STM, I will investigate U.S. agency Wikipedia pages as my set of documents to create my corpus to detect which types of agencies are more likely to have higher levels of transparency. By taking data from the response and process rates for FOIA requests for each individual agency as my metadata, I can use this method of analysis to reveal underlying patterns and topics that predict transparency. I then go to show how previously studied determinants affect my original model to then create a theoretical model for how types of agencies may lead them to be more or less transparent.

4 An Investigation Into FOIA Requests: Research Design

Data

Independent Variable

Text Data for Structural Topic Modeling

To assemble the data, I take each agency and search their main Wikipedia.com pages for textual data on their overall description. This data was collected in November 2019 to March 2020. My corpus consisted of 356 Wikipedia pages on U.S. agencies and their reporting executive agencies. When agencies or components do not have a designated Wikipedia page available, the agency's home government website (.gov) and their "About" or "Mission and Values" page text were substituted in place. Some agencies are decentralized by their regional locations such as the Department of Education that direct requests across their eleven regional locations. These components were omitted due to the lack of textual data available. Agencies that had neither Wikipedia page articles nor government home pages were also omitted.

Dependent Variable

Transparency

To measure the prime research subject of transparency, I draw on annual data from FOIA requests submitted for 119 agencies and their components in the year 2018 from FOIA.gov. Agency processing and response rates can reveal compliance to FOIA law and how accessible these agencies are to disclosing records directly requested from the public. The data available provides the pending status of reports made at the beginning and end of each fiscal year, the number of requests received each year, and the number of requests processed for each year. To

measure the response rates of each agency, I take the total number of requests processed divided by the total number of requests received. This resulting continuous variable for each agency can take on a value greater than 1 as some agencies may be responding to requests from previous years.

Control Variables:

Public Facing

To measure how agencies may be by default more public facing, I attain data from FOIA.gov on each agency's number of proactive disclosures. Proactive disclosures are records made publicly available by agencies without the requirement of a specific FOIA request. Some information disclosed by agencies is preemptively required by law including final opinions and orders, specific policy statements, certain administrative staff manuals and frequently requested records. Other information is up for each agency to decide. I analyze how each agency voluntarily shares data. This numerical variable indicates whole datasets or reports that are openly and actively available by each agency.

Agency Size

Agency size is a numerical variable that represents the total number of full-time FOIA staff employees for each agency. This data is also available on FOIA.gov.

Fiscal Pressure

Fiscal pressure is measured as the net costs incurred for each agency in 2018. The data was taken from FOIA.gov. This is calculated from an accumulation of the total costs (of which consist of both procedural costs and litigation-related costs) minus the total amount collected by the agency. The resulting numerical variable can take on values ranging from both negative and positive. For ease of calculations, this value is measured in tens of millions.

Model Inference

Using a bag-of-words approach, STMs discard the ordering of words within documents and only seek to discover topics based on the observed correlations between words. The assumption behind each document in the corpus is to be a mixture of multiple, correlated topics, each with characteristic words and its own distribution. I focus my analysis on the prevalence of specific topics, as a way of capturing the values and functions devoted to agency topics across the Wikipedia web pages. The STM algorithm requires some basic input parameters, such as the number of K topics and the topic distribution. At this point of analysis, Roberts, Stewart and Tingley (2014) note that "[t]here is no right answer to the appropriate number of topics. More topics will give more finegrained representations of the data at the potential cost of being less precisely estimated. […] For a small corpora (a few hundred to a few thousand) 5–20 topics is a good place to start." My corpus contains 356 agency descriptions. Based on this, I select a topic number of 20 for my primary analysis.

5 Trends of Topical Predictors

In the following subsections I present my findings on the data classification and discuss potential trends.

Discovering Topics

In total, I collect 274 Wikipedia.com page descriptions and 82 official government website descriptions to create a corpus total of 356 documents, after omitting missing values. The final STM model identifies the 20 latent topics that best characterize the webpage descriptions of U.S. federal agencies. Each topic represents an underlying word distribution across each word in the corpus. I obtain two types of posterior distributions, a topic distribution of each document, and a word distribution of each topic. I show the topic distribution of each document as word clouds in Figure 2, wherein the size of each word is in proportion to its probability i.e., apparent larger words are more frequent per topic where words smaller in size are less. The words with the highest probability of assignment to a specific topic can be considered representative of that topic.

The result is used as a classification scheme for areas in government agency functioning. For example, Topic #1: "offic, energi, doe, technolog, nation, oper, nuclear,…." is mostly related to energy and technology and Topic #5: ‘‘command, oper, unit, state, forc, militari, marin… " centers on military operations and command.

Topic #1 Energy & Technology	Topic #2 Security & Benefits	Topic #3 Review & Regulation	Topic #4 Legal Justice
			
Topic #5 Military Command	Topic #6 Safety Oversight	Topic #7 Economy & Business	Topic #8 Enforcement
			
Topic #9 Land & Conservation	Topic #10 Public Service	Topic #11 Banks & Credit	Topic #12 Veterans
			
Topic #13 Health	Topic #14 Inspection	Topic #15 Memorial	Topic #16 Defense
			

Topic #17 Management	Topic #18 Aerospace	Topic #19 Intelligence & Security	Topic #20 Education

Fig. 2. Word Cloud of Topic #1–Topic #20.

To further substantively interpret the meaning of each topic based upon these probabilistic word assignments, the STM model features a measure of the weighted harmonic mean of each word's rank in terms of exclusivity and frequency (FREX). By examining special words to each topic, I can identify how the language was sorted and apply labels for each given topic. Based on the word distributions and examining the FREX of each topic formulated, I may link topics with some specific government function areas or issue intuitively. The FREX calculation offers more insight compared to words with higher probability as it gives greater weight to words that are less frequent in other topics, making it easier to discern and apply labels to each topic.

Topic	*Word Probability and FREX*	*Label*
1	Highest Prob: offic, energi, doe, technolog, nation, oper, nuclear FREX: doe, energi, nuclear, technolog, patent, privaci, wast	Energy & Technology
2	Highest Prob: benefit, secur, retir, program, headquart, employe, board FREX: benefit, retir, paid, social, rrb, end, railroad	Security & Benefits
3	Highest Prob: state, feder, unit, agenc, commiss, board, govern FREX: commiss, postal, board, regulatori, court, branch, carrier	Review & Regulation

4	Highest Prob: state, depart, unit, offic, justic, attorney, law FREX: attorney, solicitor, justic, divis, crimin, civil, labor	Legal Justice
5	Highest Prob: command, oper, unit, state, forc, militari, marin FREX: ussouthcom, marin, command, unifi, coast, oil, forc	Military Command
6	Highest Prob: agenc, health, safeti, mine, act, state, unit FREX: mine, safeti, injuri, environment, diseas, health, food	Safety Oversight
7	Highest Prob: state, agenc, unit, develop, busi, econom, provid FREX: econom, agricultur, busi, help, export, statist, small	Economy and Business
8	Highest Prob: bureau, state, feder, depart, unit, intern, enforc FREX: bureau, census, revenu, tax, alcohol, tobacco, treasuri	Enforcement
9	Highest Prob: depart, land, state, interior, agenc, manag, indian FREX: interior, land, indian, conserv, reclam, congression, legisl	Land & Conservation
10	Highest Prob: servic, nation, feder, agenc, state, public, unit FREX: preserv, scholarship, park, histor, telecommun, noaa, passeng	Public Service
11	Highest Prob: bank, feder, insur, system, credit, market, agenc FREX: bank, insur, credit, market, deposit, reserv, union	Banks & Credit
12	Highest Prob: offic, veteran, servic, will, mission, work, polici FREX: judg, opportun, complaint, will, vision, practic, ogc	Veterans
13	Highest Prob: health, secretari, administr, transport, program, offic, servic FREX: transport, health, care, secretari, vha, administr, highway	Health
14	Highest Prob: general, oig, inspector, offic, investig, program, report FREX: oig, fraud, inspector, audit, general, investig, abus	Inspection
15	Highest Prob: state, nation, unit, cemeteri, govern, congress, agenc FREX: cemeteri, memori, copyright, usg, monument, mediat, archiv	Memorial
16	Highest Prob: defens, depart, secretari, forc, agenc, unit, state FREX: defens, dod, forc, air, contract, navi, armi	Defense
17	Highest Prob: manag, offic, financi, provid, inform, depart, agenc FREX: financi, manag, acquisit, foia, perform, chief, department	Management
18	Highest Prob: center, nasa, space, mission, oper, research, flight FREX: nasa, flight, spacecraft, space, goddard, rocket, center	Aerospace
19	Highest Prob: intellig, agenc, director, secur, unit, nation, state FREX: cia, fbi, intellig, nsa, dia, violenc, director	Intelligence & Security
20	Highest Prob: educ, program, student, depart, secretari, research, feder FREX: educ, student, adult, school, aid, research, evalu	Education

Fig 1. Results with twenty estimated topics, the most frequent and exclusive (FREX) words for each, and assigned labels.

Topics 1, 4, 5, 7, 8, 11-13, and 15-20 are the most intuitive to interpret and apply labels where the FREX words reveal similar concepts. They are labeled accordingly Energy & Technology, Legal Justice, Military Command, Economy & Business, Land & Conservation, Banks & Credit, Veterans, Health, Memorial, Defense, Management, Aerospace, Intelligence & Security, and Education. The similarity of words for each topic made it appropriate to apply these labels. For example, the Legal Justice topic includes words like attorney, justice, civil, and criminal, ideas that are all along the same issue. Many of these labels are inspired from words directly from the most frequent words in each topic. Moreover, many of the agencies in this study are branched from parent departments that decentralize into sub-agencies that all support a similar purpose. For example, the Department of Energy, the Department of Education, and NASA are the exclusive parent departments for the branching agencies most associated with the Energy and Technology, the Education, and the Aerospace topics. A complete list of the top five most representative agencies of each topic can be seen in Figure 3.

Labels for topics 2, 3, 6, 8, 10, and 14 were drawn more abstractly such as with the Security & Benefits topic (Topic 2) which revolves around the words benefit, security, retire, program, employee, and paid. At first glance, it is easily inferred that these words are taken from agencies that manage social security benefits, but to be sure it is helpful to see what agencies are prevalent for each topic. To create Topic 2, descriptions are taken from agencies that focus on retirement benefits including the Railroad Retirement Board, the Social Security Administration, and the Pension Benefit Guaranty Corporation. Therefore, Topic 2 is referred to as Security & Benefits. This same method is used for Topic 6, Safety Oversight. Many agencies are created to be designated watchdogs of health and safety over specific areas like food production, the

environment, and occupational workplaces. Topic 6 specifically had picked up on this pattern of agencies made for safety oversight, hence the name. Thus, examining which agencies compose the highest proportion of each topic is useful for labeling purposes as well as seeing their transparency measures. Many labels are made from the culmination of FREX measures and assessing which agencies had largely contributed to creating these topics.

In examining the text that composes each topic in Figure 2, I find similar concepts and functions that are part of many if not all agencies, such as with Topic 14 where words like inspector, office, investigate, program, and report exist due to many departments having their own Office of the Inspector General (OIG). These words were frequent enough in the overall corpus to create a topic to take this into account. This is similar to the Review & Regulation topic (Topic 3) where commissions and external review boards exist for many agencies and thus appear as a topic in the resulting model. These topics may be specific fields of governance but are essential functions that can reveal how they enable agencies in being more or less transparent.

Topic	*Top 5 Representative Agencies*	*Parent Department*
1: Energy and Technology	Idaho Operations Office National Energy Technology Laboratory Office of Scientific and Technical Information Savannah River Office Golden Field Office	Dept of Energy (DOE) DOE DOE DOE DOE
2: Benefits and Security	Railroad Retirement Board Social Security Administration Board of Veterans' Appeals Federal Retirement Thrift Investment Board Pension Benefit Guaranty Corporation	RRB SSA Veteran Affairs (VA) FRTIB PBGC
3: Review and Regulation	Federal Trade Commission Federal Energy Regulatory Commission United States Parole Commission Foreign Claims Settlement Commission United States International Trade Commission	FTC FERC Dept of Justice (DOJ) DOJ USITC

4: Legal Justice	Department of Justice Civil Rights Division Community Relations Service Department of State United States Attorney General Wage and Hour Division	DOJ DOJ Dept of State (DoS) DOJ Dept of Labor (DOL)
5: Military Command	Special Operations Command Indo-Pacific Command Southern Command Marine Mammal Commission United States Coast Guard	Dept of Defense (DOD) DOD DOD MMC Homeland Security (DHS)
6: Safety Oversight	Federal Mine Safety & Health Review Commission Environmental Protection Agency Headquarters Food and Drug Administration Food Safety and Inspection Service Mine Safety and Health Administration DOL	FMSHRC EPA Health & Human Services (HHS) DOA DOL
7: Economy & Business	US Agency for International Development Foreign Agricultural Service United States African Development Foundation Bureau of Economic Analysis Tennessee Valley Authority	USAID Dept. Agriculture (DOA) US ADF Dept. Commerce (DOC) TVA
8: Enforcement	Internal Revenue Service Alcohol and Tobacco Tax and Trade Bureau Office of Public Affairs Appraisal Subcommittee United States Mint	Dept. Treasury (USDT) USDT DOL ASC USDT
9: Land and Conservation	Office of Legislation and Congressional Affairs Congressional and Intergovernmental Affairs Bureau of Land Management Bureau of Reclamation Office of Navajo and Hopi Indian Relocation	Dept of Education (ED) DOL DOI DOI ONHIR
10: Public Service	National Railroad Passenger Corporation (Amtrak) Presidio Trust Institute of Museum and Library Services United States Postal Inspection Service Corporation for National and Community Service	NRPC PT IMLS USPS CNCS
11: Banks and Credit	Selective Service System Federal Deposit Insurance Corporation National Credit Union Administration Federal Open Market Committee Southwestern Power Administration	SSS FDIC NCUA FOMC DOE
12: Veterans	Office of the General Counsel Office of Information and Technology Office of Diversity and Inclusion Office of Disability Employment Policy Office of Resolution Management	Federal Labor Relations Authority (FLRA) VA VA DOL VA
13: Health	Office of the Secretary (of Transportation) National Institutes of Health Federal Transit Administration	DOT HHS DOT

	Office of Assistant Secretary for Health Veterans' Employment and Training Service	HHS DOL
14: Inspection	Inspector General Office of Inspector General Treasury Inspector General for Tax Administration Office of Accountability and Whistleblower Protection Office of Inspector General	FLRA National Labor Relations Board (NLRB) USDT VA Legal Services Corporation (LSC)
15: Memorial	National Cemetery Administration United States Institute of Peace United States Copyright Office United States Geological Survey National Archives and Records Administration	VA USIP CO Dept of Interior (DOI) NARA
16: Defense	Secretary of Defense/Joint Staff United States Department of the Air Force Department of the Navy Strategic Command Defense Threat Reduction Agency	DOD DOD DOD DOD DOD
17: Management	Office of the Chief Financial Officer Office of the Chief Information Officer Office of Finance Office of the Chief Financial Officer Office of Management	DOL ED VA ED VA
18: Aerospace	Armstrong Flight Research Center Kennedy Space Center Goddard Space Flight Center Ames Research Center NASA Management Office - Jet Propulsion Lab	NASA NASA NASA NASA NASA
19: Security & Intelligence	National Security Agency Central Intelligence Agency Defense Intelligence Agency National Geospatial-Intelligence Agency Immigration and Customs Enforcement	DOD CIA DOD DOD DHS
20: Education	Office of Career, Technical & Adult Education Institute of Education Sciences Office of the Under Secretary Federal Student Aid Office of the Deputy Secretary	ED ED ED ED ED

Figure 3. Agencies highly associated per topic along with their parent departments.

Topics and Predictions on Transparency

The complete list of inspected topics, control variables, and their estimated effects on transparency rates is available in Table 1. The most prevalent topic with the highest corresponding rate of transparency (measured as response per FOIA request received) is Banks and Credit (1.131) followed by the Memorial topic (1.104), Energy and Technology (1.081), Review and Regulation (1.074), and Safety Oversight (1.024). The least transparent topics are dedicated to the Veterans topic (0.702), Education (0.784), and the Land and Conservation (0.794). It is notable that the spread of the estimated effect that these topics have on transparency is not dramatically different, nor is it insignificant enough to ignore. The difference from the largest estimated effect from the smallest is estimated to be 0.3473, or a 34 percentage point difference in response rate divided by caseload. A list of the order of greatest to least estimated effect can be found in Figure 4.

The topics with the highest ratings for more processed FOIA requests per received request are best explained by examining the agencies that are highly associated with each topic. Referring back to Fig. 2 and inspecting the Banks and Credit topic,— composed of words including bank, insurance, market, reserve, and deposit— it is inferred that types of agencies characterized as and operate revolving around Banks and Credit are predicted to process 1.31 reports per FOIA request received. Agencies highly associated with this topic include the Selective Service System (SSS), the Federal Deposit Insurance Corporation (FDIC), the National Credit Union Administration (NCUA), and the Federal Open Market Committee (FOMC). These agencies overall are designed to protect the credit and assets of citizens and are arguably made to

Table 1: Topical Predictors of Transparency

	Dependent variable:				
	FOIA Responses Processed Per Request Received				
	(1)	(2)	(3)	(4)	(5)
Topic 1 Energy and Technology	1.081***	1.083***	1.081***	1.081***	1.083***
	(0.082)	(0.083)	(0.083)	(0.083)	(0.083)
Topic 2 Security and Benefits	0.997***	0.997***	0.995***	0.997***	0.989***
	(0.103)	(0.103)	(0.103)	(0.103)	(0.104)
Topic 3 Review and Regulation	1.074***	1.074***	1.074***	1.074***	1.074***
	(0.063)	(0.063)	(0.063)	(0.063)	(0.063)
Topic 4 Legal Justice	0.894***	0.894***	0.891***	0.893***	0.896***
	(0.065)	(0.065)	(0.065)	(0.066)	(0.066)
Topic 5 Military Command	0.902***	0.901***	0.901***	0.901***	0.901***
	(0.082)	(0.082)	(0.082)	(0.082)	(0.082)
Topic 6 Safety Oversight	1.024***	1.024***	1.020***	1.023***	1.027***
	(0.077)	(0.077)	(0.078)	(0.079)	(0.080)
Topic 7 Economy and Business	1.007***	1.007***	1.006***	1.007***	1.005***
	(0.067)	(0.067)	(0.067)	(0.067)	(0.067)
Topic 8 Enforcement	0.981***	0.981***	0.980***	0.981***	0.981***
	(0.088)	(0.089)	(0.089)	(0.089)	(0.089)
Topic 9 Land Conservation	0.794***	0.794***	0.793***	0.794***	0.791***
	(0.091)	(0.091)	(0.091)	(0.091)	(0.091)
Topic 10 Public Service	0.875***	0.875***	0.875***	0.875***	0.875***
	(0.083)	(0.083)	(0.083)	(0.083)	(0.083)
Topic 11 Banks and Credit	1.131***	1.131***	1.131***	1.131***	1.131***
	(0.095)	(0.095)	(0.095)	(0.095)	(0.095)
Topic 12 Veterans	0.702***	0.701***	0.702***	0.702***	0.701***
	(0.084)	(0.084)	(0.084)	(0.084)	(0.085)
Topic 13 Health	0.944***	0.944***	0.942***	0.944***	0.940***
	(0.074)	(0.074)	(0.075)	(0.074)	(0.075)
Topic 14 Inspection	0.938***	0.938***	0.939***	0.938***	0.938***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.072)
Topic 15 Memorial	1.104***	1.115***	1.103***	1.104***	1.115***
	(0.090)	(0.095)	(0.090)	(0.090)	(0.095)
Topic 16 Defense	0.913***	0.913***	0.907***	0.913***	0.901***
	(0.075)	(0.075)	(0.078)	(0.076)	(0.078)
Topic 17 Management	0.927***	0.927***	0.927***	0.927***	0.926***
	(0.097)	(0.097)	(0.097)	(0.097)	(0.097)
Topic 18 Aerospace	1.014***	1.014***	1.014***	1.014***	1.014***
	(0.092)	(0.092)	(0.092)	(0.092)	(0.092)
Topic 19 Intelligence and Security	1.011***	1.011***	1.002***	1.009***	1.011***
	(0.084)	(0.084)	(0.089)	(0.091)	(0.091)
Topic 20 Education	0.784***	0.784***	0.784***	0.784***	0.784***
	(0.091)	(0.091)	(0.091)	(0.091)	(0.091)
Proactive Disclosures		−0.000			−0.000
		(0.00000)			(0.00000)
Agency Size			0.0001		0.001
			(0.0004)		(0.001)
Net Costs				0.0002	−0.004
				(0.004)	(0.008)
Observations	350	350	350	350	350
R^2	0.928	0.928	0.928	0.928	0.928
Adjusted R^2	0.924	0.924	0.924	0.924	0.923
Residual Std. Error	0.275 (df = 330)	0.275 (df = 329)	0.275 (df = 329)	0.275 (df = 329)	0.276 (df = 327)

be transparent from its inception. For example, the FDIC was born out of a series of collapsing banks during the Great Depression. It is an independent agency made with the responsibility to insure deposits in banks and thrift institutions. The FDIC's mission is "to maintain stability and public confidence in the nation's financial system by: (1) insuring deposits, (2) examining and supervising financial institutions for safety and soundness and consumer protection, and (3) managing receiverships." Since its establishment in 1933, maintaining this goal for public confidence is key to its purpose and one that cannot be "without the trust and confidence of the institutions we supervise and insure, our fellow regulators, and the public" as stated by FDIC Chairman Jelena McWilliams (FDIC, n.d.). Committees and administrations designed to ensure the safety of people's assets are also by default designed to support economic confidence. An essential way to maintain this is to be open about the agency's information.

Aside from analyzing the agencies that compose this topic, the finding that Banks and Credit has the highest rates of transparency is consistent with previous empirical literature review, where a positive relationship between transparency and financial management is consistently supported (Cucciniello, 2010; Benito and Bastida, 2009). This is based on the theory that government transparency is most often a combination of both political, economic, and administrative factors. While it is ideal for many agencies to have goals for transparency and openness, how this is exactly carried out is important for how the public may actually feel the effects. Fiscal pressures on the public may incentivize citizens to hold their government more accountable where they desire to maximize the public services they receive while minimizing the amount they pay for it (Alcaraz-Quiles et al., 2015). Due to the sensitivity that the public has over their personal funds and the trust required to allow federal agencies to manage them,

Rank	*Model 1*	*Model 2 (Controlling for Proactive Disclosures)*	*Model 3 (Controlling for Agency Size)*	*Model 4 (Controlling for Fiscal Pressure)*	*Model 5 (All Control Variables)*
1	11: Banks & Credit (1.131)	11: Banks & Credit (1.131)	11: Banks & Credit (1.131)	11: Banks & Credit (1.131)	11: Banks & Credit (1.131)
2	15: Memorial (1.104)	15: Memorial (1.115)	15: Memorial (1.103)	15: Memorial (1.104)	15: Memorial (1.115)
3	1: Energy & Tech (1.081)	1: Energy & Tech (1.083)	1: Energy & Tech (1.081)	1: Energy & Tech (1.081)	1: Energy & Tech (1.083)
4	3: Review & Regulation (1.074)	3: Review & Regulation (1.074)	3: Review & Regulation (1.074)	3: Review & Regulation (1.074)	3: Review & Regulation (1.074)
5	6: Safety Oversight (1.024)	6: Safety Oversight (1.024)	6: Safety Oversight (1.020)	6: Safety Oversight (1.023)	6: Safety Oversight (1.027)
6	18: Aerospace (1.014)	18: Aerospace (1.014)	18: Aerospace (1.014)	18: Aerospace (1.014)	18: Aerospace (1.014)
7	19: Intelligence & Security (1.011)	19: Intelligence & Security (1.011)	19: Intelligence & Security (1.002)	19: Intelligence & Security (1.009)	19: Intelligence & Security (1.011)
8	7: Economy & Business (1.007)	7: Economy & Business (1.007)	7: Economy & Business (1.006)	7: Economy & Business (1.007)	7: Economy & Business (1.005)
9	2: Security & Benefits (0.997)	2: Security & Benefits (0.997)	2: Security & Benefits (0.995)	2: Security & Benefits (0.997)	2: Security & Benefits (0.989)
10	8: Enforcement (0.981)	8: Enforcement (0.981)	8: Enforcement (0.980)	8: Enforcement (0.981)	8: Enforcement (0.981)
11	13: Health (0.944)	13: Health (0.944)	13: Health (0.942)	13: Health (0.944)	13: Health (0.940)
12	14: Inspection (0.938)	14: Inspection (0.938)	14: Inspection (0.939)	14: Inspection (0.938)	14: Inspection (0.938)
13	17: Management (0.927)	17: Management (0.927)	17: Management (0.927)	17: Management (0.927)	17: Management (0.926)
14	16: Defense (0.913)	16: Defense (0.913)	16: Defense (0.907)	16: Defense (0.913)	16: Defense (0.901)
15	5: Military Command (0.902)	5: Military Command (0.901)	5: Military Command (0.901)	5: Military Command (0.901)	5: Military Command (0.901)
16	4: Legal Justice (0.894)	4: Legal Justice (0.894)	4: Legal Justice (0.891)	4: Legal Justice (0.893)	4: Legal Justice (0.896)
17	10: Public Service (0.875)	10: Public Service (0.875)	10: Public Service (0.875)	10: Public Service (0.875)	10: Public Service (0.875)
18	9: Land and Conservation (0.794)	9: Land and Conservation (0.794)	9: Land and Conservation (0.793)	9: Land and Conservation (0.794)	9: Land and Conservation (0.791)
19	20: Education (0.784)	20: Education (0.784)	20: Education (0.784)	20: Education (0.784)	20: Education (0.784)
20	12: Veterans (0.702)	12: Veterans (0.701)	12: Veterans (0.702)	12: Veterans (0.702)	12: Veterans (0.701)

Figure 4. List of topics in order of greatest estimated rate of process responses per FOIA request received with control variables.

agencies that have this fiscal responsibility become by default more open about their information, as found here in the model. Not doing so not only can induce pressure from public outcry when peoples' monies are mishandled, but can introduce broader undesirable political and economic consequences. For instance, Roberts (2012) reported that the United States underwent major financial cuts towards fiscal transparency measures in 2007. He suggests that this, at least in part, is what brought on the subsequent financial crisis as institutions had operated in the dark, and the result of the meltdown was even less federal support for transparency measures.

These results may be consistent with the Banks and Credit topic, but differ in how they relate to the eighth most transparent topic of Economy and Business (1.007). The agencies highly associated with this topic include the United States Agency for International Development (USAID), Foreign Agricultural Service (FAS), United States African Development Foundation (USADF), and the Bureau of Economic Analysis (BEA). While these agencies are also responsible for finances, they are oriented towards larger economic operations on a national level than the Banks and Credit agencies that are often more interactive with citizens. The BEA, for instance, is dedicated to providing macroeconomic and industry statistics, most notably reports about the gross domestic product (GDP) of the U.S. on a yearly and quarterly basis (BEA.gov, n.d.). As an agency with a jurisdiction at a national level, there is not as strong of a need to release requested records since there is less of an incentive to maintain accountability to citizens as compared to agencies of the Banks & Credit type.

The second most transparent is the topic on Memorials (Topic 15), represented by agencies like the National Cemetery Administration (NCA), the United States Institute of Peace (USIP), the Copyright Office (CO), United States Geological Survey (USGS), National Archives

and Records Administration (NARA). When taking a closer look at these agencies, they are all driven to honor public values. For example, most obviously, the NCA is made to honor military service by providing a "dignified burial and lasting memorial for veterans" (VA.gov, 2008). The USIP works to promote peace, specifically, the "nonviolent prevention and mitigation of deadly conflict abroad" (USIP.gov, n.d). The Copyright office is designed to protect titles and property and NARA is devoted to preserving historic records. The topic of Memorial it appears is apportioned to agencies that build on values like right to property, American history, peaceful relations, and veterans. Thus, these agencies are designed to be interactive with the public in order to promote these values to U.S. citizens and are therefore more open about the information they may keep on record.

Analyzing the bottom tier, the topic of Veterans has the lowest rate of transparency, which may be due to more administrative reasons than anything else. It is also important to note that this topic is composed based on a substantial part of the corpus that has agency descriptions that belong to the parent department of the VA. Nonetheless, this is a rather interesting finding in that the VA itself has historically struggled in their performance with processing FOIA requests. Every few years the Center for Effective Government, a think tank organization focusing on government transparency practices, releases scorecards on how agencies perform in processing FOIA requests. Their criteria for grading is based on the presence of clear agency rules guiding the release of information and communication with those requesting information; quality and "user-friendliness" of the agency's FOIA website; and the timely completion of processing (Moulton, 2015). Among the lowest scores was the Department of Veteran Affairs where they had received an "F" in 2014 and a "D" in 2015. This finding supports the theory of the

exogenous nature of what leads to greater transparency. As mentioned before, how access to information is realized is based on largely political, economic, and administrative variables where some factors stemming from these categories are more influential than others. Due to the VA administration itself lacking the resources and processing efficiency is a possible explanation for why it ranks last in this model. This issue in agency administration can help explain similar patterns in the topic of Education. What both of these topics have in common is that they are prevalent in the corpus based on the frequency of the parent departments as referred to in Fig. 3. This suggests that transparency rates can be due to these larger parent departments themselves. The Department of Education had recently gone under review by the Office of Government Information Services (OIG) on their FOIA program and procedures. In a 2019 evaluation, OGIS reported that (1) the department's FOIA regulation had not been updated since 2010 and is out of compliance with the statute; (2) template letters and standard language were not in compliance with the statute and could be clearer; (3) technology that the department uses to administer FOIA is not seamless or being used to its fullest extent; and (4) management controls and a decentralized FOIA program challenge Education's effectiveness and efficiency (OGIS, 2019). Regulatory non-compliance appears to be the issue for the Department of Education to execute FOIA processing and can explain their low relative ranking.

Second to last in transparency ranking is the Land and Conservation topic (Topic 9) where it is predicted that agencies of this type process 0.794 requests per FOIA request received. Most representative agencies that stem from a variety of parent departments under this topic include the Office of Legislation and Congressional Affairs, the Congressional and Intergovernmental Affairs, Bureau of Land Management, Bureau of Reclamation, and the Office

of Navajo and Hopi Indian Relocation. What these agencies have in common are the mission for conserving and protecting land. Why this topic is rated so lowly in our model may be due to the fact that publicly funded land conservation operations are simply not disclosing information about these private land sites for security and safety reasons. These lands are conservatory with reasons such as for the purpose of protecting rare species or cultural artifacts. Disclosing their location and information may pose a danger where collectors know where to find them. The privacy of landowners is also an incentive to not disclose information to protect their location and other information about specific citizens (Coniff, 2019). As seen here it is the actual content of the information that is kept away to explain why the topic of Land Conservation may be rated so low due to the conflicting interests involved. A part of FOIA law is a list of nine exemptions for disclosure for certain types of information that can be harmful or invasive if released, and therefore rejected upon request. Among them are (1) information that is classified to protect national security; (2) information related solely to the internal personnel rules and practices of an agency; (3) information that, if disclosed, would invade another individual's personal privacy; and (4) geological information on wells. Exemptions based on privacy or even geological information of wells can easily be applied to the Land & Conservation topic. A full list of FOIA exemptions is available in Table 2.

A pattern worth noting is that topics on the more extreme ends of the transparency rankings are more specific and have also been easier to label with their straightforward concepts and similar issues. The same cannot be said with middle-tier-ranked topics that take from a more diverse set of agencies to compose the topic. This may suggest several key findings. One is that, as seen with topics like Energy & Technology and Veterans, these topics predominantly source

Table 2: FOIA Exemptions to Disclosure of Information

Exemption 1: Information that is classified to protect national security.

Exemption 2: Information related solely to the internal personnel rules and practices of an agency.

Exemption 3: Information that is prohibited from disclosure by another federal law.

Exemption 4: Trade secrets or commercial or financial information that is confidential or privileged.

Exemption 5: Privileged communications within or between agencies, including those protected by the:

- Deliberative Process Privilege (provided the records were created less than 25 years before the date on which they were requested)
- Attorney-Work Product Privilege
- Attorney-Client Privilege

Exemption 6: Information that, if disclosed, would invade another individual's personal privacy.

Exemption 7: Information compiled for law enforcement purposes that:

- 7(A). Could reasonably be expected to interfere with enforcement proceedings
- 7(B). Would deprive a person of a right to a fair trial or an impartial adjudication
- 7(C). Could reasonably be expected to constitute an unwarranted invasion of personal privacy
- 7(D). Could reasonably be expected to disclose the identity of a confidential source
- 7(E). Would disclose techniques and procedures for law enforcement investigations or prosecutions, or would disclose guidelines for law enforcement investigations or prosecutions if such disclosure could reasonably be expected to risk circumvention of the law
- 7(F). Could reasonably be expected to endanger the life or physical safety of any individual
-

Exemption 8: Information that concerns the supervision of financial institutions.

Exemption 9: Geological information on wells.

Source: FOIA.gov

from singular parent departments. Thus these transparency levels based on topics may be due to more about these agencies themselves than the mission that they set themselves to achieve. Alternatively, this finding may suggest that more narrow and specific topics that are easier to identify are also areas that are clearly more open or more secretive. This could reflect blanket policies, for example, information related to Legal Justice issues (ranked fifth to last in transparency) might require advanced review before releasing information to the public. More general topics such as Management, Inspection, and Security & Benefits, all of which are mid-ranked in terms of transparency, may include many different kinds of cases on which kind of information is acceptable to be disclosed to the public. The variation in the content of the records may discern the FOIA responsiveness for these types of agencies.

What is surprising about these findings is that Intelligence & Security is ranked rather high (seventh) for relative transparency, considering that information on national security is an official FOIA exemption. This may be explained by the fact that it ranks much higher than related topics like Defense and Military Command, both of which information of national security concern may be more often designated to these types of office. Also, information on issues concerning national security may pertain more towards current operations. Information can be kept so under wraps by Intelligence & Security agencies that the public is not even aware that there is information to request. Moreover, records kept by agencies like the National Security Agency (NSA), the Central Intelligence Agency (CIA), the Defense Intelligence Agency (DIA), and the National Geospatial-Intelligence Agency (NGIA) may have preexisting reputations of being security-sensitive that requesters are already deterred, fully aware of the exemptions and anticipate denial.

Influence of Control Variables on Topical Content

Public-Facing Agencies

Overall the control variables did not make large changes to our topics in their effects on transparency ratings. The number of proactive disclosures agencies had readily available had the least effect. For every agency that had a proactive disclosure counted by FOIA.gov, there was a minute negative (−0.000) effect on the number of processed responses per FOIA request received. This, however, is not statistically significant and can be due to chance error. This disproves my first hypothesis of how agencies that are more public-facing will show higher levels of transparency. The model here shows that there is no effect.

While having more or less proactive disclosures available may not determine transparency, holding it as a constant did have some effect on the prevalent topics. In this regard, this constant also made the least changes on the other topics as compared to the other control variables. Holding proactive disclosures constant had increased the effect for the second most and third most transparent topic on Memorial and Energy & Technology by 0.011 and 0.002, respectively. It also decreased the effect on transparency for the topic of Veterans and Military Command both by 0.001. Overall, I find that the topical content of agency descriptions is still predictive of the level of transparency, even when controlling for proactive disclosures. While more public-facing agencies (measured by the number of proactive disclosures of information) may be more transparent according to previous research, my model still captures how agencies categorized under more public-facing topics such as Banks & Credit and Memorial are more responsive to the public.

Agency Size

Agency size in itself as a variable had an estimated effect of 0.0001 which was also not statistically significant. Agency size alone does not make a difference on FOIA processing rates and disproves my original hypothesis. However, agency size has affected the most agencies in how they respond to FOIA requests when holding it constant as compared to the other control variables. Model 3 in Fig. 4 shows that it overall decreased the effect of eleven topics on predicting transparency rates, especially topics that had ranked higher to mid-tier across the transparency rankings. Only one topic had an estimated effect that was increased by 0.001. The topic of Inspection, ranked twelfth of the twenty on transparency, had bumped up to an estimated effect of 0.939, holding agency size constant. However, the relative ranking of agencies remains stable from the first model, with the exception of Economy & Business exceeding Intelligence & Security for seventh place. This is the only instance of which rankings change.

Financial Conditions

Fiscal pressure, as recalled, was measured by the number of dollars the agency had to pay to carry out FOIA requests and processes by the tens of thousands. Although financial condition is a determinant of the disclosure of public information (Baber 1983; Ingram 1984; Giroux and Deis 1993; Evans and Patton 1987; Robbins and Austin 1986; Cheng 1992; Laswad, Fisher, and Oyelere 2005), my findings show that an estimated effect of 0.0002 that is not statistically significant. Based on this, net costs for each agency does not alone determine the rate of transparency for an agency and disproves my third and final hypothesis. When holding net costs

constant, there are decreases in the effect for the fifth, seventh, fifteenth and sixteenth most transparent topics. Specifically, the estimated topics that decreased are Safety Oversight, Intelligence and Security, Military Command and Legal Justice.

Most notably, the rankings of the topics do not change when controlling for net costs. This is supportive evidence that the topical function of the bureaucracy is an important explanatory variable to include while predicting transparency.

6 What the Future Holds for Government Transparency

Discussion

This study investigates the main determinants of information transparency in U.S federal agencies. By using STM to inductively produce independent variables based on text data on agency descriptions, I identify how different types of agencies respond to FOIA requests. I apply and test the validity of empirical results of prior research and investigate the underlying variations in these previous studies in a linear regression model. The contextual research background of these previous findings provide meaningful inferences to the model as well as reveal how they may affect these different topics of agencies.

The findings obtained for the topics analyzed indicate that the characteristics of each topic predict the rate at which transparency may be achieved. Controlling for proactive disclosures, the size of the government agency examined, and the costs induced by requests do not substantially change the ordering of how the topics are related to transparency. My study shows how the different missions and functions of agencies can provide insight into the level of transparency provided by these agencies, where agencies that are financially responsible under

the Banks Credit topic exert more openness in their data. Moreover, agencies with specified watchdog roles also score higher on transparency as they categorize under the topics of Review & Regulations and Safety Oversight. This finding is different from what we see with the topic of Inspection which captures the wide range of parent departments that house Offices of Inspector Generals. Ranked twelfth out of the twenty topics, due to the wide range of other topics that are latent within this topic, its general nature places it lower than the other oversight topics like Review & Regulations and Safety Oversight.

Looking closer to how agencies may be more or less transparent based on their topic, my analysis captures the exogenous nature of transparency. Types of agencies have provided explanations for why when agencies have similar missions they also respond similarly to FOIA requests. Whether this be due to the similar values in the sensitive content that is captured by the topic, or if the topic captured agencies that have struggled with administrative issues, my primary results model provides a broader framework to incorporate the many previously studied factors on what determines transparency. In this regard, more variables not yet discussed in this study require further analysis.

Previously studied determinants of transparency have shown contradictory evidence. This is in part due to the fact that determinants can change on a case by case basis where there are certain conditions of which these factors may have a real effect (Rodríguez Bolívar, 2013). Differences in government level, the time of which specific laws are passed, how funding is allocated per fiscal year that in turn affects which laws are passed, etc. can change widely. This broader framework of assessing types of agencies helps reveal how certain factors may affect transparency through a wider lens that might be more consistent over time. In regards to this,

fully understanding variation in the topical prediction of transparency requires scrutiny to multiple possible sources of such variation, as well as the limitations of an approach based on online Wikipedia pages. My findings suggest a few sources of variation that are consistent with previous empirical research.

Before assessing the possible sources of variation, it is worth noting how these sources have come about in research. At a broader theoretical level, transparency is studied in the context of two central theories: principal-agent theory (also known as agency theory) and the legitimacy theory (Eisenhardt, 1989; Suchman, 1995). According to the principal-agent theory, government officials (the agents) act in the interest of citizens (the principal) and they are to be accountable for their actions as they represent the people (Lane, 2005). When officials do not have the same interests as citizens, they must be held accountable for their deviant actions to demonstrate that they have been acting according to their responsibilities (Lane, 2005). Therefore, greater transparency reduces information asymmetries and increases the degree of confidence and public trust in political actors, limiting conflicts between citizens and the government.

Legitimacy theory, on the other hand, describes transparency as about organizations' activities as a way to legitimate actions to their stakeholders and award legitimacy to organizations (Patten, 1992). If this organizational legitimacy is threatened, organizations will disclose information in an effort to reinvigorate it. Disclosure of information may be used to regain public confidence and enhance the organization's reputation (Pina, Torres, & Royo, 2010).

Factors that can potentially determine transparency fall along these lines of theory and can explain variations in topical predictions. First, topics can shift due to a host of political

factors. Examples of this include the political strength an agency may wield such that it does not need to cooperate with executives and can govern using its own bureaucratic power. In line with agency theory, agencies with such strength will tend to be less transparent (Alcaraz-Quiles, Navarro-Galera, and Ortiz-Rodríguez, 2014). Another example is the gender composition of an agency that can influence bureaucratic administration and FOIA processing procedures. Studies show that women's participation in leadership can improve the quality of information because they are more ethically minded than men (Khazanchi, 1995). Thus, women's participation in bureaucratic government can minimize the conflicts between the principal (citizens) and the agent (government). In accordance with this, Tejedor-Romero and Arujo (2015) find a positive and significant relationship between gender and transparency in public administration especially in times of financial crisis. Moreover, citizens themselves may be watching their government more closely which in turn influences how agencies may behave. Based on legitimacy theory, if agencies know they are being watched, they may be more likely to disclose information for a more favorable view of the agency. Higher voter turnout may be representative that citizens have stronger interests in government activities (Hollyer et al., 2011) and agencies may react accordingly based on this information.

Second, a host of economic factors may explain some variations across topics. The literature demonstrates a positive relationship between economic status and transparency in public administration (Hameed, 2005). This study has taken into account the personal fiscal pressures an agency may experience brought on by increasing FOIA requests. However, financial pressures may also be felt on the side of citizens that can in turn pressure agencies to be held more accountable if the public is unsatisfied with government performance. Times of

economic stress such as increased taxes or high unemployment rates can act as signs for citizens that their government is not working to their best interest. Consequently, to win back the favor of the public, agencies may be more transparent about their practices.

Third, beyond theories of transparency and its relationship with the public that it serves, administrative issues that have to do more with agency design are a factor that can vary topical predictions of transparency. As discussed in this study, the size and number of staff members is reasoned to provide an agency with more resources and in turn more procedural efficiency with responding to FOIA requests. Sufficient funding for these agencies to pay employees may also have an effect. Moreover, research has shown that agency structure can have an effect on responsiveness such that the extent of high-level personnel politicization affects the degree of how agencies may be cooperative with other political forces such as with the President, the public, and Congress (Berry, 2016; McCubbins, 1987). Political appointees arranged in the upper echelons of agency management hold legal discretion over agency operations and hold the role as an agent in accordance to the principals, which is the party that gave them their current position (Barry 2017). Thus, political agendas of larger government such as the president can affect agency operations including responsiveness to FOIA requests.

Finally, the specific content of information that is being requested may be a key factor in how agencies adjudicate on whether to grant or deny access. FOIA law provides particular exemptions for agencies to not be required to disclose certain kinds of information. However, going beyond this, if it is not the content that causes the agency to refuse it can be the requestor. Rodríguez (2018) found discriminatory biases of FOIA processing procedures where the requestor gender had influenced responses.

My findings of variations in the topics of agencies and how they may predict transparency rates complement existing research that discuss the potential for political and economic effects and fluctuating standards of accountability. I posit topical predictions of transparency as a highly related dimension of transparency and accountability research, as they can capture the content of which information is being disclosed as well as which agencies are similar in how they respond to requests. Topical variations in transparency can have both formal sources, due to new changes in FOIA law requirements or guidelines, or informal sources, including internal administration dynamics, public demands, and discriminatory biases against requestors. Assessing the model before us, many of these factors can offer explanations for why certain types of agencies are apparently more transparent than others. These measures of topics can be combined with existing measures to further research to account for agency typology within analyses of existing determinants.

Model Limitations and Future Work

Challenges for topic modeling include trusting the output in creating hypotheses (Ramage, et al. 2010). When intuitions do not match trends in the data, it is questionable how if it is the researcher's intuition that is faulty or if the model itself needs to change. Ultimately, a topic model's trustworthiness must be determined by informed human judgments, which in itself may be erroneous. In particular, the model must find broad trends to be consistent with the larger domain of transparency research. Lacking a supportive context, topic models can possibly have limited value in discovering the unknown.

Another weakness is data validation. Agency descriptions are mainly sourced from Wikipedia.com articles, an online encyclopedia where authors are volunteer editors of the content posted. This content can be changed by anyone at any time where no substantial verification of the sources of information is required. While STM can independently take text that draws out the effects of the author, incentives for publication, and other variables, the text source itself can arguably be not representative as the text can be changed at any time. As for the control variables and my outcome variable, much of this data is dependent on FOIA.gov and what they decide to update. Their data goes as far back as 2008, however, transparency research had exceptionally peaked in the early 2000s, which is a possible time range route of interest for studying FOIA requests (Cucciniello, 2010)

Aside from extensive analysis of the previously mentioned factors that may influence topical predictions of transparency, future research could identify the main stakeholders who influence the disclosure of information. Agencies may often be caught in between working in the best interest of being politically insulated or being accountable under larger governance and to the public. Additional research is needed to determine how STM can be used to predict the transparency of agency types spanning an arbitrary time range. But this article offers a first step in achieving a framework to design the complex model previously mentioned and my findings show that topical predictive power still remains even holding for proactive disclosures, agency size, and net costs.

Conclusion

In this paper, I examine the wide variations as to what may determine increased access to government information and analyze the possible reasons U.S. federal agencies may want to disclose more records. To capture the various reasons and conflicting evidence of past findings, I take a broader approach by using STM to create my independent variables: topics of federal agency descriptions. Scraping text data from Wikipedia.com descriptions as well as official government websites, the statistical software model 20 topics that establish what types of agencies are extant based on data from the corpus. I then compare how these topics differ in how agencies of these types respond to FOIA requests as my measure of transparency. To provide context for this model, I take into account proactive disclosures, the size of each agency, and the net costs of these requests. Ultimately, I find that topics are highly predictive of transparency, controlling for the variables mentioned. Topics that involve more public-facing agencies appear more transparent than those designed for different areas of government such as economics or foreign policy. Topics on veterans, education, and land conservation are found to be the least transparent of all.

These results are significant in a number of ways. First, my findings suggest that the type of work and functions that an agency carries out, the institutional legacy, and administrative design of particular types of agencies, is important to understand how responsive they are to FOIA requests. Second, my approach demonstrates the utility of STMs in government transparency through agency descriptions. Future research should build on the approaches taken here to further study the connections between agency function, structure, and legacy and access to information. Given the insight from the presented model, understanding the conditions and

contextual backgrounds of these topics has given way to better understanding the theory behind predictors of transparency in government.

Works Cited

Aksu, M., & Kosedag, A. (2006). Transparency and disclosure scores and their determinants in the Istanbul Stock Exchange. Corporate Governance: An International Review, 14(4), 277-296.

Alt, J. E., & Lassen, D. D. (2006). Fiscal transparency, political parties, and debt in OECD countries. European Economic Review, 50(6), 1403-1439.

Alcaraz-Quiles, F. J., Navarro-Galera, A., & Ortiz-Rodriguez, D. (2015). Factors determining online sustainability reporting by local governments. International Review of Administrative Sciences, 81(1), 79-109.

Bureau of Economic Analysis. (n.d.). Retrieved from http://www.bea.gov/

Benito , Bernardino , and Francisco Bastida . 2009 . Budget Transparency, Fiscal Performance, and Political Turnout: An International

Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

Chaffetz, Jason. (2016) U.S. House of Representatives Committee on Oversight and Government Reform, 2016. FOIA Is Broken: A Report. Staff Report 114th Congress. [online] Available at: https://republicans-oversight.house.gov/wp-content/uploads/2016/01/FINAL-FOIA-Report-January-2016.pdf

Cucciniello, M., Porumbescu, G. A., & Grimmelikhuijsen, S. (2017). 25 years of transparency research: Evidence and future directions. Public Administration Review, 77(1), 32-44.

Banisar, D. (2006). Freedom of information around the world 2006: A global survey of access to government information laws. Privacy International.

Berliner, D. (2014). The political origins of transparency. The journal of Politics, 76(2), 479-491.

Berry, C. R., & Gersen, J. E. (2016). Agency design and political control. Yale LJ, 126, 1002.

Birkinshaw, P. (2006). Freedom of information and openness: Fundamental human rights. Admin. L. Rev., 58, 177.

Bridis, T. (2018, March 12). US sets new record for censoring, withholding gov't files. Retrieved from https://apnews.com/714791d91d7944e49a284a51fab65b85/US-sets-new-record-for-censoring,-withholding-gov't-files

Cate, F. H., Fields, A., & McBain, J. K. (1994). The right to privacy and the public's right to know: The "central purpose" of the Freedom of Information Act. Administrative Law Review, 46(1), 41–74.

Carroll, Jean (1967). "A Note on Departmental Autonomy and Innovation in Medical Schools," Journal of Business, 49 (March), 531-534.

Coglianese, C. (2009). The transparency president? The Obama administration and open government. Governance, 22(4), 529-544.

Conniff, R. (2019). Why Isn't Publicly Funded Conservation on Private Land More Accountable? Retrieved from https://e360.yale.edu/features/why-isnt-publicly-funded-conservation-on-private-land-more-accountable

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. Poetics, 41(6), 570-606.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. Academy of management review, 14(1), 57-74.

Exec. Order No. 13,292, § 1.1 (c), 68 Fed. Reg. 15315 (Mar. 25, 2003).

Federal Deposit Insurance Corporation. (n.d.). Retrieved from https://www.fdic.gov/transparency/

Florini, Ann. (2007). The Right to Know: Transparency for an Open World. New York: Columbia University Press.

Fung, A. (2013). Infotopia: Unleashing the democratic power of transparency. Politics & Society, 41(2), 183-212.

Grimmelikhuijsen, S. (2009). Do transparent government agencies strengthen trust?. Information Polity, 14(3), 173-186.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Political analysis, 21(3), 267-297.

Geraats, P. M. (2005). Transparency and reputation: The publication of central bank forecasts. Topics in Macroeconomics, 5(1).

Hameed, F. (2005). Fiscal transparency and economic outcomes. IMF Working Paper No. 05/225. Washington, DC: International Monetary Fund

Herz, M. (2008). Law lags behind: FOIA and affirmative disclosure of information. Cardozo Pub. L. Pol'y & Ethics J., 7, 577.

Hollyer, J. R., Rosendorff, B. P., Vreeland, J. R. (2011). Democracy and transparency. The Journal of Politics, 73, 1191-1205.

Khazanchi, D. (1995). Unethical behavior in information systems: The gender factor. Journal of Business Ethics, 14, 741-749.

Kirtley, Jane E. (2006). Transparency and Accountability in a Time of Terror: The Bush Administration's Assault on Freedom of Information, Communication Law and Policy, 11:4, 479-509, DOI: 10.1207/s15326926clp1104_2

Lane, J. E. (2005). Public administration and public management: The principal-agent perspective. London, England: Taylor & Francis.

McCubbins, Matthew D., Roger G. Noll & Barry R. Weingast, Administrative Procedures as Instruments of Political Control, 3 J.L. Econ. & Org. 243 (1987) [hereinafter McCubbins et al., Administrative Procedures]; Mathew D. McCubbins, Roger G. Noll & Barry R. Weingast, Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies, 75 Va. L. Rev. 431 (1989) [hereinafter McCubbins et al., Structure and Process].

Mishler, A., Crabb, E. S., Paletz, S., Hefright, B., & Golonka, E. (2015, August). Using structural topic modeling to detect events and cluster Twitter users in the Ukrainian crisis. In International Conference on Human-Computer Interaction (pp. 639-644). Springer, Cham.

Moffitt, S. L. (2010). Promoting agency reputation through public advice: Advisory committee use in the FDA. The Journal of Politics, 72(3), 880-893.

Moulton, S. (2015). Making the Grade Access to Information Scorecard 2015. Retrieved from https://www.foreffectivegov.org/sites/default/files/info/access-to-information-scorecard-2015.pdf

Munksgaard, R., & Demant, J. (2016). Mixing politics and crime–The prevalence and decline of political discourse on the cryptomarket. International Journal of Drug Policy, 35, 77-83.

Office of Government Information Services (OGIS). (2019, September 25). Review of U.S. Department of Education Freedom of Information Act (FOIA) Program . Retrieved from https://www.archives.gov/files/ogis/assets/education-compliance-assessment.pdf

Pack, B. (2004). FOIA Frustration: Access to Government Records Under the Bush Administration. Ariz. L. Rev., 46, 815.

Patten, D. M. (1992). Intra-industry environmental disclosures in response to the Alaskan oil spill: A note on legitimacy theory. Accounting, Organizations and Society, 17, 471-475.

Pina, V., Torres, L., Royo, S. (2010). Is E-government promoting convergence towards more accountable local governments? International Public Management Journal, 13, 350-380.

Posen, A. S. (2003). Six practical views of central bank transparency. Central Banking, Monetary
Theory and Practice: Essays in Honour of Charles Goodhart, 1, 53-172.

Ramage, D., Dumais, S., & Liebling, D. (2010, May). Characterizing microblogs with topic models. In Fourth international AAAI conference on weblogs and social media.

Roberts, M. E., Stewart, B. M., & Tingley, D. (2014). stm: R package for structural topic models. Journal of Statistical Software, 10(2), 1-40.

Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. Journal of the American Statistical Association, 111(515), 988-1003.

Roberts, A. S. (2012). Transparency in troubled times. Tenth World Conference of the

International Ombudsman Institute, Suffolk University Law School. Research A.J. Wagner Government Information Quarterly

Rodríguez, R. P., & Rossel, C. (2018). A field experiment on bureaucratic discretionary bias under FOI laws. Government Information Quarterly, 35(3), 418-427.

Rodríguez Bolívar, M. P., Alcaide Muñoz, L., & López Hernández, A. M. (2013). Determinants of financial transparency in government. International public management journal, 16(4), 557-602.

Scalia, A. (1982). The freedom of information act has no clothes. Regulation, 6(pp. 14–19).

Sack v. United States Department of Defense. Decide May 2016. United States Court of Appeals, District of Columbia Circuit.

Sinrod, E. J. (1994). Freedom of information act response deadlines: Bridging the gap between legislative intent and economic reality. The American University Law Review, p 363

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of management review, 20(3), 571-610.

Tejedo-Romero, F., & de Araujo, J. F. F. E. (2018). Determinants of local governments' transparency in times of crisis: evidence from municipality-level panel data. Administration & society, 50(4), 527-554.

United States. Congress. House. Committee on Government Reform. Special Investigations Division, & Waxman, H. A. (2004). Secrecy in the Bush administration (p. 81). US House of Representatives, Committee on Government Reform-Minority Staff, Special Investigations Division.

United States Institute of Peace. (n.d.). Retrieved from https://www.usip.org/

Wagner, A. J. (2017). Essential or extravagant: Considering FOIA budgets, costs and fees. Government Information Quarterly, 34(3), 388-395

Wang, C., & Blei, D. M. (2011, August). Collaborative topic modeling for recommending scientific articles. In Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 448-456)..

Wasike, B. (2016). FoIA in the age of “Open. Gov”: An analysis of the performance of the Freedom of Information Act under the Obama and Bush administrations. Government information quarterly, 33(3), 417-426.

Watts, R. L., & Zimmerman, J. L. (1986). Positive accounting theory.